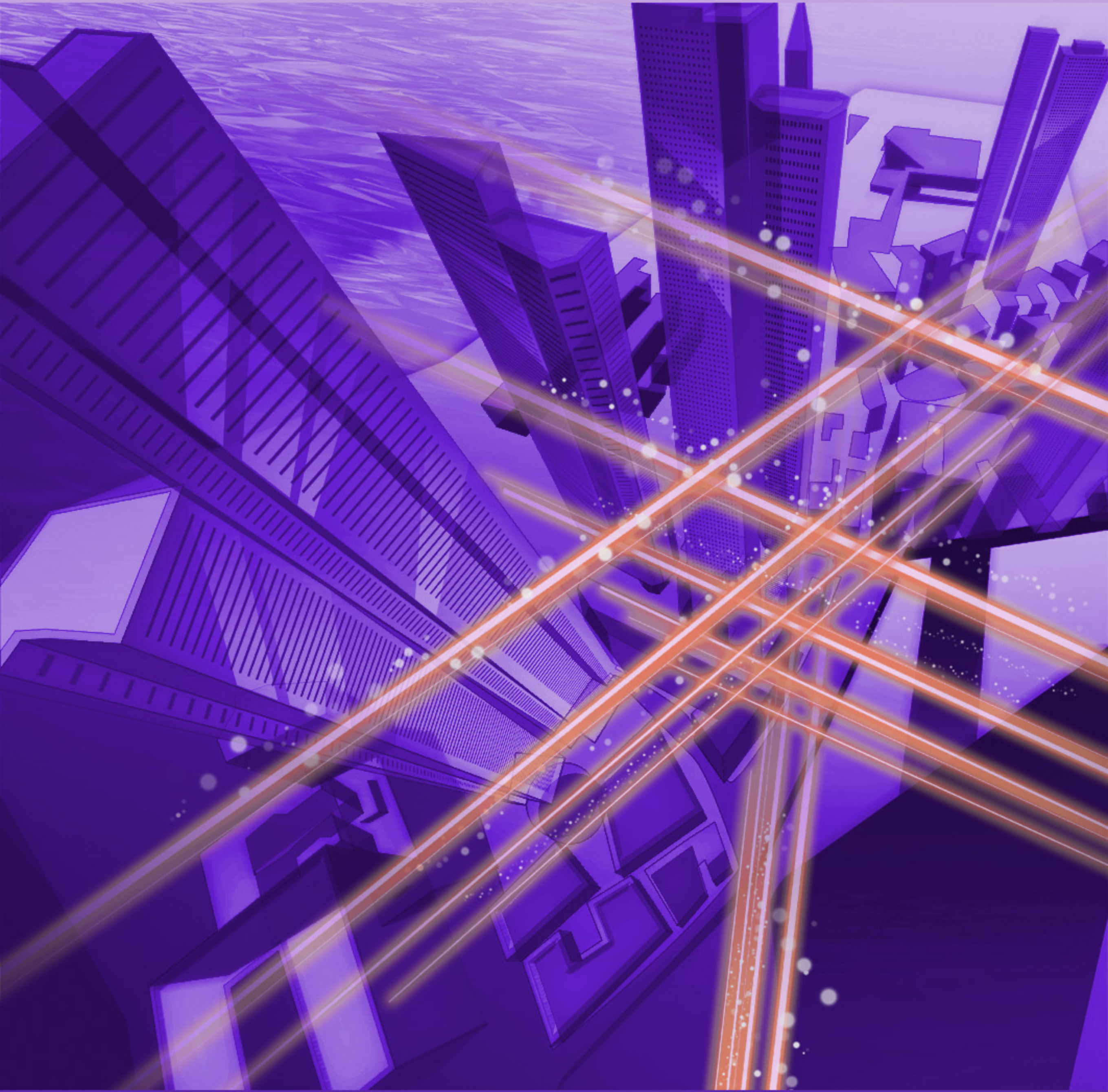
NTT Technical Review

12
2020



December 2020 Vol. 18 No. 12

NTT Technical Review

December 2020 Vol. 18 No. 12

Practical Field Information about Telecommunication Technologies

- Manual for Assisting with Troubleshooting in PHS Digital Cordless Telephone Systems

External Awards/Papers Published in Technical Journals and Conference Proceedings

Start with Yourself, Improve Together, and Contribute to Society

Hidemune Sugahara
Senior Executive Vice President,
NTT Communications

Overview

Digital transformation is accelerating in a society that has changed dramatically due to the spread of the novel coronavirus. NTT Communications, which recently celebrated its 20th anniversary, has reorganized with the aim of providing a safe and secure platform for data utilization to customers around the world. The company is striving to contribute to building a data-driven society on the basis of the concept called 3S (simple, smart, speedy) + S (secure) while developing technology and systems to implement the Innovative Optical and Wireless Network (IOWN). We asked Hidemune Sugahara, senior executive vice president of NTT Communications, about his enthusiasm and philosophy as a member of senior management.

Keywords: digital transformation, data utilization, IOWN

Growing into a company that can deepen and expand its business and contribute to society

—How do you approach business as senior executive vice president?

I would like NTT Communications to become a company that excites our employees, customers, and partners. We would like to use our technology to create value and an exciting future for all stakeholders. We would like to grow into a company that can deepen and expand its business and contribute to society.

In July 2019, the overseas business unit of NTT Communications, Dimension Data, and NTT Security, were integrated to become NTT Ltd. While seamlessly collaborating with NTT Ltd. inside and outside Japan, NTT Communications has resolved to respond firmly to Japanese customers operating global business.

In Japan, under the circumstances embodied by the keywords *data utilization*, *digitization*, and *data-driven*, the management and business of our customers have entered a period of major transformation, and digital transformation (DX) has been attracting attention. To respond to these trends with the concept called 3S (simple, smart, speedy) + S (secure), we reorganized the company in April 2020 to focus on an industry-specific solution-provision system and platform-service provision system. I am supervising platform service, technological innovation, information security, and DX areas.

—So, what is 3S + S?

The concept of 3S + S was created to accelerate efforts toward the actualization of a *smart world*, and it means delivering simpler and smarter services and solutions than other companies at the customers' desired pace with more secureness. On the basis of

this concept, we are developing our business as *one team* centered on the following four major organizations.

At the Business Solutions Headquarters, we are aiming to create a smart world and support DX of our customers. In particular, through industry-specific consulting-type sales focused on the financial, public, manufacturing, distribution, and foreign-capital/information-technology (IT)-service industries, we provide optimal solutions that match the needs of our customers' business, industry, and society. We have set up the Smart World Promotion Office to envision a new world—focusing on data utilization—together with our customers and partners. We are working to solve social problems by using data from fields such as education and manufacturing. In fiscal 2020, we plan to start full-scale activities with a team of 200 people and gradually expand into new fields. To be a *true partner* who is closest to the customer, the Business Solutions Headquarters is promoting collaboration with customers and other NTT Group companies with the aim of maximizing our capabilities and expanding B2B2X (business-to-business-to-X) business.

The Platform Services Headquarters will add more value to our core competencies of networks and datacenters and will be responsible for developing and providing safe and secure services demanded by the digital society. We will continue to enhance functions, such as data collection, connection, storage, and analysis, on our Smart Data Platform (SDPF). In particular, we are developing access network-related services such as the introduction of a zero-trust Secure Access Service Edge model and local fifth-generation mobile communication networks (5G). Moreover, we plan to create a new service that offers network operations while strengthening security of IT and operational technology (OT) and to provide these as one-stop managed services. To develop services from the perspective of user experience, it is important to have a mechanism to ensure the smart world and SDPF work in unison and to maintain a market-oriented mindset.

At the Innovation Center, we are creating new value and view of the world—based on technology—that are not bound by conventional business domains. The four divisions of technology, namely, strategy, management, technology, and design, are creating and promoting new business through open innovation with an eye toward the future, five to ten years from now. Gathering 200 people, including business planners, digital engineers, and designers as a center of excellence (CoE), the Innovation Center supports the above-mentioned Business Solutions Division and Platform Services Division for proposing and providing services to customers. The Innovation Center is also promoting open innovation to create a not-yet-seen exciting future together with our partners and customers while providing a venue for fostering a co-creation community called C4BASE.

The Digital Reform Promotion Department will provide customers with successful use cases regarding our own DX and expertise in data-driven management. The Innovation Center helps customers develop the data infrastructure, implement data-driven management through our CoE for data science and information security, reform work styles, and improve operational efficiency and productivity. To create a smart world, DX must be promoted throughout society. We believe that showcasing our DX efforts will help drive DX and eventually contribute to society. The spread of the novel coronavirus has changed society completely, and we are all trying to do our best to meet the expectations of the so-called *new normal* society.

Prepare an environment in which employees can do their best

—There is a great responsibility behind the phrase "Let's do our best." What are your thoughts on this phrase?

That's a difficult question. It is easy to say, "Let's do our best." However, if I don't do my best, I can hardly tell others to do their best. Even if a problem is difficult to solve from the standpoint of an employee, it might be possible to solve it from the standpoint of senior management, like myself. We, senior management, should do our best to prepare an environment in which employees can do their best. Last year, we celebrated the 20th anniversary of the company's founding and took that opportunity to advocate "Start with yourself, improve together, and contribute to society." I want to foster an organizational culture that ensures employees can do their best (take on challenges), regardless of their positions.

I also think we should support our customers to take on new challenges. To achieve this, both pioneering efforts and reading technological and societal trends are important; however, we must not be too pioneering nor focus too much on the support side. Although this balancing act is very difficult, we must maintain it. I'm sure we can do it.

In terms of pioneering efforts, researchers at NTT laboratories are conducting basic research and technological development. I would like them to make pioneering efforts while listening to the voices at the front lines to anticipate business trends. If cutting-edge sales and cutting-edge researchers collaborate, we will have no rivals.

I was involved in the satellite-communication business in NTT and have a long-standing relationship with NTT laboratories, with whom I have worked on the development and launch of artificial satellites. I was also involved in the founding of NTT Resonant, at which time, I worked with about 200 people seconded from NTT laboratories. With this experience, I understand how fortunate we are to be able to use the resources of the laboratories in our business. By sharing the responsibility that was previously assigned to each laboratory, we support the commercialization of research results and deliver them to our customers so that researchers can focus on their research. I'd like them to make full use of their talents and create something new. Moreover, our employees assigned to the Innovation Center will probe the research results produced by researchers and set up systems to commercialize them.

—You have high expectations for NTT laboratories, right?

We are grateful for NTT laboratories, which have everything from the world's most-advanced basic research to its application and practical implementation. I think that contributing to society by using their research results and converting them into products and services is our way of responding to those benefits. Going beyond that point, I think it is also

necessary to consider collaboration with customers at the research phase. To that end, I'd like to actively discuss such collaboration with the laboratories. Two of the key technologies that I have recently become interested in are cloud-based 5G and quantum cryptography. Regarding technology being researched at NTT laboratories, we are interested in security technology required for data utilization, especially the zero trust model, and are working on such technology at the Innovation Center.

Our efforts toward the implementation of the Innovative Optical and Wireless Network (IOWN) has also been underway, and we have created an IOWN-promotion project team at the Innovation Center to collaborate with NTT laboratories. In the project, we are focusing on implementing the All-Photonics Network and overlay networking solutions at an early stage. We first expect to apply them to the network in our datacenters. Moreover, we want to collaborate with NTT laboratories to create a showcase using IOWN technology at the NTT Musashino R&D Center in Tokyo. Considering the coming space era, we would like to explore space business in collaboration with NTT Space Environment and Energy Laboratories.

To implement Cognitive Foundation®, a key element of IOWN, it is important to share technology and architecture within the NTT Group. This sharing will help us build a foundation for providing one-stop services to customers. IOWN is a good target for standardization, and I'd like to promote that standardization in cooperation with NTT laboratories.

One step forward, one step ahead, one step outside

—Lastly, would you tell us what you have valued in your work?

"One step forward, one step ahead, one step outside" is my motto. The same phrase is true for individuals and organizations, and the world you see when you go one step forward, ahead, or outside will differ. When you experience that difference, different ideas and actions are born. However, sometimes that step is not taken or cannot be taken. That's why I cherish this phrase.

Looking back on making mistakes and misjudgments, I should have taken a step forward. Such failure may cause problems for others and regret in myself. I sometimes think I was not prepared enough. Some of these problems can be beyond our control. It is necessary to read the trends in society; however, sometimes you can't obtain satisfactory results even if you are fully prepared. That's why I think it's important to believe in yourself and take that step forward.

Even if you make a mistake, you can always learn from it. It is important to think about how to learn from the experience and better prepare for your next step. Regarding a recent inconvenience caused to our customers and other stakeholders due to unauthorized access, I regret that it occurred because we were convinced that we had taken all possible measures. From investigating the cause to dealing with it, however, many issues concerning this incident were addressed by the considerable efforts of our employees, reaffirming their high skill level. Taking advantage of the lessons learned at that time, we would like to support our customers through our skilled employees—who are our assets—and grow into a company that excites our customers and partners. To that end, I want us to be an attractive company where employees can play an active and exciting role.

Interviewee profile

■ Career highlights

Hidemune Sugahara joined Nippon Telegraph and Telephone Public Corporation (now NTT) in 1987. He became senior vice president of NTT Communications in 2016, president & CEO of NTT Com Solutions in 2018, and executive vice president of NTT Communications in 2019. He has been in his current position since June 2020.

Unexpected Experimental Results against Common Beliefs Bring Opportunities for New Discoveries

Hideki Yamamoto
Senior Distinguished Researcher,
NTT Basic Research Laboratories

Overview

Hideki Yamamoto, Senior Distinguished Researcher at NTT Basic Research Laboratories, and his co-researchers are creating novel materials nonexistent in the natural world and elucidating the property of such materials. They apply NTT's unique and state-of-the-art apparatus and technologies to stack the elements making up a material through control of their reaction at the atomic and molecular levels. We asked Dr. Yamamoto about his current research activities and the mindset he adopts as a researcher.

Keywords: superconductor, magnetic material, thin-film synthesis, molecular beam epitaxy

Synthesis of novel materials and discovery of new properties through original and unique methods

—Dr. Yamamoto, please tell us about your current research activities.

My research theme is the design and creation of novel superconducting and magnetic materials through thin-film synthesis methods and explanation of their properties (**Fig. 1**) [1]. The ultimate goals of this research are to contribute to the (i) development of lossless power transmission/supply and wiring, (ii) further reduction in power consumption in devices, and (iii) shift to green-power generation and power storage.

To be more specific, I am involved in creating novel materials not seen in the natural world by growing thin films with a thickness from 0.1 nm (atomic monolayer) to 1 µm on a base of a single crystal (a crystalized solid with regularly arranged atoms) called a substrate. I am also involved in accounting for various, and sometimes novel, properties emerging in such specimens. We use molecular beam epitaxy (MBE) for growing these thin films. With this method, we use ultrahigh-vacuum (UHV) chambers (having a vacuum of one ten-trillionth that of ambient pressure), in which each constituent element of the designated compound is supplied in the form of atoms or molecules to give rise to reaction on heated substrates leading to the formation of thin films [2].

I believe that we are an extremely advanced research team on a global basis working on the creation of completely new compounds by using MBE. Specifically, we are conducting research on thin-film growth of oxides that contain two or more cations (complex oxides) by using a method called oxide

(1) Chemical reactions under low-reaction temperatures and non-equilibrium conditions

Reactions between ultimately small particles such as atoms, molecules, and ions

(2) Stabilization of metastable phases through epitaxy (mutual interaction with the underlying crystal substrate)

(3) Uniform oxidation in the case of oxides (strong oxidation possible through the use of O_3 (ozone) and O)

Tenuity and large surface area to volume ratio of specimens

(4) Impurity-free synthesizing environment (ultrahigh vacuum)

(5) High-throughput syntheses

(6) Resource-saving (usage of small amount of raw material)

(7) Specimens in the form of single-crystal thin films, highly compatible with future device fabrication

Fig. 1. Advantages of searching for and synthesizing new materials by using MBE—a method for growing high-quality thin films of known materials.

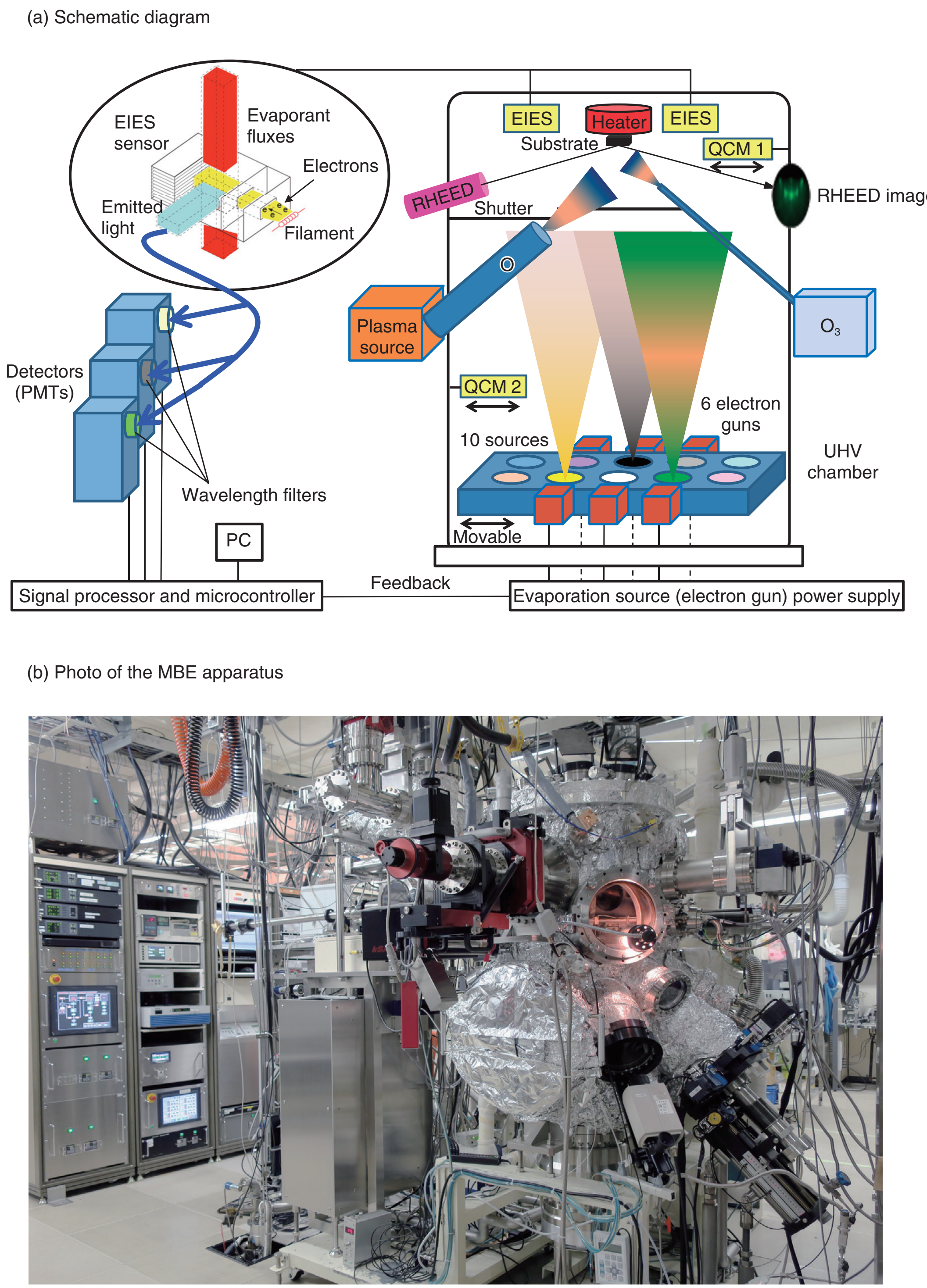
MBE. Our custom-made MBE apparatus *sui generis* (**Fig. 2**) has two key features. First, each atomic/molecular flux of constituent cations can be supplied in a stable manner for a prolonged time by monitoring the flux rates of those elements in real time and feeding the results back to the power supply of the evaporation source. Second, it is capable of strong oxidation in a vacuum by supplying oxygen gas, which typically exists as O_2 (oxygen molecules), in the form of highly reactive atomic O (oxygen) or O_3 (ozone) gas. By taking advantage of these features, we have been able to synthesize crystals not found in the natural world.

Using this apparatus, we have been involved for some time in creating new superconducting materials and elucidating their properties. In superconducting materials, electrical resistance becomes zero under certain conditions. In particular, superconductivity occurs at low temperatures below a critical temperature (superconducting transition temperature: T_c). In this state of zero electrical resistance, electricity flows with no power loss within the superconducting material. The highest T_c under ambient pressure (1 atmosphere) is currently –140ºC, which is about 60ºC lower than the temperature at which dry ice sublimes from a solid (approximately –79ºC). However, there have recently been a series of reports on materials having a T_c of –70ºC or even –25ºC closer to room temperature while they can be stable only under ultrahigh pressures (approximately 2 million atmospheres). These reports have led many researchers to believe that room-temperature superconductors must exist but simply have not yet been discovered and synthesized. A report on room-temperature superconductivity in a carbonaceous sulfur hydride, though under 2.7 million atmospheres, has just been published (October 2020) [3]. It strongly suggests that room-temperature superconductivity is no longer an impossible dream.

We, of course, aim to synthesize and discover such materials, but as a prior step, we have been using the thin-film growth methods that I just mentioned to create new superconducting materials having a variety of properties. Examples include the discovery of diverse phenomena such as induction of superconductivity in materials previously thought to be insulators and strain-induced increase in T_c. More recently, we have taken up the challenge of inducing superconductivity in an artificial structure (artificial superlattice) consisting of alternating layers of the simplest building block of cuprate superconductors and oxides containing no copper (Cu); note that the cuprate superconductor family exhibit the highest T_c under ambient pressure. We are now only one step away from superconductivity in such artificial superlattices.

While known as a ferromagnetic metal and neither a new material nor superconductor, we most recently prepared the world's highest quality $SrRuO_3$ (Sr: strontium, Ru: ruthenium) thin film and observed the emergence of an exotic state called a magnetic Weyl semimetal* [4], the quest of which had so far been

* Magnetic Weyl semimetal: A special state that has only recently been found to occur in some magnetic materials. Weyl quasi-particles that emerge in this state behave as if they have no mass.



EIES: electron impact emission spectroscopy
PC: personal computer
PMT: photomultiplier tube
QCM: quartz crystal microbalance
RHEED: reflection high-energy electron diffraction

Fig. 2. Our MBE apparatus for fabricating complex-oxide thin films.

hampered due to difficulties in preparing high-quality specimens. This accomplishment shows the effectiveness of our high-level thin-film growth technology and measurement techniques for electrical properties under magnetic fields. In future research, we hope to conduct measurements with which the electronic state can be visualized so that the existence of this exotic state can be recognized at a glance.

—Your team is achieving world-class results based on an original method.

The creation of new substances and materials has helped to advance the natural sciences while simultaneously contributing to the development of high-performance and highly functional devices and breakthroughs in device design. This is clearly shown by the discovery of superconductivity in ceramic materials (1987 Nobel Prize in Physics), discovery of fullerenes (molecules consisting of 60 carbon atoms in a soccer ball structure) (1996 Nobel Prize in Chemistry), fabrication of graphene (ultimately thin graphite having a thickness of a single layer of atoms) (2010 Nobel Prize in Physics), and invention of blue light-emitting diodes using a nitride semiconductor (2014 Nobel Prize in Physics). There have also been a variety of proposals and initiatives, in which elements and compounds not previously used will be exploited, for overcoming the limits of miniaturization (the limit of Moore's law) in Si (silicon) integrated circuits, the basis for modern electronics.

Against this background, we have been pursuing the creation of new materials using oxide MBE and successfully created not only superconducting materials but also new magnetic materials. A magnet is a substance whose magnetic property (ferromagnetism) weakens as the temperature increases until eventually disappearing at a certain temperature (Curie temperature). Accordingly, magnetic materials having higher Curie temperature allow for a magnet that can be operated at higher temperatures. With this in mind, we conducted a materials search, which has evolved into the synthesis and discovery of the new ferromagnetic material Sr_3OsO_6 (Os: osmium) [5]. This novel material exhibits the highest Curie temperature (above 780°C) among insulating materials, smashing the record for the first time in 88 years. In addition, Sr_3OsO_6 is free from iron or cobalt unlike most magnetic materials in existence today, which blazes a new trail in the search for magnetic substances. Together with further new materials which, I believe, will be found along the guidelines obtained through the discovery of Sr_3OsO_6, applying these materials to magnetic/spintronic devices that can be operated stably above room temperature will become possible.

—Please tell us how you got started in your world-class research.

In 1987, my former supervisor in this research field moved from Stanford University to NTT Basic Research Laboratories with the aspiration of contributing to both basic and applied research and founded this research unit. About six years later, I entered the laboratories and became involved in this research together with the supervisor and another senior colleague forming a three-person team. At that time, technology for achieving a long-term and stable supply of cation fluxes in a vacuum chamber by providing feedback to the power supply source had not been seen outside of NTT Basic Research Laboratories. I found this to be astonishing, and at the same time, I intuitively felt that this was technology that I should focus on throughout my career. I was fascinated by the possibility of creating novel materials not existing in the world by using this technology.

In 2004, I inherited this research from my former supervisor, and I have since been involved in this work for about 15 years. A variety of experiences have taken place during this time.

We have searched for and synthesized new materials by using the unique thin-film growth techniques that have been cultivated and built up over many years at NTT. However, the manufacturer of the rate controller, which is indispensable for our technologies of high-precision control of the flux rate of each element making up a thin film by electron impact emission spectroscopy (EIES), discontinued the product in 2001. Consequently, thinking that it would be fatal to my research if this technology were to vanish, I consulted with my former supervisor and in the end formulated a specification and ordered a custom-made product from a startup company in Silicon Valley, USA. For about one year starting about three months after that order, I had the good fortune of researching at Stanford University for joint research, which allowed me to visit that company frequently during breaks in my research. Thanks to this state of affairs, I was able to personally participate in the completion of this custom-made product from the operation-verification stage of the first prototype (**Fig. 3**). This technology has further evolved over the 15 years since the development of the prototype, and

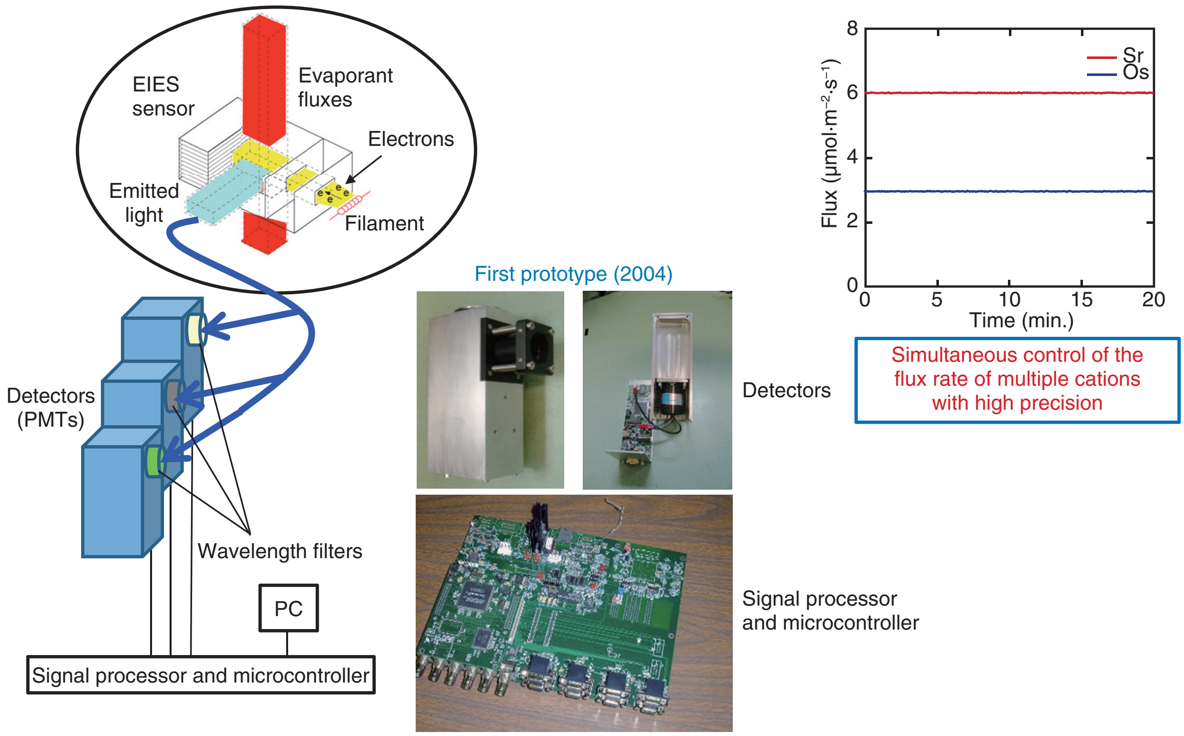


Fig. 3. Flux-rate control equipment using EIES (first prototype of custom-made product).

looking at the way that it has supported our research is deeply moving for me.

When involved in research, there are tough years to contend with, but also exciting years when discoveries, even if small, occur one after another. My first thrilling experience occurred in my third year at NTT. I was able to interpret the experimental results from the research that I had taken up immediately after entering the company and to then publish my first paper on those results. My good fortune then continued with our discovery of the first new superconductor synthesized using the thin-film growth method and other achievements. Our research team was given a boost when my senior colleague received the Young Scientist Presentation Award of the Japan Society of Applied Physics in 1996 and I the same award the following year. On the other hand, I remember well how our supervisor at that time admonished us young researchers excited about our string of successes with the words "perform your experiments simply and honestly." Experimental science is the process of understanding nature by changing experimental conditions in a variety of ways to query nature on its laws and using the answers received. For this reason, the significance of the words "simply and honestly" spoken by him sticks with me to this very day and at every turn.

In subsequent years, more encouragement came my way by receiving the Superconductivity Science and Technology Award (2016) that I had privately longed for and by having the opportunity to publish a number of papers in leading journals. More recently, however, I am frequently called upon to submit scientific manuscripts as the last author (usually, the person in charge of those research projects), so it's more a sense of relief than a feeling of joy when we've gotten a paper published. Within the laboratories, I sometimes hear young researchers talk about their tears of joy over getting a paper published, which I feel a bit envious about. I guess this is a clear sign of my age!

Setting big goals (high mountains to climb) and taking a safe and steady approach

—How have your views of things and role changed after becoming a senior distinguished researcher?

I am grateful for being put into a stable position that allows me to concentrate on my research. I believe I can take appropriate and legitimate risks for making breakthroughs and acquiring high returns by exploiting this position. Taking risks is vital to achieving breakthroughs in basic research, especially in materials research. However, there are many cases in which there is no other choice but to select research themes for which results and papers can be readily produced to secure the next positions of team members and win competitive research funding. The reality is that taking risks is not easy. I therefore want to make good use of my very favorable position and environment as an NTT senior distinguished researcher to pursue research that takes appropriate and legitimate risks.

Among the 11 core technologies for making the NTT vision of the Innovative Optical and Wireless Network (IOWN) a reality, 6 are related in some form to materials research, so in this sense too, I feel the weight of responsibility. Although my theme of design and thin-film synthesis of novel superconductors and magnetic materials with elucidation of the underlying physics is not necessarily a mainstream one in IOWN, there are many examples in history of major breakthroughs coming out of non-mainstream research themes, and I have expectations in this regard.

I also feel that forming research teams that are diverse not only in gender, nationality, and language but also in expertise is important for achieving breakthroughs. It is easy to form and manage a research team whose members are in the same field of expertise, but I think it's difficult for such a team to make major breakthroughs. I think it's great if a young researcher who joined a research team but with a different field of expertise occasionally thinks, "Why don't my seniors know about this?" From a short-term perspective, many points of difference can make things all the more difficult, but from a long-term perspective, the importance of such diversity will be felt.

—Is there anything that you have kept in mind when searching for problems or themes to work on?

Since entering NTT Basic Research Laboratories after receiving my doctorate, setting big goals (high mountains to climb) and taking a safe and steady approach is something that I have always kept in mind. For example, the search for room-temperature superconductors included in my research theme is a high mountain to climb and an appropriate theme for a researcher in my present position. The search for artificial photosynthetic material on par with plants is likewise a high mountain to climb. However, when you're young and have a dream, you might take on a significant challenge from the beginning, and I was the same way. This can lead to a situation of trying to push a massive rock with all one's might and failing to move it even 1 mm. To avoid taking such an all-or-nothing approach, it is advisable to achieve small milestones and write papers on those results as they happen. I learned this approach from my seniors at NTT Basic Research Laboratories.

Basic research does not go well if the theme is one that the researcher does not truly enjoy. It is not always a short run to the top of the mountain, and it is not unusual to have to take detours or pull back as needed. In such circumstances, it is important to adopt a mindset in which you never quit looking up at the top of the mountain. Good basic research can have a major impact on creating or changing a certain concept. In this sense, an example of good basic research that I'm particularly proud of was our research of novel superconducting materials that we synthesized and discovered over about ten years from 2003. This research presented counterexamples to previously established superconductivity emergence conditions in cuprate superconductors. I felt disheartened when our paper was initially not accepted by academic journals, but at least the research was a thrilling experience. Although these materials were established as novel superconductors, a dispute as to whether our interpretation of the superconductivity-emergence mechanism is truly a counterexample to the established theory is continuing. This is also a common occurrence in basic research, so I maintain a positive attitude.

Imagining one's ideal research life given ample funding

—What would you like to say to junior researchers?

Materials informatics has recently appeared as the fourth paradigm of science after experimentation, theory, and computational science. At the same time, the accuracy of predicting the electronic structure within materials through theoretical computations has significantly improved. As a result, it has become possible to conduct materials search in a considerably more efficient manner than before and optimize thin-film growth conditions. I would like junior researchers to use such efficient methods and avoid the detours that I took due to my own lack of ability. However, the reality is that there are still many things you won't understand unless you actually conduct experiments in materials science, so I would like junior researchers to keep in mind the following advice.

Before doing anything else, construct a high-reliability experimental system that can reproduce experimental results. Once you complete this to a reasonable extent, try conducting some experiments. It's important to pursue experiments with good efficiency, but it's also important to try detours without emphasizing only *short-distance runs*. While the reproduction of results is a prerequisite in science, experimental results that differ from what was expected or from commonly accepted theory present a great opportunity. I believe that there's a strong possibility of encountering an unknown or unexpected phenomenon through those experimental results. Moreover, experimental results thought to be a failure in the sense of achieving a certain goal, cannot alone be judged a failure. For example, while it would certainly be disappointing to fail in preparing a thin film with the target level of quality, it is not uncommon for some hints to be hidden in those experimental results.

I sometimes ask young researchers, "What kind of research life would you like to lead and what kind of research would you like to pursue if your pay was enough to lead a worry-free life and you could receive ample funding for research?" The reply to the question should be, "the research that I have a great desire to pursue." In the long run, it would be better to pursue a research theme determined in this manner.

As you know, originality is more important than anything else in basic research. On the other hand, a former supervisor of mine once told me that research on current technology is not something to detest as long as it concerns truly important technology. Avoiding researching current technology and falling to a technological level at which you cannot even reproduce fascinating results that some other research groups have provided is worse than researching it. I believe that you have sound science if multiple top-level research teams and institutions can reproduce good basic research results.

—Dr. Yamamoto, can you tell us about your future research topics?

Until several years ago, the elements that we used to grow thin films of complex oxides included transition metals on the relatively upper rows of the periodic table such as Cu. Recently, however, we have expanded our target to transition metals such as Ru, Pd (palladium), and Os that belong to lower rows than that of Cu on the periodic table, which, for example, has led to the discovery of a material having magnetic properties up to the highest temperatures among insulators. Our next theme is to what extent we can systematically and strategically (but not haphazardly) expand our use of these transition metals. In addition, the discovery of materials that can exist stably only under ultrahigh pressures of 2 million atmospheres but make a superconducting transition at temperatures closer to room temperature has been reported for hydrides, not oxides. Up to now, our target for growing thin films has been essentially limited to complex oxides, so how we should expand our target beyond oxides is also in my things-to-do list.

Tackling these issues is going to take some time—it is not something that can be completed within one's own generation. Training up-and-coming researchers who will become the principal investigators of next-generation research while continuing to produce research results is of prime importance. I feel the need for passing on knowledge as well as technical skills. This is necessary not only for our researchers but also for equipment and component manufacturers that support our research, so I think it would be worthwhile to develop some means of passing on technology on both sides.

References

[1] H. Yamamoto and H. Gotoh, "Overview of Novel Materials Creation Research at NTT," NTT Technical Review, Vol. 17, No. 10, pp. 1–6, 2019.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201910fa1.html

[2] A. Ikeda, Y. Krockenberger, Y. Taniyasu, and H. Yamamoto, "MBE Growth and Element-distinctive Atomic-resolution Characterization of High Temperature Superconductors," NTT Technical Review, Vol. 17, No. 10, pp. 18–23, 2019.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201910fa4.html

[3] E. Snider, N. Dasenbrock-Gammon, R. McBride, M. Debessai, H. Vindana, K. Vencatasamy, K. V. Lawler, A. Salamat, and R. P. Dias, "Room-temperature Superconductivity in a Carbonaceous Sulfur Hydride," Nature, Vol. 586, 373, Oct. 2020.
https://doi.org/10.1038/s41586-020-2801-z

[4] K. Takiguchi, Y. K. Wakabayashi, H. Irie, Y. Krockenberger, T. Otsuka, H. Sawada, S. A. Nikolaev, H. Das, M. Tanaka, Y. Taniyasu, and H. Yamamoto, "Quantum Transport Evidence of Weyl Fermions in an Epitaxial Ferromagnetic Oxide," Nat. Commun., Vol. 11, 4969, Oct. 2020.
https://doi.org/10.1038/s41467-020-18646-8

[5] Y. K. Wakabayashi, Y. Krockenberger, Y. Taniyasu, and H. Yamamoto, "Creation of Novel Material Sr_3OsO_6 with the Highest Ferromagnetic Transition Temperature among Insulators," NTT Technical Review, Vol. 17, No. 10, pp. 7–11, 2019.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201910fa2.html

■ **Interviewee profile**
Hideki Yamamoto

Senior Distinguished Researcher, Supervisor, Executive Manager of Materials Science Laboratory, NTT Basic Research Laboratories.

He received a B.S., M.S., and Ph.D. in chemistry from the University of Tokyo in 1990, 1992, and 1995. He joined NTT in 1995, and his principal research fields are thin-film growth, surface science, and condensed-matter physics. He was a visiting scholar at the Geballe Laboratory for Advanced Materials, Stanford University, USA (2004–2005). He received the 2nd Young Scientist Presentation Award (1997) from the Japan Society of Applied Physics (JSAP) and the 20th Superconductivity Science and Technology Award (2016) from the Forum of Superconductivity Science and Technology, the Society of Non-traditional Technology. He is a member of JSAP, the Physical Society of Japan, the Japan Society of Vacuum and Surface Science, the American Physical Society, and the Materials Research Society.

NTT's Global Business Initiatives

NTT, Inc.

Abstract

The NTT Group positioned its global business as one of its pillars of growth in the NTT Medium-Term Management Strategy announced in 2018, and in FY 2019 revenue grew to US$19.5 billion. After active pursuit of international mergers and acquisitions, approximately one-fourth of the NTT Group's workforce (135,000 people) now works outside Japan, and the global business has grown in importance. The NTT Group's strength lies in its ability to offer full stack support for customer business transformation through the global collaboration of operating companies. This full stack support encompasses everything from networks, datacenters, hybrid clouds, and other information technology infrastructure to management services and business applications.

Keywords: global, smart world, brand

1. Expanding global business

In November 2018, the global holding company NTT, Inc. was established under the NTT holding company's umbrella, and NTT Communications, Dimension Data, NTT DATA, NTT Security, and NTT Innovation Institute Inc. (NTTi3) were moved under NTT, Inc. The aim of this is to improve global governance, rapidly incorporate into NTT, Inc. management the expertise and experience of talent specialized in global markets, and strengthen NTT's global competitive advantage and profitability. In July 2019, NTT Ltd. was established to promote customer business development and the creation as "One NTT" (**Fig. 1**), a global business growth strategy, which combines support for customer business development with innovative creation, to strengthen the competitive strength of NTT's global business (**Fig. 2**).

2. Support for customer business transformation

NTT is using software-defined technology to deploy IT (information technology) as a service in the form of networks, datacenters, and cybersecurity essential to customer business development. In addition to these services, NTT is deploying industry-specific advisories and outcome-based solutions that directly contribute to customer value across a variety of industries. NTT is also cutting total supply costs across the NTT Group, mainly through a global procurement company.

3. Contributing to the creation of a *smart world*

Through these activities, NTT is transforming the NTT Group's homegrown technology and capabilities developed in collaboration with its clients into integrated solutions to contribute to the creation of a *smart world*. In Las Vegas, Nevada, in October 2018, NTT signed a memorandum of understanding regarding the creation of new use cases in transportation, parks, and roadside monitoring for deployment in cities in Nevada, across North America, and throughout the world. NTT's achievements in public safety solutions with advanced situational awareness capabilities in cooperation with the City of Las Vegas was due to the NTT Group's ability to fully use its expertise by drawing on its full IT stack that serves as a one-stop source of innovation and its stance that "customers own their data."

4. Commitment to innovation

NTT has established a global innovation fund, an organization that boosts innovative creation, and an overseas research hub to catalyze its innovative

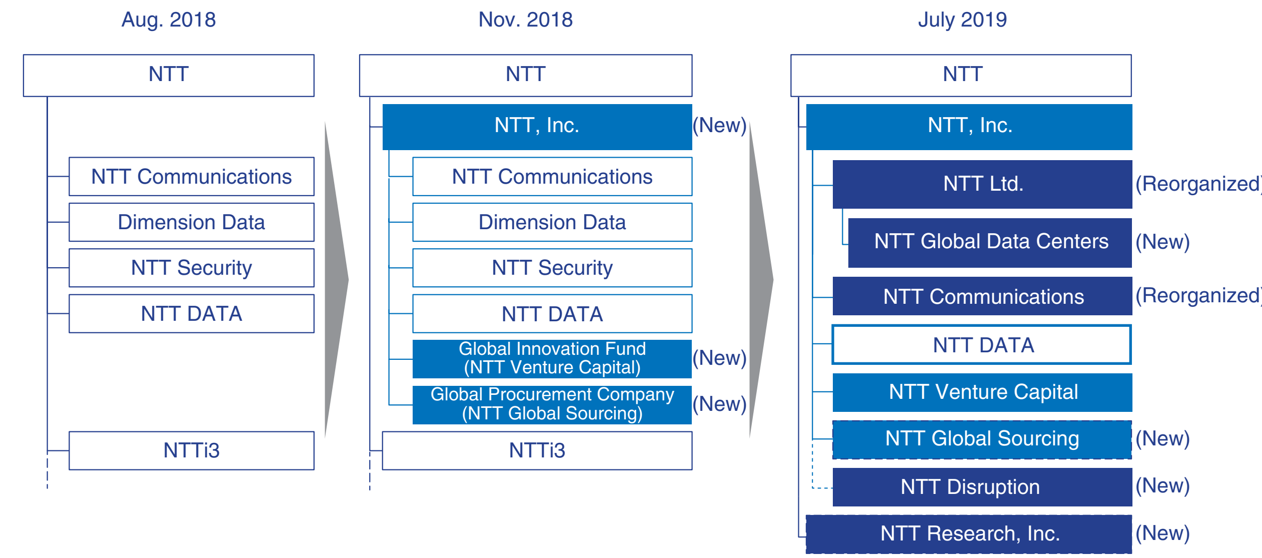


Fig. 1. Reorganization of global business.

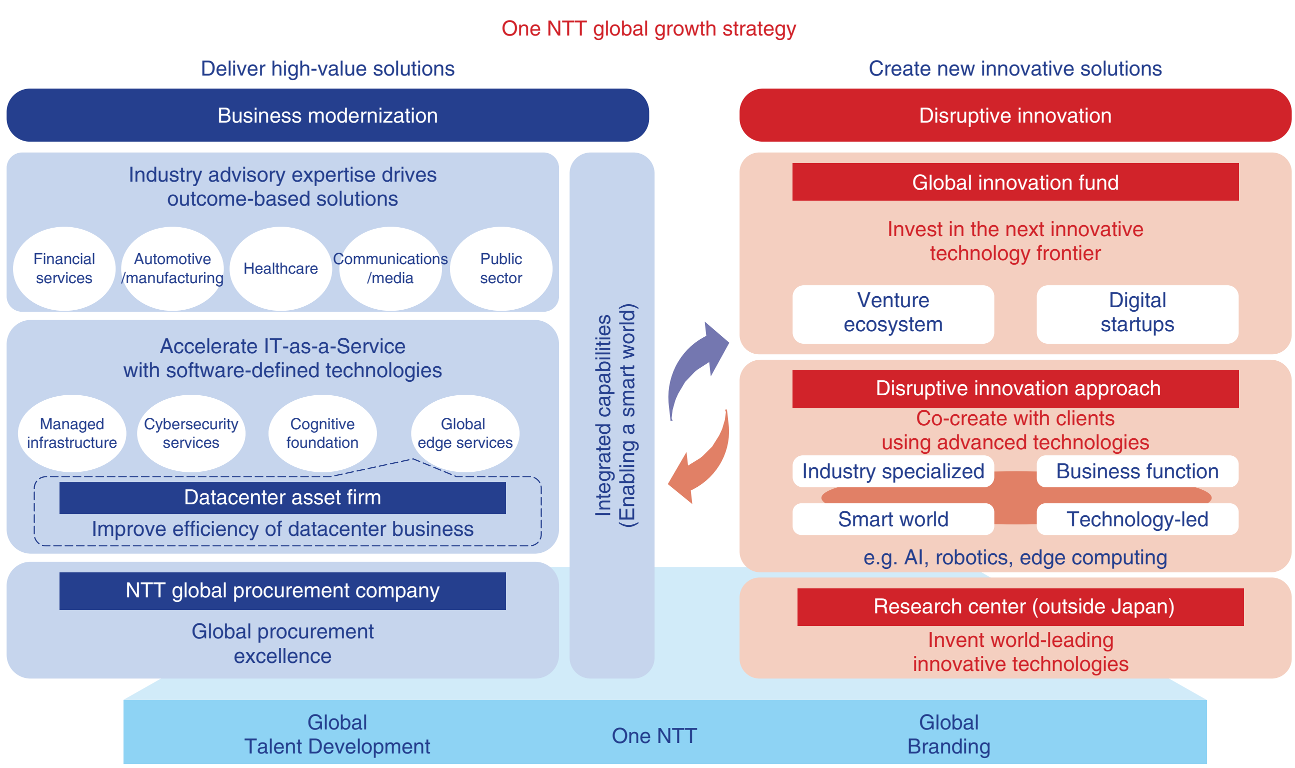


Fig. 2. One NTT global growth strategy.

initiatives aimed at creating something totally new. NTT Venture Capital was established in November 2018 as the NTT Group's global innovation fund to promote synergies and new business creation within the group. Through NTT Venture Capital, NTT is investing primarily in technology sectors that are expected to grow in global markets. NTT Disruption boosts innovative creation, using cutting-edge technology in artificial intelligence, robotics, and edge computing, to create new value in collaboration with clients. In particular, the company aims to commercialize new technologies through proofs of concept that affirm the value of the technologies with customers. Finally, NTT established NTT Research, Inc., an overseas research hub that leads the world in research on innovative technologies.

5. Strengthening the NTT brand

To further enhance its global presence, NTT has become the leading sponsor of the NTT IndyCar Series of top-of-the-line professional motorsports in North America. NTT has also signed a technology partnership agreement with Major League Baseball in the USA and is also an official technology partner of the Tour de France. Through partnerships like the above, NTT is working to provide an innovative experience for sports fans using cutting-edge technology and building the NTT brand through television and the web, enhancing its global competitive advantage.

By combining support for customer business development with its commitment to innovative creation, NTT, with its talents and brand, is enhancing its competitive strength in global business and accelerate growth as One NTT.

A New Training System Using VR Will Rev Up the Baseball Industry Like Never Before

Tomoko Ara and Hitomi Nakamura

Abstract

By accurately reproducing various pitches of actual pitchers in virtual space and visualizing the movements of athletes using a sensing function, the baseball-simulation system that we provide achieves more efficient and effective training for athletes. This system also provides baseball fans with a new means of enjoying the sport by enabling them to *stand*—in virtual space—in the batter's box and play against the pitching of a real pitcher. This system represents a new initiative that has never been available before, and it is generating interest within the industry.

Keywords: virtual reality, baseball, innovation

1. Baseball-simulation system using virtual reality

The baseball simulation system that we provide uses virtual-reality (VR) technology to highly accurately reproduce an actual baseball ground and the pitching of actual pitchers in a 360-degree virtual space by matching the trajectory and speed of the ball. This system was developed using the research results from NTT Media Intelligence Laboratories and NTT Communication Science Laboratories (Sports Brain Science Project) and commercialized after several verification tests.

The key feature of this system is its high reproducibility. It is technically possible to reproduce pitching images with computer graphics (CG). An actual batter standing in the batter's box focuses on the stance or even the interval between the pitcher's movements and uses that knowledge to decide how to swing the bat or determine the type of pitch that will be thrown. The system therefore makes full use of actual throwing images to precisely reproduce positioning and timing when the pitcher releases the ball. By referring to actual pitching data, the system realistically reproduces the trajectory and speed of the pitch that the pitcher threw in a specific inning of a specific game. The same reproduction applies to the baseball ground in the background. By incorporating data taken onsite using a 360-degree camera, the system virtualizes a situation in which the player (batter) can receive the pitch as if he or she were standing on home plate. In an actual baseball ground, the atmosphere in the stadium and the appearance of the pitching differ between night and day games. This system can reproduce such environments in the VR space. By changing the background according to the purpose, it is possible to experience a more realistic space (**Figs. 1** and **2**).

2. Sports × VR

Although it may seem difficult to imagine using VR in the sports industry, VR can be advantageous for sports. Video analysis and the utilization of data are gradually becoming widespread in all sports disciplines. However, in many cases, such data are two-dimensional information, so they are understood using one's intellect. VR makes it possible to *experience* things through the body in a highly immersive three-dimensional space, which may lead to more effective training. It is also possible to practice repeatedly with different ball types and speeds that

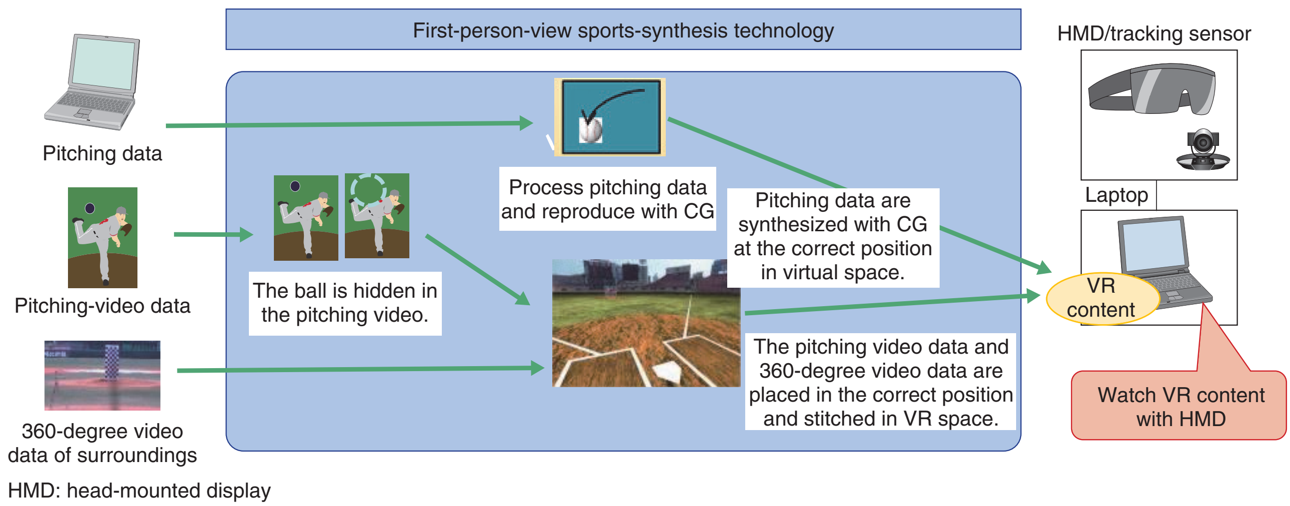


Fig. 1. The baseball-simulation system.

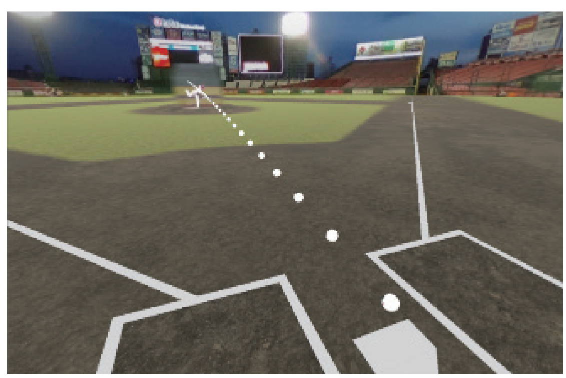


Fig. 2. Reproduction of baseball ground with the baseball-simulation system.

are difficult to reproduce in the real world. Since the necessary equipment is compact, full-scale training is possible regardless of time, place, or weather.

3. Expansion to Japanese and US professional teams

The system is currently being offered for the purpose of training (improving performance of athletes) and entertainment (engaging with fans). There are two use cases regarding training. The first case is *imagery training* before a game. Reproducing, in VR space, the balls thrown by the starting pitcher of the opposing team in the previous game enables the batter to understand the pitching tendency of the pitcher and create a more concrete swing image. As mentioned above, more and more professional baseball teams are checking game videos and pitching data of opponents on tablets and other devices before a game. It can be said that being able to *experience* such data with the body, as opposed to understanding it with the mind only, is very effective just before a game. Although this system is basically for batters, we have had many requests from pitchers. Pitchers only have a limited number of ways to evaluate their pitching in the real world; however, with VR, it is possible to actually stand at the home plate and see what their pitching looks like from the batter's perspective.

The other use case is *visualization* of a player's performance. By monitoring the movement of the player (batter) via sensors while watching the pitch in VR space, it is possible to visualize the movement

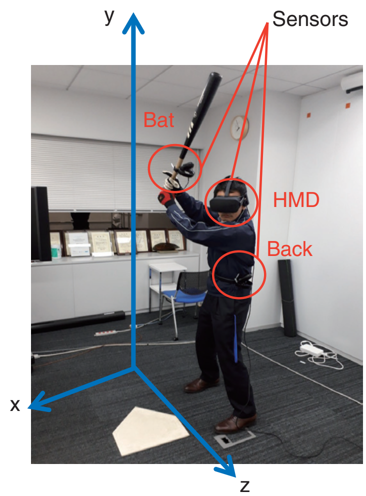


Fig. 3. Swing movement is measured in three directions by using sensors.

and timing of the swing. In addition to the head-mounted display (HMD) worn on the batter's head, sensors are attached to the lower back and bat to monitor the swing movement in three axial directions (X, Y, and Z) (**Fig. 3**). For example, by comparing the pitching timing in the VR space with the movement of the swing, it is possible to determine, for example, delay and hesitation in the swing with objective data (**Fig. 4**). In this manner, it becomes possible to visualize—with objective data—things that are recognized as good or bad, which was previously difficult to verbalize, and athletes and coaches can better embody the training athletes need to overcome their weaknesses. In addition, by monitoring using VR for observing fixed points, it will be possible to determine the effects of training and certain changes.

The system was initially introduced through verification testing to the Rakuten Eagles, a Japanese professional baseball team, in 2017. At that time, the purpose was to conduct imagery training to beat the pitcher of the opposing team before the game. However, we received feedback stating that "As well as watching the pitching, I want to check my swing movement while watching," accordingly, we developed a sensing function for visualizing the batter's swing movement while watching the pitching in VR space and added it to the system. We then focused on introducing the system to Major League Baseball (MLB) in the United States in consideration of MLB's market size and environmental conditions.

The basis of the system for entertainment is the same as that for training, and the aim is to reproduce a ball thrown by an actual pitcher in an actual game with high accuracy. In the real world, it is unlikely a person will ever experience standing in the batter's box in a real baseball stadium and receiving a ball thrown by a professional pitcher in a real game; however, this is possible in VR space. Being neither a batting cage nor a video game, we believe the system will provide a new means to enjoy baseball.

We launched a service for fans of Rakuten Eagles from the 2018 season onwards and a similar service for fans of Hiroshima Toyo Carp in 2019. The system is attracting attention as an effective tool for maintaining the relationship between teams and their fans in circumstances in which it is difficult to directly contact with fans during off-season or due to the current COVID-19 pandemic (**Figs. 5** and **6**).

4. Differences between Japanese and US sports markets

If baseball alone is looked at, we see that the market size in the United States is more than five times larger than in Japan, and the number of teams in MLB is 30 (that in Nippon Professional Baseball (NPB) is 12). Although that fact alone makes MLB valuable as a potential market, environmental conditions have a major impact on introduction of the system. Specialized staff, such as data analysts, are employed in all MLB teams. On top of that, if the system stands a chance of making a team stronger, even if it is a new tool that has never been available in the industry, MLB teams will actively try and find ways to use the

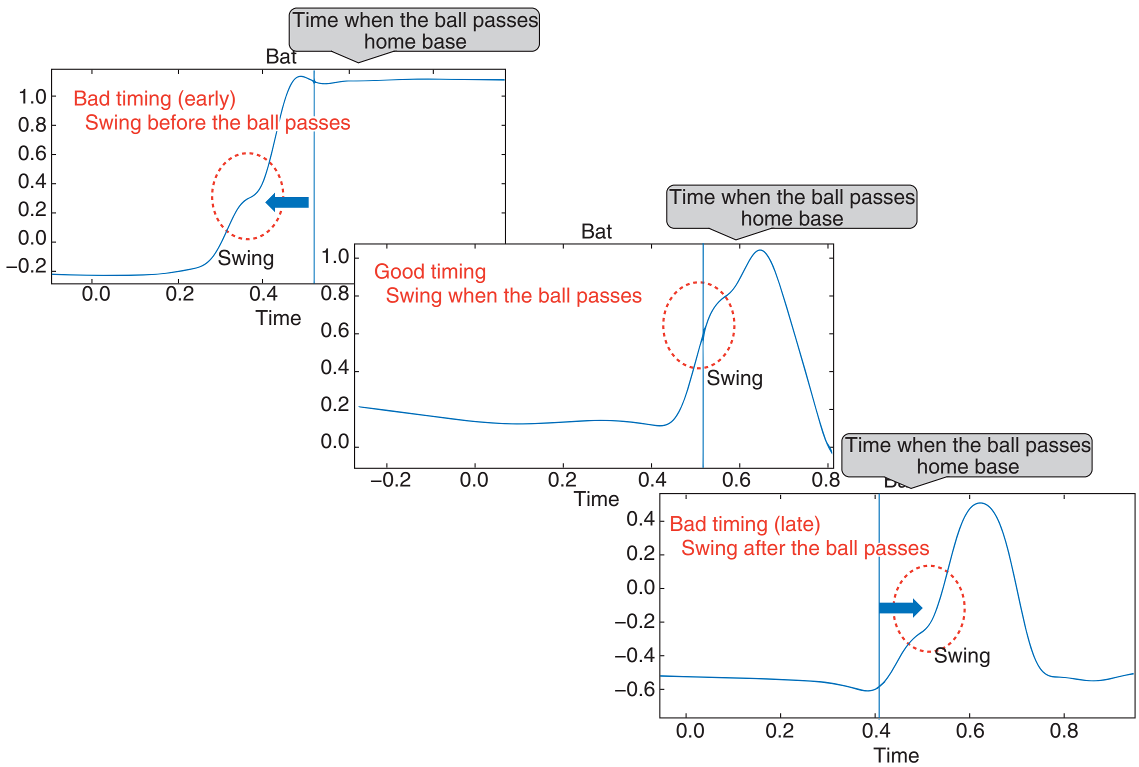


Fig. 4. Sample results.

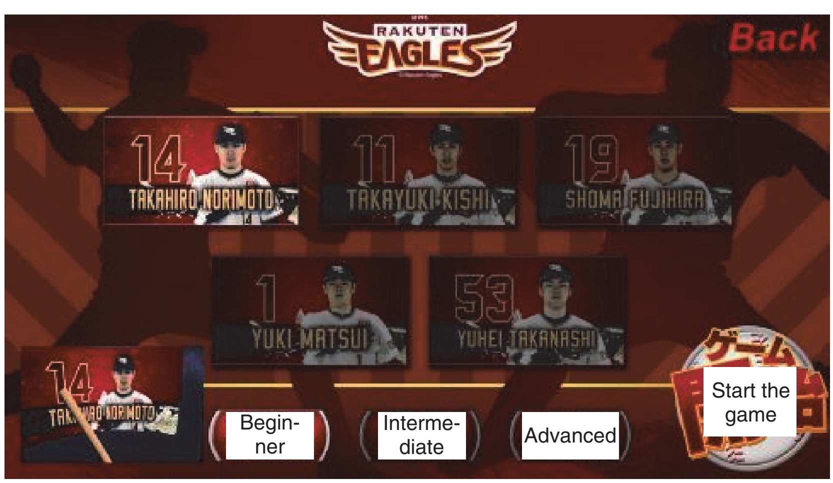


Fig. 5. Top screen of entertainment version.

system with the mindset of first giving it a try. As a result of their efforts, we were also able to compile effective use cases, acquire knowledge, and identify issues. On the basis of the usage record of the system in the United States, we are considering deploying the system to other teams such as more NPB teams.

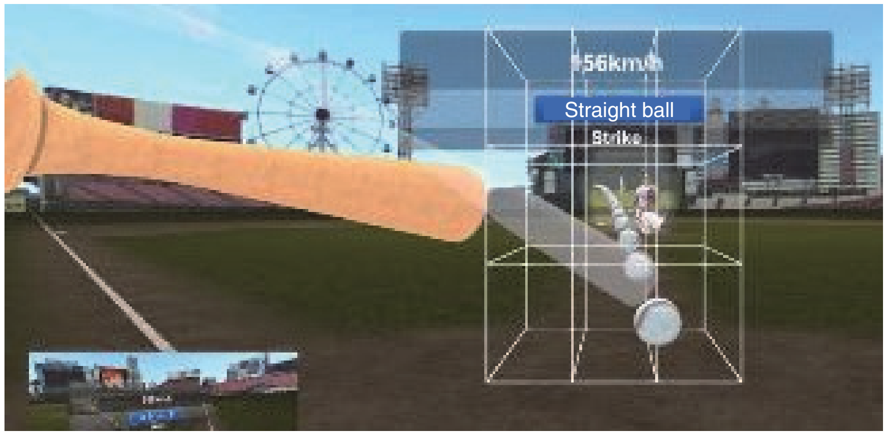


Fig. 6. Illustration of a play in the entertainment version.

5. Future developments: bringing change to the baseball industry

On the basis of our efforts thus far, we are currently redefining the system's effectiveness, simplifying it to expand it to more NPB teams, and building and verifying a new service model. The VR equipment is not manufactured under the assumption that it will be used by sportspeople, so we want to optimize it as the manufacturer improves the VR equipment.

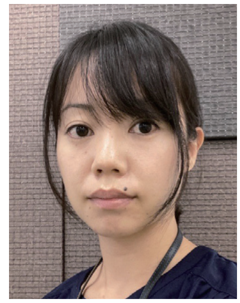
The introduction of the system to amateur baseball teams is also one of our goals. Compared to professional teams, amateur teams have a limited training environment and time. The types and speeds of pitches that can be experienced in actual games are also limited. Training using our system makes it possible to repeatedly practice for ball types and speeds that are difficult to train for in the real world. In the actual verification tests, we heard the comment that players who trained at a ball speed of 150 km/h in VR space began to feel that the ball speeds of 130 km/h and 140 km/h were slow in real practice, and it became easier for them to adjust their timing.

The ultimate goal of this project is to establish a platform that can visualize and manage player performance. The evaluation of a player's performance has traditionally been highly opaque. An evaluation index for a batter is the batting average—the number of hits divided by the number of at-bats—which, however, is a comparison under different conditions such as opponent pitcher, time, and weather. In contrast, the reaction data for pitching in VR space is measured under constant conditions, including ball type and speed, so that data can be used to compare players. By managing the performance data of players from amateurs to NPB and MLB players in the form requested by each user, it is possible to support the development of young players within the team and the formulation of game strategy. It will also be possible to use the system to provide new reference data for scouting and drafting players, which used to be largely dependent on individuals. With these possibilities in mind, we would like to continue discussions with the individuals concerned.

This year's National High School Baseball Championship, which is held in the spring and summer every year in Japan, was canceled due to the COVID-19 pandemic. Thus, it is necessary to establish a new format for the draft that provides amateur players a path to becoming professional players. As can be seen from the introduction of a limited number of pitched balls per pitcher per week at the National High School Baseball Championship, attention is also being paid to the viewpoint of a player's career length. We thus believe that visualizing the player performance and making training more efficient will become even more important. We will contribute to the further development of the baseball industry by using new technologies and ideas regarding VR.

Tomoko Ara
Deputy Manager, NTT DATA Corporation.

She received a B.A. from the Department of Humanities and Social Sciences, Keio University, Kanagawa, in 2004 and an MBA from Waseda University, Tokyo, and University of Washington, USA, in 2016. She joined NTT DATA in 2004. Her main interest is how to create and commercialize new business with technologies developed by NTT R&D. She has been researching immersive technology including VR since 2016 and is currently conducting field experiments of the VR Batting Training system in the Japanese and US markets.



Hitomi Nakamura
Assistant Manager, NTT DATA Corporation.

Navigating Today's Computing Model Landscape

Nadeem Ahmad

Abstract

We are witnessing a shift in the computing model landscape. We are increasingly seeing enterprises move away from on-premise hardware deployments while seeing growth in cloud and hyperscale providers. Enterprises are now dealing with a hybrid computing environment encompassing four different computing architectures and multiple cloud-deployment services. This article recaps a recent review by the Innovation and Technology group at NTT Ltd. on the shift in the computing model landscape and the four main computing architectures being deployed within enterprise environments today. This article also covers where NTT Ltd. have helped clients navigate the decision-making process around deploying the various computing models and finally touches on our focus around providing high-value services to continue enabling and supporting clients with their hybrid computing challenges.

Keywords: cloud, application, data analytics

1. Changing dynamics in the market

The market is going through multiple transitions with a strong drive for enterprise clients to consider and adopt cloud computing models. In the datacenter and more traditional datacenter infrastructure space (servers, storage, and datacenter networking), there is no doubt that the new paradigm will consist of a hybrid computing model, incorporating a combination of on-premise computing resources (as per the classic bare metal servers and infrastructure), private cloud platforms (on-premise and/or hosted), and public cloud platforms.

2. Four computing models are challenging the enterprise application landscape

The market is also seeing new computing architectures emerging in quick succession, creating further challenges for enterprises as to which models and architectures to pursue. There are four distinct computing models evolving at the same time: physical server, virtual server, container-based architecture, and serverless computing. This results in enterprises having to support a hybrid computing environment where the more diverse the mix, the more complex and expensive operations and service continuity become (**Fig. 1**).

In 2019, less than 50% of the physical servers globally have been virtualized, yet we are already seeing a rise in container-based architectures and serverless computing models. Each of these models has a relevant place and fit, and it is highly likely that enterprises will have to consider, plan, and invest in all these architectures. The answer as to which computing model or computing architecture to pursue is found in the application in question.

Many applications cannot be re-platformed to run on virtual machine (VM) architectures and will remain on a bare metal and dedicated infrastructure, most likely for years to come, as these are core business applications that will have to be rewritten to move to more modern architectures, and in many cases this is not worth the return. There are many applications that have been re-platformed to run on VM architectures, and many of the modern cloud platforms originated from these architectures and continue to grow. The next transition is defined by container-based architectures, and most applications will have to be rewritten to fully leverage container-based

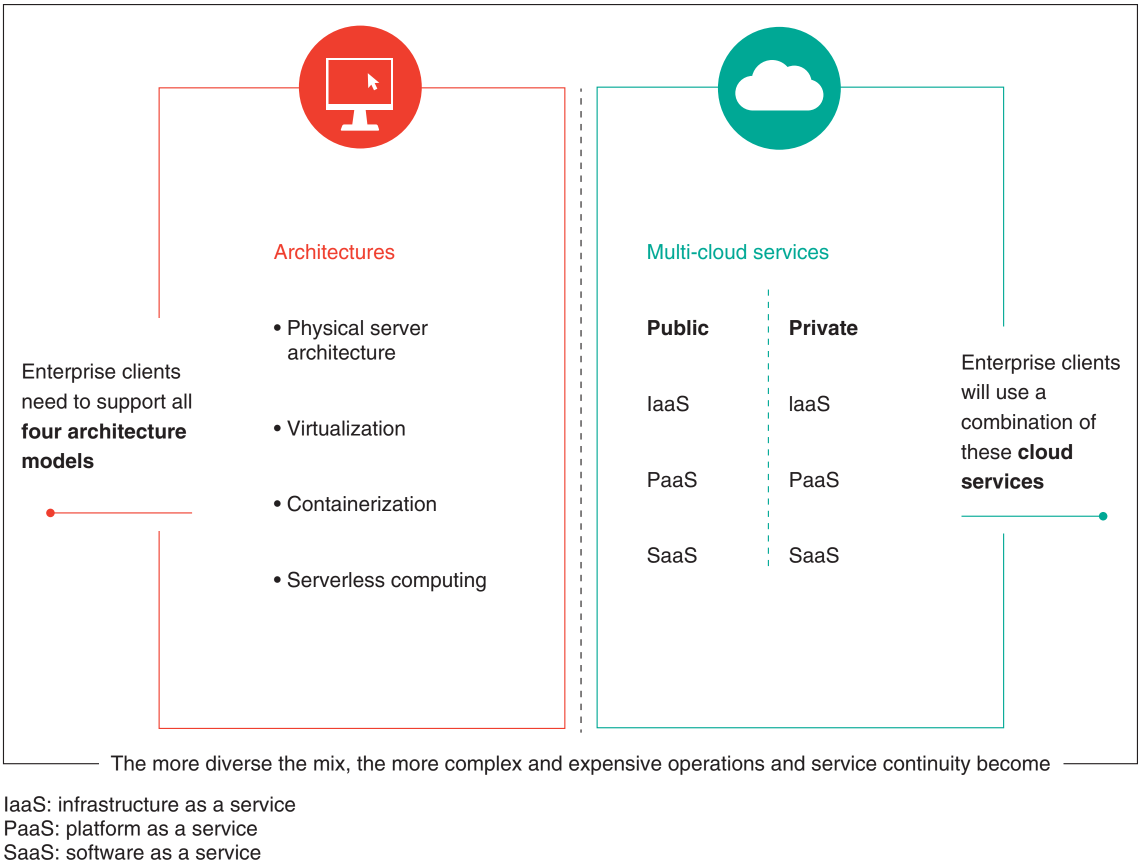


Fig. 1. Hybrid computing environment.

architectures.

At the same time, we are seeing the strong emergence of yet another computing architecture in the form of serverless computing, where value is provided to cloud-based functions, the functions are executed, and value is returned for use. Serverless computing will require further adjustment in computing architecture, especially in the context of data platforms and data abstraction, which implies an entirely different application architecture. Not all applications will be suitable for serverless computing, so the expectation is that certain applications will be suitable or execute optimally on specific computing models and architectures, resulting in a situation in which enterprises will have to invest in all these models and architectures to truly benefit from advances in technology.

This article distinguishes between computing models such as physical/dedicated infrastructure models, cloud-based models, such as infrastructure as a service (IaaS), and computing architectures such as dedicated server, VM, container, and serverless. It should be noted that hybrids occur within and across both computing models and computing architectures, resulting in a complex service challenge for most enterprises. This complexity is also contributing to the desire for an increasing number of enterprises to leverage applications in a software as a service (SaaS) service type to avoid the complexity of the underlying infrastructure or rewriting applications. This is not always possible, and an increasing complex hybrid is a reality.

3. Where are all the workloads (and money) going?

Morgan Stanley estimates 44% of computing workloads will be orchestrated in the cloud by the end of 2021, up from 21% in 2018 [1]. Meanwhile, every 1 USD of revenue growth for the largest cloud service providers (i.e. hyperscalers) has resulted in about 3 USD of revenue decline for the major legacy non-cloud infrastructure-technology providers. Data from other industry analysts and views from key market players support this trend. It is clear there is a market shift coming across three critical areas:

- The shift in information technology (IT) spending growth focused on digital transformation
- The change in the mix of enterprise workloads by an environment showing non-cloud shrinking
- The rapid growth and adoption of software-defined infrastructure

There are three clear leaders in the IaaS and platform as a service (PaaS) hyperscale market, Amazon Web Services (AWS), Microsoft, and Google.

- AWS is still the market leader. The rate of innovation and expansion of their services is significant, and part of their success with Amazon has been to compete where needed with razor thin margins, which could be dangerous as the market matures and competitive pressures rise.
- Microsoft has focused almost every part of its business on winning in IaaS and PaaS as well as making the most use of its SaaS offerings (predominantly Office 365) to drive utilization of Azure and its rich partner ecosystem to drive adoption and utilization. Microsoft also have the largest incentive in this group to drive true richness in hybridcloud linking to their partner ecosystem and massive presence and breadth on-premise at the enterprise.
- Google is a relatively late entrant; however, they are investing heavily in enterprise, have significant capacity worldwide, strength in high growth and consumption areas of data and artificial intelligence, and machine learning, and clear ambition to win.

Though hyperscale cloud providers have seen tremendous growth, they are not satisfied and continue to expand the number, variety, and scale of the services that they offer to customers to meet their demand. The number of AWS services rose from 1 in 2006 to over 140 services by 2018. To us, this represents a shift away from traditional servers and buying patterns for some legacy applications and certainly for new initiatives. The applications and scale points to the fact that AWS' (and other hyperscalers') support for migration to cloud continues to expand, supporting larger instances, more dedicated configurations with the same programmability as their heritage, and customer demand. This is reflected in the growth in their revenue.

Today, enterprises are facing challenges within and competition from external and new entrants looking to disrupt their business and even their entire industry. As a result, digital transformation and speed of innovation is top priority for both business and technology leaders. Increasingly, the line between IaaS and PaaS is blurring, and the speed and efficiency to consume one from the other is driving the shift to hyperscale providers. Cloud applications that are written to take advantage of instant scaling, latest DevOps (development & operations) practices, and toolchains and containers are often public cloud first workloads, skipping the enterprise datacenter all together. While there is a risk of provider lock-in for enterprises, the speed and richness of solutions is incredibly attractive. The stickiness with a single hyperscale provider is high for a solution given that many PaaS components are proprietary and designed to work 'better-together' with other parts of their own ecosystem and partners.

4. The struggle to maintain relevance is real: enter programmability

Considering the complexity and scale of the computing models and architectures and the adjacent infrastructures such as connectivity, the modern workplace, and diverse business models, it goes without saying that there is a major challenge to operate and service these complex environments and remain relevant in the industry, especially in the light of the speed of technology transitions. The industry does not produce enough skilled human resources to perform the operational services to maintain these environments. The problem is multi-facetted in that there is a shortage of skilled human resources, resulting in these environments becoming so complex and dynamic that humans can no longer perform the operations effectively or efficiently. All operational models are now underpinned by a large degree of automation, which in turn requires all these infrastructure and architectural platforms to be fully programmable.

Many infrastructure vendors have been redeveloping their products to be programmable, and one must consider the life cycle of the installed base, legacy of multiple platforms, operating systems, product families, and lack of modern software architectures. There is an increasing demand for all vendor products to be 'API (application programming interface)-first,' implying that products need to be programmable first and foremost and console or command line configurability secondary, if at all.

Companies that have a true API-first model will most often have their console access via the very same APIs, whereas companies that have to transition from a console or command-line-first approach to an API have to do so via complicated stages. Typically, one finds that not all the configurational items are accessible via their APIs, which results in a complex hybrid model. The latter is not desirable as it requires 'old' and 'new' skills and does not support the automation that operational model demand, and all of this becomes costly and lethargic. Considering the current market and key technology vendors, it goes without saying that those who transition to API-first and truly programmable models will flourish, be preferred in the market, and accelerate their growth, and those who do not move swiftly enough will see market share losses and face severe market challenges.

As information and communication technology (ICT) providers consider their objective to be leading IT and services companies for their clients, they will need, if they have not already, to establish a preference for vendor partners that offer API-first technology that is fully programmable. At the same time, they will need to proactively remove technology products from their portfolio that do not provide programmability, since there is no longer a sustainable business model for non-programmable infrastructure for ICT providers or their clients.

5. With no end in sight—maturity and expansion of SaaS continues

The mass adoption of enterprise SaaS continues across a wide range of application categories, and each of these applications previously required servers and storage inside the enterprise and are now consumed as a service. There are SaaS providers in the following major categories:

- Productivity: Office 365, Google G Suite, Atlassian Confluence and JIRA
- File store: OneDrive for Business, Dropbox, Box, Google Drive
- Collaboration: WebEx, Zoom, High Five, Slack, MS Teams, Yammer
- Enterprise resource planning (ERP): SAP Hana & cloud variants, Oracle, Dynamics 365, Intacct, Intuit, NetSuite
- People and payroll: Workday, ADP
- Service/sales/customer relationship management (CRM): ServiceNow, Salesforce.com, Siebel

PaaS and serverless computing are poised to make further inroads into reducing the requirement for enterprise clients to buy servers, storage, and networking infrastructure, not to mention the consulting, professional, and technical services that went into their design, build, and operation.

6. Hybrid computing environment in practice

This concept of four computing architectures is not just theory or an academic exercise. At NTT Ltd., we have practical use cases of implementing a hybrid mix of architectures. However, decision-making for clients is not always centered around the best technical approach. Often, the factors that go into choosing to implement one architecture over another or a mix of architectures are a matter of operational challenges, financial discipline, and risk tolerance. Let us explore one of our practical examples of our multi-architectural approach to meet the needs of one of our more well-known clients.

7. Implementation of four computing architectures for Tour de France

NTT Ltd.'s Advanced Technology Group focus is on using sports as a technology incubator within NTT. They are focusing on the strategic innovation projects that NTT are executing with clients. The Amaury Sports Organization (A.S.O.) organizes the Tour de France, a bicycle race that takes place over 21 days around France involving riders traveling about 200 kilometers every day. Supporting the race is what is commonly referred to as a traveling circus as a large number of support and infrastructure vehicles are required. When we started and first partnered with the A.S.O., there was a real concern around the level of connectivity, e.g., just basic supply of items such as electricity, in many of the remote locations the tour visits.

Our first solution was using an enormous truck, i.e., a traveling datacenter. The truck traveled hundreds of kilometers each day and night to support the race. With this truck, we collected the data from the bicycles and riders, provided a large amount of analytics, and provided the summary and result of data analytics to the broadcasters. Our truck would pull up next to the broadcasters' trucks, and a cable would be plugged in from one truck to the next. We had a physical infrastructure, i.e., physical servers in the truck. This worked, but it was quite challenging as one had to take all the technical staff on the road to ensure everything worked smoothly.

In the 2016 race, we had a stage of the race located

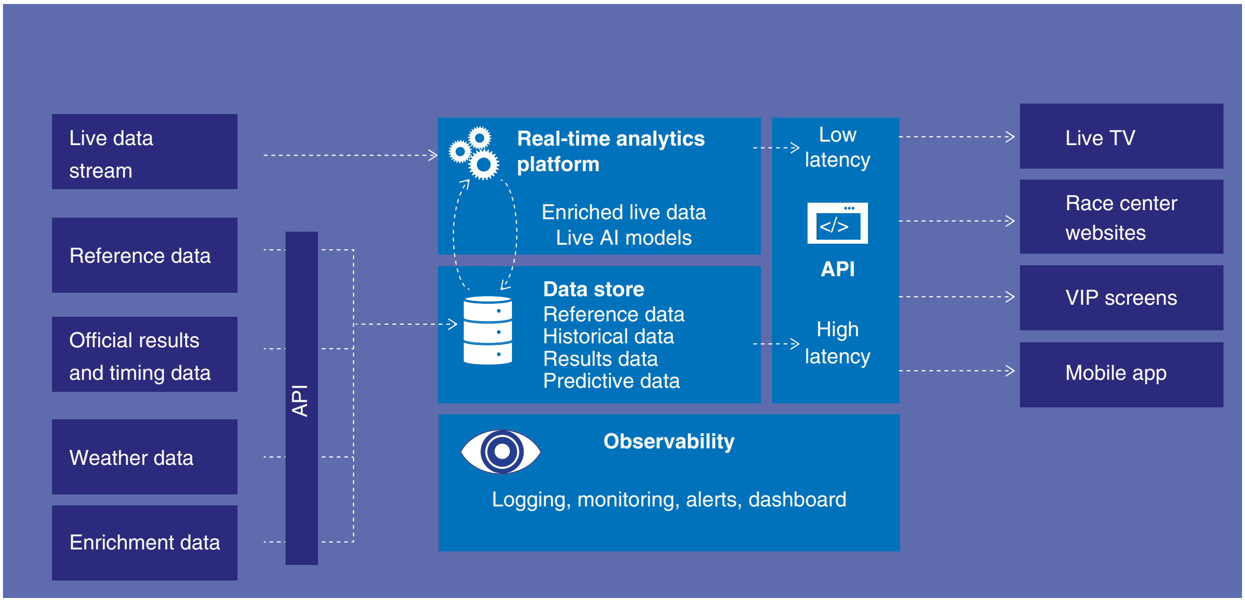


Fig. 2. Data platform.

at a mountain called Mont Ventoux, which is known for having extremely windy conditions and severe storms. During that stage, a severe storm came through. It was so bad that the finish line, which was meant to be at the ski resort at the top of the mountain, had to be changed, and they finished the race only halfway up the mountain.

This does not seem like a significant operational challenge, but all support trucks that were going to be parked in the large parking lot up at the ski resort were distributed at schools, fields, and supermarket parking lots across several villages at the bottom of the mountain, and our truck was no longer parked next to the broadcasters' trucks. In fact, it was two and a half kilometers away.

Therefore, the decision was made at around two o'clock in the morning that we basically had to take everything that we have been running on physical servers and shift to running on VMs in the NTT cloud overnight, and we were actually up and running for the next day. We ran successfully on that model for the rest of the race, and that convinced everyone including our client that the applications and functions running on physical servers could run in a virtual world in a virtualized environment, and we stayed with that virtualized environment until 2018.

Over the last eight months or so, we have been making some very fundamental architectural changes, which are in line with what we have been talking about thus far, i.e., the four computing models. If we look at **Fig. 2**, we see a representation of our data structure, which is the core of our solution. Very broadly, we obtain data from a range of sources on the left side, carry out some data processing and analytics in the middle, and distribute data to a variety of consumers on the right side. The real-time data analytics platform is grouped into four main areas because these areas were essentially built as different platforms and have different drivers in each area.

As an example, we need a robustly managed API. However, we had a hard look at our previously built API and through the course of last year, as much as everyone prefers to do it themselves, we realized that there was actually nothing about our API that was different from anyone else's. There was no additional benefit to building it ourselves, and what we could bring would be only additional cost, complexity, and the management burden across the platform and solution.

Therefore, we ended up going with a set of PaaS services and serverless functions for the API. We used a third party's API, which has the advantage in that we do not need to worry about any of the infrastructure or mechanics of the API infrastructure and simply plug in our functions and transformations we want and structure the data that we want to process.

This has significantly narrowed the scope, focus, and effort of that workstream. It also means that the support requirement on our end significantly decreases. We do not need to worry about servers. We were previously running an API that can scale to support a million users and requires a very significant infrastructure and large amount of looking after, but using a third party's API basically bypassed that whole lot. The last race of the Tour de France in which we used such APIs, was uneventful for the people looking after the API infrastructure because it just worked. At the end of the day, this is the desired result for all IT operational support people both on the client and provider sides.

The real learning has been trying to understand the specific features of the hyperscalers' solution. One needs to get into such specifics, and we found the best way of doing this is running proofs of concept and proving out the technologies because one needs to not only make decisions in terms of technology but also in terms of costing models.

Previously, the approach was just about obtaining the right technologies, then we would run proofs of concept and end up with a choice of technology that supports the desired functionally, however, the cost of the technology solution was quite high due to our large amount of data. Today, with most hyperscalers, if there are multiple overlapping options in many areas, we can switch to another option that is functionally just as effective and have a far better pricing model. This is something to consider, particularly when one starts looking up at the higher end of the stack of solutions from the hyperscalers.

Thus far, we have seen operational challenges move our architecture from physical to virtual. We then saw that financial considerations cause a need to adopt serverless and "as a service" models from hyperscalers. Technologies are not always the deciding factor into which a computing model will work for our client environments.

Looking at Fig. 2 again, the data store is primarily an SQL (structured query language) server; it is just a data store. It has a large amount of data, and we have made the choice, in this case, to stay with the VM-based solution not because there are no other alternatives or because it is the best option. It is just that in our case for 2020, we have several other parts of the solution changing, so we did not want to also change our data store architecture. This represents another item for organizations to consider and is their tolerance of risk. We have deliberately chosen not to do anything with that data store and chose to leave it on a VM because we have enough technology risk elsewhere in the stack and do not need to change it. We will probably change it next year because there are some advantages to shifting to PaaS solutions, but it is not an imperative this year. We now have two computing models incorporated into the environment, i.e., we have serverless functions on the API and VMs on the data store.

The last area is the real-time analytics platform, which is where NTT Ltd. brings a large amount of its intellectual property into the services that we provide to the A.S.O. and broadcasters. It is where we take the sensor data from the bicycles and enrich the speed and location data we obtain from the bikes and apply machine learning to turn a couple of data points into over fifty data points that then support a range of consumers and consumption models. Therefore, as opposed to the above third party's API, it made sense for us to build our own data analytics platform.

However, building our own real-time analytics platform is complex, and we want to be able to continue to implement changes and rapidly evolve that solution. In this case, therefore, we have gone with a containerized solution that is built across a range of Docker containers, which makes deployments much easier. One might ask, "How does running on containers make things easier?" It is because one does not need to deploy an entire VM, operating system, plus everything else. One just deploys the container so what one needs to run on is much smaller and simpler in terms of footprint and resources. We can describe this in code, which means we can say we want a container that is this big and looks like this and has these items in it much more easily in code. This means we can make a deployment directly from code. The code for the infrastructure can be managed, just like application code, and run just like the rest of the application. In fact, it becomes a single package—the container and everything that is in it. We can then make deployments multiple times a day, which means that we can make changes and test the changes rapidly, giving us the velocity that we really want around the development in that space. The above gave a bit of context and insight in the three computing models being implemented. Remember, we moved away from an entirely physical infrastructure to VMs, serverless computing, and container-based architectures.

There is one last point around the complexity of having all three architectures in play. Referring to Fig. 2 once again, there is a small box at the bottom called "Observability." Running hybrid architecture solutions like this introduces much complexity. The challenge is that if one changes one piece of the puzzle, another piece somewhere else can entirely break or malfunction, and it can become too complicated to understand all the ways the system can fail. While one can put in place alerts, one can be guaranteed that what will break is not something one has placed an alert on.

What we have built across the whole solution is an observability platform for visualizing the outputs of all the layers then all the functions of the layers. This allows us to monitor the outputs and understand what functions are working or not working. Complexity is definitely a challenge and is the reason we have implemented that fourth observability component in our solution to address and manage that challenge.

The Tour de France platform we built provides an excellent example and provides interesting insights into the factors considered when implementing a hybrid mix of computing architectures. We started out with a traditional physical infrastructure then, due to operational challenges, we had to pivot to VM implementation. Due to financial considerations, we incorporated serverless computing, and further operational, portability, and efficiency considerations drove the move to a container-based architecture.

The lesson here is not always a matter of the best technology but the best fit of technology driven by business considerations, aside from the technology itself. In the Tour de France example, the team had to address considerations such as operational challenges, financial considerations, security, and of course risk tolerance.

8. Summary and high-value services

Unlike physical computing, virtual and container-based computing make it easy to move applications, as they are decoupled from the underlying computing infrastructure. The use cases for these models are becoming clearer and adoption is on the rise, and there is now the promise of serverless computing, which offers even greater agility and cost savings because applications do not have to be deployed on a server. Instead, functions run from a cloud provider's platform, return outputs, and immediately release the associated resources.

We have four distinct computing models evolving simultaneously: physical server, virtual server, container-based architecture, and serverless computing. This raises interesting questions for organizations: What do we do with our applications? Which workloads should move to the cloud? What must stay on-premise? Can I re-platform some applications to run on VMs? Can I rewrite others for the container or serverless world? Most importantly, how long will it take to get this done so I can unlock the benefits of these new models? The computing-architecture decision has a direct relationship with the future of each application, in fact, it may even dictate the future direction of each application.

This article reviewed some of these decisions and considerations during a practical client implementation of all four computing architectures where factors such as operational challenges, financial considerations, portability, efficiency, and risk tolerance all played a role in making the decision to deploy each of the four computing architecture. ICT service providers, including NTT Ltd., are adapting their go-to-market and partner strategies given the rapid growth of hybrid and hyperscale deployments. Hyperscale providers have such a breadth of service offerings that staying up to date across multiple clouds will require a massive investment on the part of enterprise clients as well as some complexity. In this regard, ICT service providers can provide true value and guidance to their clients.

What is also evident is the need for ICT providers to drive high-value service offers to enterprise clients to guide and support them on the journey ahead. This is why NTT Ltd. is now focused on selling and architecting over forty high-value services. Some relevant services include:

- **Infrastructure consulting** services encompassing several services focused on optimization and transformation wherever our clients are in their transformation journey. Whether it is resolving ageing infrastructure, managing complexity, securing infrastructure, moving to a multi-cloud environment, or improving user experience, we can help our clients evaluate their current operating state and transform and improve the performance of their infrastructure while optimizing their spending.
- **Cloud consulting services** including a hybrid migration assessment to obtain clear insights into a client's current infrastructure and applications to create a roadmap to managed services on a hybrid cloud. Cloud consulting services also include cloud migration services, which are end-to-end services beginning with auditing workloads for suitability for migration, planning the migration to the best-fit venue, then migrating these workloads to the cloud without disrupting daily operations and business.
- **Managed Hybrid Infrastructure Services**

offering comprehensive management and monitoring of on-premise, cloud, and hybrid IT infrastructures to improve operational agility, reduce risk, and optimize cloud and technology infrastructure investments. With our highly trained vendor-accredited people, processes, and tools necessary to manage the day-to-day operations of our clients' multi-platform multi-site IT infrastructures, we ensure the resilience and performance of such infrastructures so one can focus on achieving business outcomes.

- **Managed application services** around the implementation and management of ERP and CRM applications on premise, in the cloud, or in hybrid IT operating environments with consistent governance, security, and management practices.
- Finally, our **Hybrid and Multicloud service** allowing our clients to use the best platform for each workload and taking away the difficulty of managing a hybrid IT environment. With systems spread out across multiple environments, our client's need a way to manage all these environments from a single interface. With our 'Solution Insight' cloud management platform, we enable our clients to gain insight into how all their systems work together. This helps them better understand which applications can be moved to the cloud and which need to stay in their current environments.

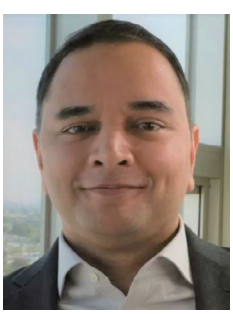
This short list is not an all-inclusive one, as we have several other high-value services we are emphasizing in our engagements with clients across digital transformation, workplace, datacenter, networking, and security.

The shift in the computing landscape is upon us as enterprise clients need to support four different architecture models at the same time. Enterprises will use a combination of cloud services across public and private clouds. The more diverse the mix, the more complex and expensive operations and service continuity become. Through our high-value service offers, we can help enterprise clients navigate decisions and considerations when embarking on a hybrid computing-environment-implementation journey.

Reference

[1] Morgan Stanley, https://www.morganstanley.com/ideas/it-hardware-2018

Trademark notes



Nadeem Ahmad

Group Senior Vice President, Technology Strategy, NTT Ltd.

Nadeem Ahmad is a technology professional with 25 years of ICT industry experience and is a social constructivist at heart. He believes technology does not determine human action but that human action shapes technology. He joined NTT Ltd. (previously Dimension Data) over twenty years ago initially leading technical delivery and operations teams within the United Kingdom, Europe, and Americas but has spent the bulk of his time in the Group CTO Office (now the Innovation and Technology group) focused on business and technology innovation initiatives across the Group.

His role at the Innovation and Technology group involves defining NTT Ltd.'s technology strategy and positioning, in collaboration with go-to-market practice leaders, across relevant technology domains. This includes identifying and monitoring key technology trends and emerging domains, assessing their impact on enterprise clients and their potential to affect new solutions or services, as well as providing input into new offer creation. This extends to providing group guidance on emerging technology domains and translating key trends and market developments into business impact via thought leadership, internally and externally.

LinkedIn profile: https://www.linkedin.com/in/nadeemahmad1/

NTT Pro Cycling: Data-driven Innovation in Action

Nadeem Ahmad

Abstract

With thousands of riders and hundreds of races to choose from each year, the NTT Pro Cycling team uses data and advanced analytics to optimize their performance by ensuring they recruit the best possible riders for their budget and for selected races. Training is then optimized for every individual rider to ensure they are in peak condition for their target races, giving them the best chance to maximize UCI (Union Cycliste Internationale) points. NTT Ltd. supports the team through innovative technology that monitors and manages the performance of the riders. We also use advanced data analytics and data science to help the team make key decisions. Working together within a support ecosystem, we give the team every possible advantage to be a high-performing pro sports team, both on and off the bike.

Keywords: data analytics, technology, NTT Pro Cycling

1. The NTT Pro Cycling team and its mission

In the world of professional cycling, NTT Pro Cycling is unique. It is the first team with roots in Africa competing on the world stage and that is riding to uplift communities in Africa through their support of the Qhubeka charity. Qhubeka is an African not-for-profit organization that donates bicycles as part of the World Bicycle Relief's charity program and has supplied more than 75,000 bicycles through fundraising programs targeted at children in rural African communities to make it easier for them to get to and from school, helping to improve their attendance and performance.

The team strives to be the benchmark within the sport through being a performance-driven, technology-enabled team underpinned by their unique purpose-led approach in their support of the Qhubeka charity. Their commitment to changing lives together with Qhubeka remains stronger than ever, and they will be looking to do so with even greater enthusiasm in the years to come. Together with staff, riders, and valued partners, this core belief remains a foundation of the team. This truly is a team that was born in Africa but created for the world.

As the NTT Pro Cycling team's technology innovation partner, NTT Ltd. helps the team to be the most purpose-led, performance-driven, and technology-enabled team in pro cycling today. We work together to fulfill their ambition of being ranked among the top 10 cycling teams in the world. Racing in the Union Cycliste Internationale (UCI) World Tour, the premier men's elite road cycling tour, means competing against some of the best funded teams in the most iconic cycling races. To achieve this, the team leverages the power of advanced data analytics and data platforms. This enables them to harness all their resources in the best possible way, based on real data, to perform at the highest level possible to win as many points as possible, retaining World Tour status and delivering value for their sponsors who provide their primary revenue stream.

2. Creating a technology-enabled data-driven pro-cycling team

In 2020, the team adopted a fully integrated athlete performance management solution for training analysis, health and wellness tracking, in-race hydration and nutrition management, injury reporting, and overall athlete condition (health and mental wellness) monitoring. Additionally, advanced data analytics

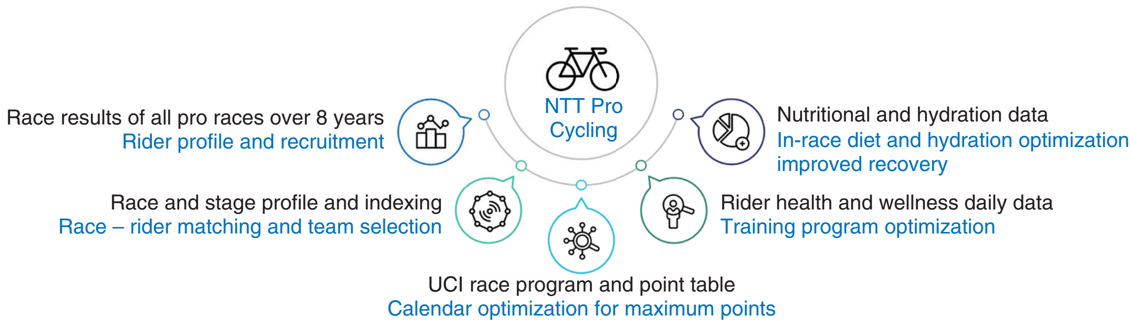


Fig. 1. NTT Pro Cycling's data-driven approach.

help the team with regards to talent recruitment and retention as well as establishing a race calendar and team selection to maximize the potential UCI points and race results (**Fig. 1**).

The team uses data science to drive key decision-making to meet the following three critical performance criteria:

- Recruiting and retaining the best talent to create the best team within the available budget
- Optimizing athlete performance so they can perform at their best
- Constructing the most effective race calendar and team selection to maximize the potential UCI points and race results

NTT Ltd. provides several technologies and services to help NTT Pro Cycling meet these performance criteria. These technologies cut across advanced data analytics, data platform, hybrid cloud, and artificial intelligence and machine learning. We provide a range of services in support of these technologies to address the team's performance criteria including:

- Digital and analytics consulting
- Data platform and data architecture definition
- Data science services
- Systems integration
- Application managed services
- Program management

Over the past year and a half, NTT Pro Cycling, in conjunction with NTT Ltd., has been working on a comprehensive data platform that helps the team climb the UCI rankings. The solution supports the team through innovative technology focused on advanced data science, which is used to help make key decisions faster.

Innovative technology not only supports NTT Pro Cycling's performance but also its business processes. Data analytics help the team select riders, prepare the race calendar, and assign the right riders to the right races. Such technology also helps the team monitor the health of its riders by tracking physiological data through a health and wellness application. To compete at the top of cycling's World Tour, the team is looking to drastically improve and force its way into the top 10 of the World Tour team ranking. The team is challenging themselves to be better. They leverage innovation, technology, and data in new ways to enhance their performance daily and speed up and improve decision making. Like all other organizations, professional sports teams have embarked on a digital transformation journey to remain relevant and competitive. It is only through this that they improve their efficiencies and become more effective organizations. For NTT Pro Cycling, this is no different.

3. Rider recruitment

Every pro cycling team knows their riders are their most important assets. Having the right riders with a range of skills and capabilities is an important foundation for building a successful cycling team and to secure those all-important race victories. NTT Pro Cycling uses data to analyze just about every pro cyclist in the world. Using a combination of historical data and performance metrics, they can identify which riders they want to recruit. By leveraging data analytics, team management can make more informed decisions to build the best possible team while still staying within budget.

Over the past 18 months, NTT Pro Cycling and the NTT Ltd. Advanced Technology Group have jointly developed several statistical and analytical models for classifying different riders to study career trajectories

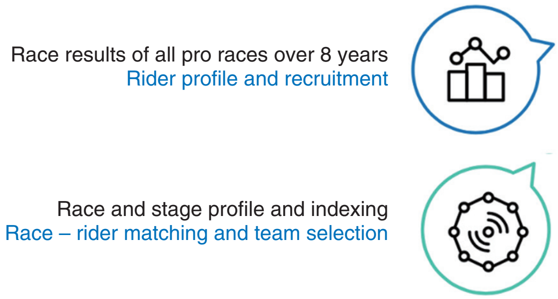


Fig. 2. Rider recruitment data.

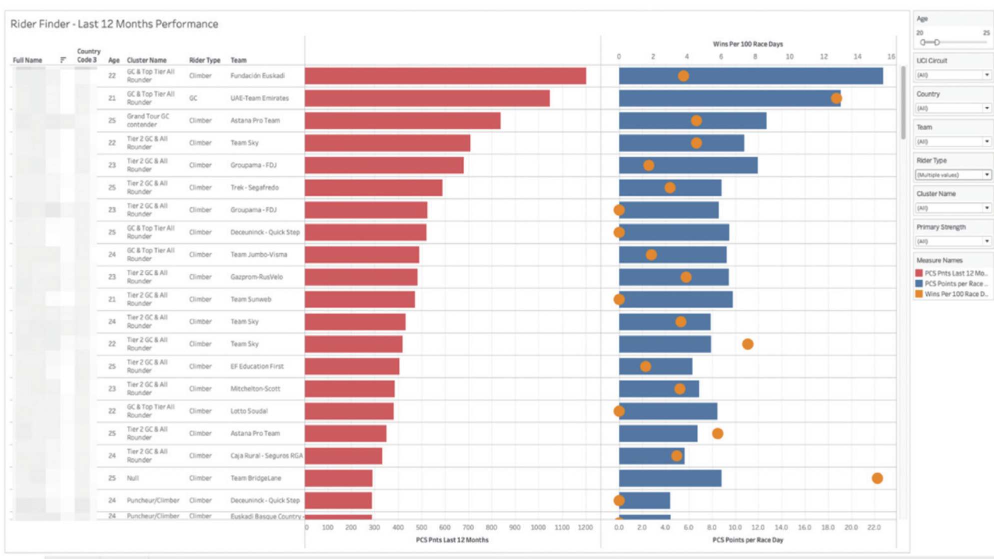


Fig. 3. Example of NTT Pro Cycling rider dashboard.

and form. Built upon data from ProCyclingStats.com, the analytics package allows the team to distil every rider down to a matrix of strengths and weaknesses. It is a process that runs daily in NTT Ltd.'s cloud platform, scraping required data, calculating all the key metrics, then encapsulating them into a dashboard that team management can use to analyze different groups of riders, deep-dive into individual rider profiles, and calculate the trajectory of that rider in the future (**Fig. 2**).

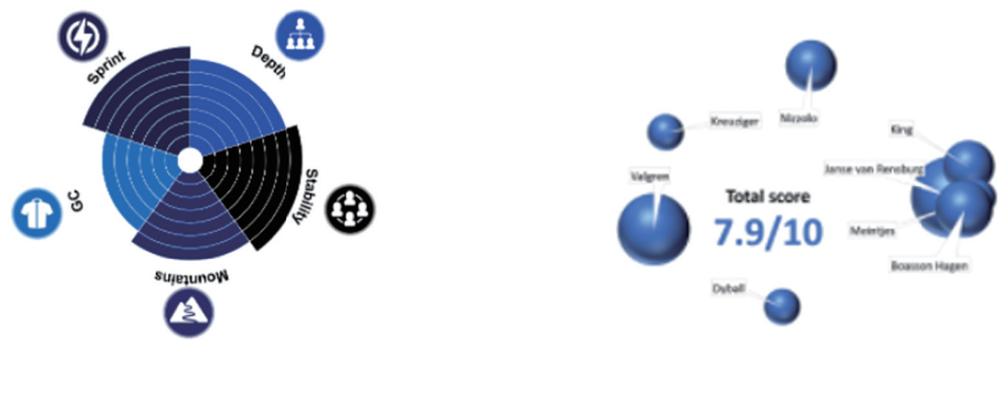
3.1 Building a team

To give the team the intelligence they need, NTT Ltd.'s advanced data analytics solution analyzed over eight years' worth of race results and team performances. The application programming interface (API) from the Pro Cycling Stats platform creates a view of the current composition of the team (see **Fig. 3** for an example). This is then used to identify the specific strengths and weaknesses within the team across five categories of general, sprint, depth, stability, and mountain (example in **Fig. 4**).

Starting with this overall understanding of the team's performance data and metrics, team management can work out which riders work best together. They can then use this to create a 'team stability' score.

3.2 Gap analysis

Once the status of the current team is established, team management can identify any gaps the riders need to fill (see **Fig. 5**). By using the analysis of the



Team profile

Team stability

Fig. 4. Team strengths, weaknesses, and stability.

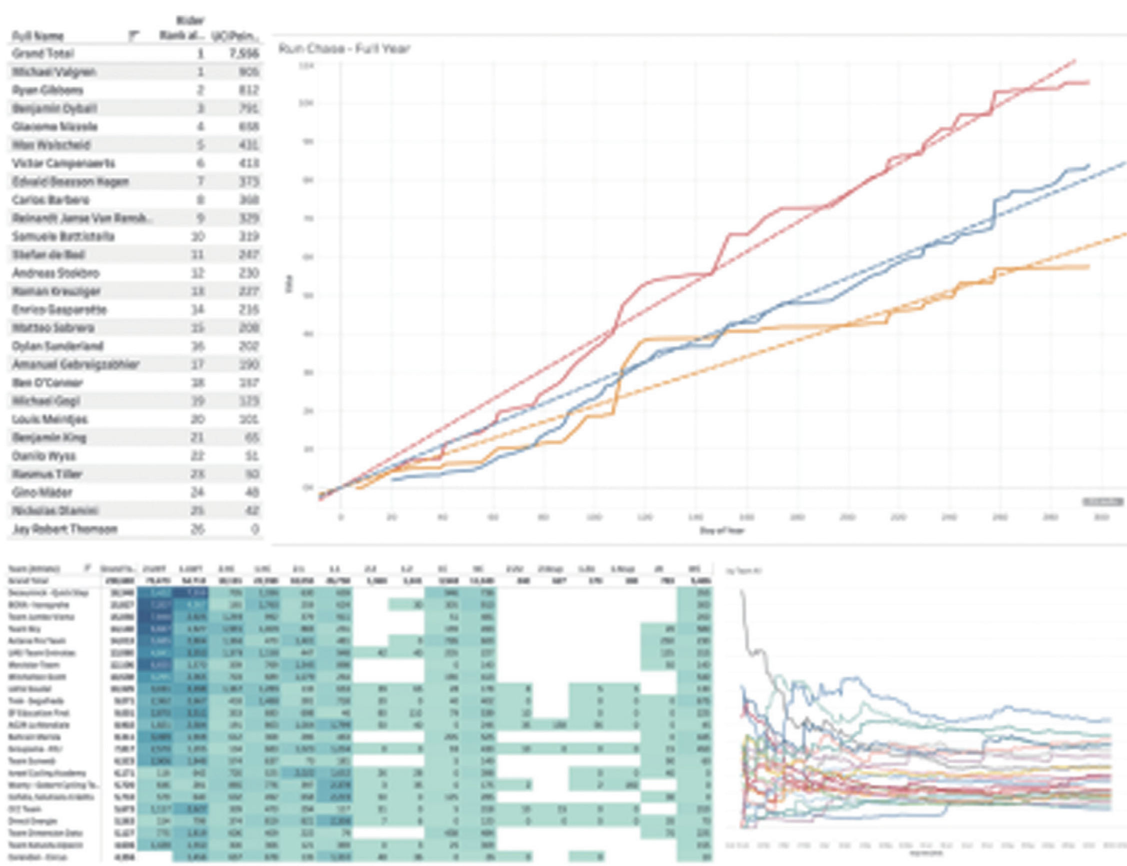


Fig. 5. Rider ranking and gap analysis.

global pool of pro cyclists, they can see which riders would deliver the biggest 'bang for buck.' In addition to examining the strength of the individual riders against the needs of the team, the data analysis solution incorporates the full pro cycling calendar, including the number of points available and the historic strength of the field. This enables team management to identify opportunities to score the maximum amount of points along with which riders and combinations are most likely to be successful.

3.3 Performance analysis

As well as using analytics to establish which riders are the best in any given situation, NTT Pro Cycling uses them to review the global pool of riders. The team understands that discovering the right resources at the right time is critical to success both in business and on the road.

With their understanding of the team's requirements, team management and performance coaches work together to narrow down the thousands of cyclists, creating a shortlist of potential talent. Factors taken into consideration include rider type, race victories, age, and recent performance, to name a few. From that information, the analytics package's algorithm deducts the likelihood of riders earning ever-important World Tour points in races to come, i.e., the 'win quotient.' Together with data-driven rider profiles, management has a view of each rider's strengths and weaknesses, allowing them to identify the skills needed to fill the gaps.

3.4 Integrate, monitor, and report

The final phase is actual recruitment. With a holistic view of the current team, the gaps they need to fill, and the potential talent available from our data and analytics solution, team management are able to make quicker, more informed decisions, taking some of the guesswork out of the process. These tools allow decision-makers to make objective choices to help them achieve their goals. All this information allows team management to sign riders and be assured that, together with their vast industry knowledge and experience, technology will enable them to improve their performance. To monitor and evaluate performance, NTT Ltd. has integrated multiple technologies, wearables, Internet-of-Things devices, platforms, and tools to provide both detailed performance analytics and a holistic snapshot for management.

One final note on rider recruitment. NTT Ltd.'s solution allows the team to look beyond the big names, to the true 'moneyball' performers, i.e., the riders who have previously been overlooked but provide a unique value proposition. Using a combination of historical data and performance metrics, experienced sports directors can supplement the process by identifying riders for specific roles to help fill out the roster.

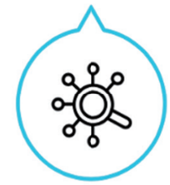


Fig. 6. Point data for calendar optimization.

4. Data-driven calendar optimization

In addition to the 36 compulsory World Tour events, there are hundreds of races for teams to choose from, so making the decision about which races to participate in with which riders is critical to success. For this reason, NTT Pro Cycling uses data analytics to help optimize their race calendar and use their resources most effectively. The team participates in 82 races on 256 days across five continents every year with up to three different races on a single day. This creates not just logistical challenges but challenges in ensuring optimal team performance. To obtain the best possible results, the right riders need to be at the right races at the right time. The logistical issue is that there are 29 riders, 54 management and support staff members, over 200 bicycles, 250 sets of carbon wheels, 22 vehicles, over 900 airplane tickets, 3.5 tons of nutrition, and dozens of partners and assets across the globe. This makes up the complex team infrastructure that is NTT Pro Cycling. As part of NTT Pro Cycling's drive to secure World Tour points, the team must make tough decisions about which races to participate in. The calendar needs to create a balance among logistical feasibility, the opportunity to earn points, rider schedules, their strengths, and their training and race load (**Fig. 6**).

At the top level of cycling—the UCI World Tour—securing enough points is critical to the team's survival. With different tiers of races and different point scales, it is essential to know where the best opportunities to score points are. Data and analytics tools help the team identify this and from that, build a race calendar.

The team uses an algorithm to take into account the points available by race, calculate field strengths based on the caliber of riders who have competed in the past, and create a points-per-race-day weighted by the field strength indicator. This shows the best

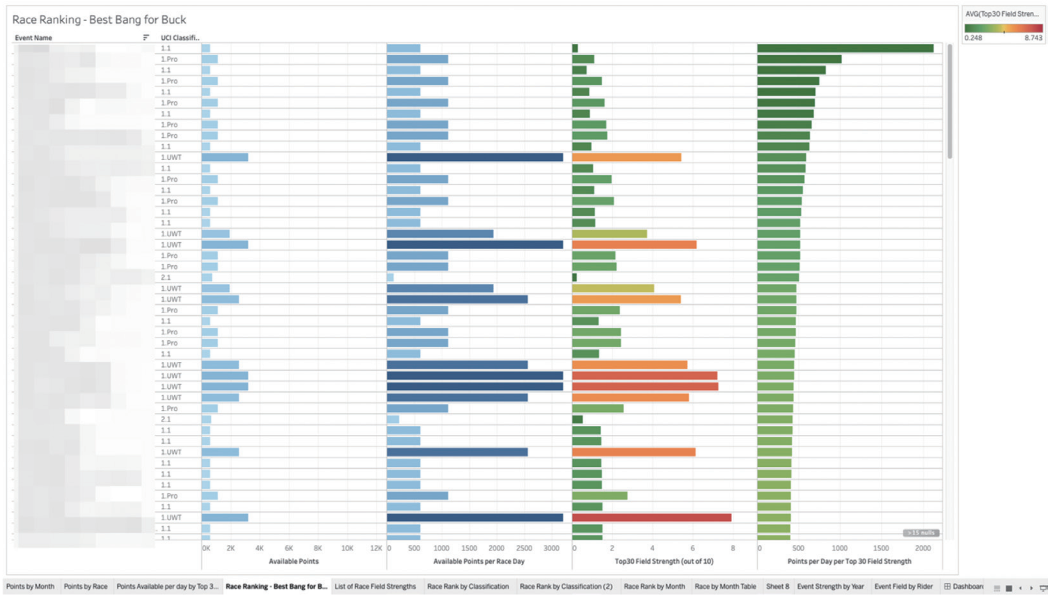


Fig. 7. Race calendar ranking and the biggest “bang for buck.”

value races where they are likely to achieve the best results, therefore earn the most points.

From there, the scoring, together with the logistical constraints of time and distance, helps the team build out the race calendar. Coupled with our rider and team composition analyses, they can pinpoint who should go to which races. The system produces multiple dashboards (see **Fig. 7** for an example), which team management can use along with other key information such as rider programs and team logistics, in determining which races they should attend.

One of the main findings from this analysis is that it is four times easier to win points in Asia Tour .HC (now ProTour) races than in European World Tour races. It also shows that the team’s riders can win more points in one-day races than stage races. There are more points on offer at World Tour races, but they are statistically more difficult to earn. The analysis proved its worth early on with strong results across the globe:

- Eight top-2 finishes with six different riders on three continents
- Stage wins at both Étoile de Bessèges and Tour de Langkawi (Ben O’Connor and Max Walscheid)
- Leader of the points classification at the Tour of Langkawi (Max Walscheid)
- National Championships of South Africa (Ryan Gibbons)

5. Rider training, health, and wellness

Training, preparation, nutrition, and health are key ingredients for any successful pro-cyclist. NTT Pro Cycling understands this as they use a combination of technology solutions from their ecosystem of partners including NTT Ltd., Alcatel Lucent Enterprise, Lumin Sports, Today’s Plan, Garmin, and Rotor, to monitor and optimize the riders training and race preparation.

Phila (an Nguni word meaning ‘to live’) is a health and wellness mobile application that provides a morning and evening questionnaire to check on the rider’s location, sleep, fatigue, mood, soreness, illness, injury, training, and race feedback. It allows the coach to provide immediate feedback on the day’s training plans (**Fig. 8**).

Assisting in optimizing hydration and nutrition for recovery, the team uses the Hydration and Nutrition Guidance (HANG), an automated application for use at races. The app is linked with Bluetooth Smart Scales to capture morning, pre-, and post-race weights to track hydration levels and provide personalized recommendations to each rider following the race. The team uses a centralized athlete data reporting tool called ARC, which integrates all the training data, health and wellness data, HANG data, and race reporting to provide the coaches and team management with performance dashboards on key metrics,

Fig. 8. Rider health and nutritional data.

including the optimal training load for each rider and race analysis and modelling training load based on rider race schedules.

6. Summary and results

For NTT Pro Cycling, innovative and advanced data science is used to help make key decisions faster to not only support the team performance but also their business processes. Data analytics help them select riders, prepare their race calendar, and assign the right riders to the right races. Innovative ecosystem technologies also help them monitor the health of its riders by tracking physiological data through a health and wellness application.

It would be naive to think that NTT Pro Cycling is the only team looking at gaps in its line-up and trying to recruit to fill those gaps. It is also unlikely to be the only team with a data analytics solution of some kind to help with the process. However, the team is far ahead of the curve in the road cycling space in the level of sophistication and detail of performance monitoring and management solution to provide access to their capabilities and all the crucial data via one unified data analytics platform. With a holistic view of the current team, the gaps they need to fill, and the potential talent available from the data and analytics solution, team management is able to make quicker more informed decisions, taking some of the guesswork out of the process.

Using data-driven recruiting, based on team performance, 2020 NTT Pro Cycling's selected riders should be expected to deliver 25% more UCI points than the 2019 team did, which would result in them moving up 7 positions in the UCI World ranking. To move into the top 15 cycling teams in the world, they require their top 10 riders to win 25% more points than in 2019. To achieve a top 10 ranking would require double the points they accumulated in 2019. Through optimizing the race calendar alone, there is potential to increase the available points they can win by 20–30%.

6.1 Initial results

At the time of the writing of this article (before the Tour de France that began in August), the 2020 racing season had already proven to be very successful, reaping the benefits from a technology-enabled, data-and-performance-driven team. With an increase of 18% more UCI points, a ground-breaking 200% increase in wins and 114% more podiums compared to last year's results, the solution in place has already proven to be of immense value to the team and its management.

When NTT Ltd. built the data and analytics platform, the team committed to use the data, applications, and platforms that are collecting the data, so algorithms can be applied to analyze the data to make accurate data-driven decisions. The key is that they trust the data and monitor the outcomes, adjusting along the way, to refine the solution over time and achieve the goal of becoming a top-10 World Tour cycling team.

Trademark notes

Nadeem Ahmad
Group Senior Vice President, Technology Strategy, NTT Ltd.

Nadeem Ahmad is a technology professional with 25 years of ICT industry experience and is a social constructivist at heart. He believes technology does not determine human action but that human action shapes technology. He joined NTT Ltd. (previously Dimension Data) over twenty years ago initially leading technical delivery and operations teams within the United Kingdom, Europe, and America, but has spent the bulk of his time in the Group CTO Office (now the Innovation and Technology group) focused on business and technology innovation initiatives across the Group.

His role at the Innovation and Technology group involves defining NTT Ltd.'s technology strategy and positioning, in collaboration with go-to-market practice leaders, across relevant technology domains. This includes identifying and monitoring key technology trends and emerging domains, assessing their impact on enterprise clients and their potential to affect new solutions or services, as well as providing input into new offer creation. This extends to providing group guidance on emerging technology domains and translating key trends and market developments to business impact via thought leadership, internally and externally.

LinkedIn profile: https://www.linkedin.com/in/nadeemahmad1/

Researching AI Technologies to Support and Substitute Human Activities

Hidenori Tanaka, Masaki Kitahara, and Yoshinori Kusachi

Abstract

The coming of digital transformation has been gaining interest, and its development seems to be accelerating during the COVID-19 pandemic. Competition in artificial intelligence (AI) technologies is also intensifying, and platformers are accumulating massive amounts of data as learning data. Amid these trends, NTT Media Intelligence Laboratories is promoting the research and development (R&D) of AI technologies to support and substitute human activities by taking advantage of the technologies and expertise that it has cultivated. This article introduces this R&D initiative.

Keywords: support of human activities, substitute for human activities, AI technology

1. Introduction

NTT Media Intelligence Laboratories has been involved in the research and development (R&D) of technologies for processing various forms of media, such as voice, audio, language, images, and video, and has put several of these technologies into practice. More recently, it has made a number of contributions to business such as operator support at contact centers [1], provision of artificial intelligence (AI) agents [2, 3], sound-collection technology for emergency call systems [4], and video-compression devices for 4K/8K broadcasting [5].

However, today's market environment is going through major changes and competition is intensifying. Every industry is experiencing digital transformation (DX), which is revolutionizing business models through digitalization, and the development of DX is expected to accelerate to deal with the COVID-19 pandemic. The advent of deep learning has ignited a third AI boom, and the basic algorithms of AI technologies are now available to anyone. Moreover, learning data that contributes to the improved performance of AI systems is now being collected on a large scale by platformers, such as Google, Apple, Facebook, and Amazon, so a world in which AI performance increases daily is becoming a reality.

In response to this external environment, the NTT Group is promoting initiatives such as B2B2X (business-to-business-to-X) and DX toward the creation of a *smart world* as part of its Medium-Term Management Strategy. It is also promoting the concept of the Innovative Optical and Wireless Network (IOWN) to build a smart world through innovative technologies.

Against this background, NTT Media Intelligence Laboratories is leveraging the technologies and expertise cultivated through R&D activities in media processing to pursue the R&D of AI technologies for supporting and substituting human activities—the source of all value—and the R&D of Digital Twin Computing (DTC) to create new value from a medium- to long-term perspective [6]. In this article, we introduce the R&D of AI technologies toward an application domain in which AI supports and substitutes human activities.

2. Overview of AI technologies to support and substitute human activities

Several possible scenarios can be considered when talking about an application domain in which AI supports and substitutes human activities. For example, in terms of increasing efficiency, AI can not only improve the productivity of operators at contact centers and AI agents but also improve business processes and productivity in offices and improve the quality of life as new value. We expect the COVID-19 pandemic to drive the penetration of teleworking and online meetings and transform the way in which they should be carried out.

However, applying current AI technologies to such scenarios is difficult. For example, when attempting to support and substitute human activities on a more personal level based on a person's environment, there would be a need to obtain data on that person and environment, but there are situations in which a large amount of data cannot be obtained. Under such conditions, it would be difficult to achieve the full potential of AI. Taking speech-recognition technology as an example, differences in its application would arise in telephone calls, meetings, etc. depending on whether only speech-to-text processing is sufficient or whether processing as far as speaker identification is necessary. In the case of online meetings, there may even be more differences in expected performance and other requirements.

NTT Media Intelligence Laboratories has begun to address these issues by focusing its efforts on achieving efficient learning with a relatively small amount of data, producing new AI technologies, and achieving breakthroughs in the performance of current technologies.

3. Current initiatives in AI technologies to support and substitute human activities

The Feature Articles in this issue introduce a group of technologies that we are currently developing. First, in the article "Media-processing Technologies for Ultimate Private Sound Space" [7], the focus is on sound as a major element of technologies that can produce new AI technologies, such as for understanding surrounding conditions from sound, listening to only the person you want to hear, and eliminating sound you do not want to hear. Our aim with these technologies is to achieve an ultimate private space that can be expected in applications such as teleworking.

Next is "Saxe: Text-to-Speech Synthesis Engine Applicable to Diverse Use Cases" [8]. Regarding speech-synthesis technology for generating speech for virtual announcers, AI agents, etc., this article introduces technology for high-accuracy reading of heteronyms according to context and deep-neural-network speech-synthesis technology for reproducing a variety of speaker characteristics at low cost. These technologies are aimed at achieving breakthroughs in the performance of current technologies and efficient learning with a relatively small amount of data.

In "Speech Recognition Technologies for Creating Knowledge Resources from Communications" [9], we envision speech recognition in scenarios such as meetings and face-to-face customer service. In addition to techniques for improving current speech-to-text technology, this article introduces technologies for producing new AI technologies such as for extracting the gender, emotions, and other characteristics of the speaker from sound.

In "Knowledge and Language-processing Technology that Supports and Substitutes Customer-contact Work" [10], we introduce technologies for achieving breakthroughs in the performance of current technologies and producing new AI technologies. These technologies include document summarizing with a certain specified length according to the application scenario and response analysis for improving the productivity of an operator involved in inside (remote) sales.

Finally, to achieve a fourth-dimensional (4D) digital platform [11] that integrates various types of sensing data in real time and enables a variety of future predictions, the article "Spatial-information Processing Technology for Establishing a 4D Digital Platform" [12] introduces technology for digitalizing real-world space and point-cloud-coding technology for efficiently compressing 3D data that include temporal changes. These are technologies for efficient learning with a small amount of data and achieving breakthroughs in the performance of current technologies.

4. Anticipated use cases

We now introduce use cases of AI technologies for supporting and substituting human activities that take into account the recent COVID-19 pandemic (**Fig. 1**). In creating personal space, the aim is to construct a pseudo personal space for teleworking and a space that ensures privacy within one's home. In online

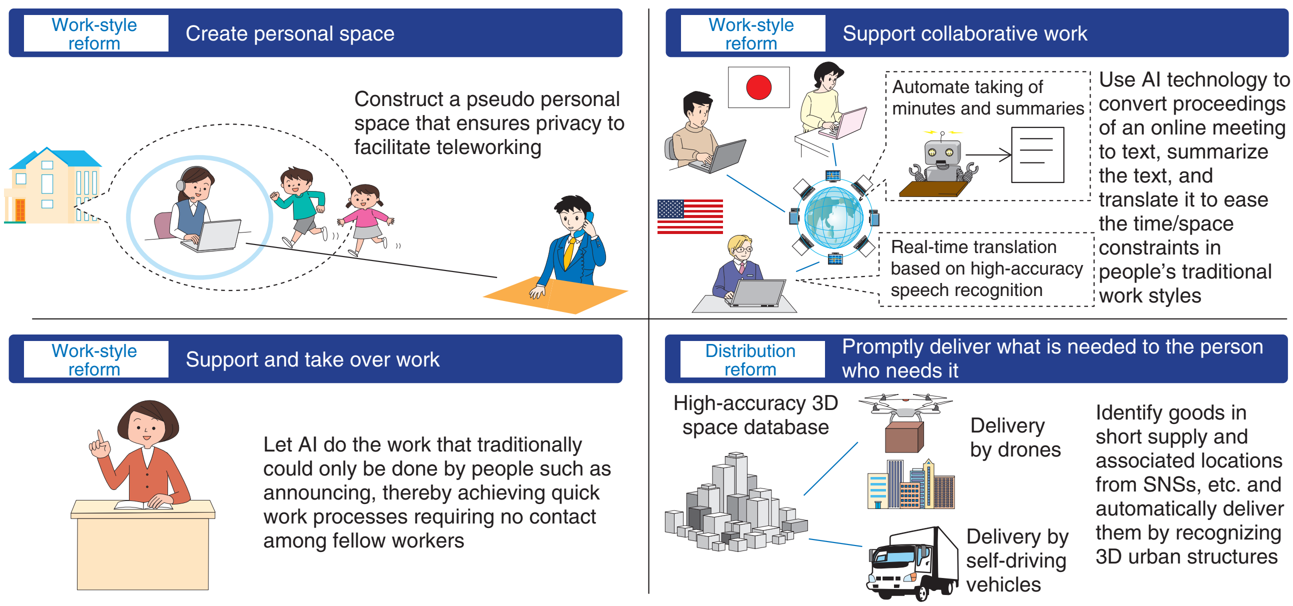


Fig. 1. Use cases of AI technologies to support and substitute human activities.

meetings, AI technologies will be used to support innovative collaborative work by converting the proceedings of a meeting to text, summarizing and translating the meeting, and easing the time/space constraints in people's traditional work styles. Additionally, by letting AI do the work traditionally done by people, such as announcing, we anticipate quick work processes that require no contact among fellow workers. Finally, from the viewpoint of distribution reform, we envision the possibility of identifying goods in short supply and associated locations from social networking services (SNSs), etc. and automatically delivering those goods by recognizing 3D urban structures.

5. Future outlook

The surrounding environment is changing rapidly. To provide various means of supporting and substituting human activities, technology must deal accordingly with these changes. To this end, we plan to promote R&D with a flexible frame of mind while understanding both the macro and micro aspects of these changes.

References

[1] ForeSight Voice Mining, https://www.ntt-tx.com/products/foresight_vm/
[2] Website of NTT Communications' AI services, https://www.ntt.com/en/services/ai.html
[3] Website of NTT DOCOMO's my daiz AI agent service (in Japanese), https://www.nttdocomo.co.jp/service/mydaiz/
[4] Press release issued by NTT on Feb. 19, 2018 (in Japanese). https://www.ntt.co.jp/news2018/1802/180219c.html
[5] Press release issued by NTT, "NTT Develops the Worlds' Best Performance 8K HEVC Real-time Encoder through Dedicated LSI Use," Feb. 15, 2016. https://www.ntt.co.jp/news2016/1602e/160215a.html
[6] White paper on Digital Twin Computing, https://www.ntt.co.jp/svlab/e/DTC/DTC_Whitepaper_en_2_0_0.pdf
[7] M. Fukui, S. Saito, and K. Kobayashi, "Media-processing Technologies for Ultimate Private Sound Space," NTT Technical Review, Vol. 18, No. 12, pp. 43–47, 2020. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202012fa6.html
[8] Y. Ijima, N. Kobayashi, H. Yabushita, and T. Nakamura, "Saxe: Text-to-Speech Synthesis Engine Applicable to Diverse Use Cases," NTT Technical Review, Vol. 18, No. 12, pp. 48–52, 2020. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202012fa7.html
[9] Y. Nakazawa, T. Mori, Y. Yamaguchi, Y. Shinohara, and N. Miyazaki, "Speech Recognition Technologies for Creating Knowledge Resources from Communications," NTT Technical Review, Vol. 18, No. 12, pp. 53–58, 2020. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202012fa8.html
[10] K. Nishida, K. Saito, T. Amakasu, K. Iso, and S. Nishioka, "Knowledge and Language-processing Technology that Supports and Substitutes Customer-contact Work," NTT Technical Review, Vol. 18, No. 12, pp. 59–63, 2020. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202012fa9.html
[11] Press release issued by NTT, "4D Digital Platform™"; Integrates Various Sensing Data in Real Time and Enables Future Predictions," Mar. 26, 2020. https://www.ntt.co.jp/news2020/2003e/200326c.html
[12] Y. Yao, K. Kurata, N. Ito, S. Ando, J. Shimamura, M. Watanabe, R.

Tanida, and H. Kimata, “Spatial-information Processing Technology for Establishing a 4D Digital Platform,” NTT Technical Review, Vol. 18, No. 12, pp. 64–70, 2020.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr202012fa10.html

Hidenori Tanaka
Senior Research Engineer/Producer, Planning Section, NTT Media Intelligence Laboratories.

He received a B.E., M.E., and Ph.D. in science and technology from Keio University, Kanagawa, in 2004, 2006, and 2011. After joining NTT in 2006, he engaged in research on computer vision and pattern recognition. He has been involved in technology marketing and R&D management.

Yoshinori Kusachi
Chief Producer, Planning Section, NTT Media Intelligence Laboratories.

He received a B.E. in computer science from Kyoto University in 1995 and an M.E. in information science and Ph.D. in information engineering from Nara Institute of Science and Technology in 1997 and 2007. He joined NTT in 1997 and engaged in research and practical application development in the fields of image processing, computer vision, and pattern recognition. He has also been involved in R&D management.

Masaki Kitahara
Senior Research Engineer, Scene Analysis Technology Group, Universe Data Handling Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. in industrial and management engineering from Waseda University, Tokyo, in 1999 and 2001. After joining NTT in 2001, he engaged in the development of compression and rendering algorithms for 3D video. He transferred to NTT Advanced Technology (NTT-AT) and managed various software development projects. In 2009, he returned to NTT and has been engaged in the development of video encoding and AI technologies and R&D planning. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Media-processing Technologies for Ultimate Private Sound Space

Masahiro Fukui, Shoichiro Saito, and Kazunori Kobayashi

Abstract

NTT Media Intelligence Laboratories has been researching and developing media-processing technologies to create ultimate private spaces for promoting digital transformation in diverse spaces, including work-style reform. This article focuses on sound, which is one of the most important elements in achieving private space and introduces our aims in establishing the following three technologies: technology to understand surroundings from sound (event-detection and scene-identification techniques), active-sound-control technology to make the sound heard only by listeners who want to hear, and active-noise-control technology to eliminate undesired sounds. This article describes our, the authors', efforts regarding these technologies.

Keywords: event detection, active sound control, active noise control

1. Introduction

Due to the work-style reforms promoted by the Japanese government and the impact of the novel coronavirus, the traditional way of working that requires employees to come to the office has been reconsidered, and flexible work styles that are not bound by location or time have been attracting attention. An important part of this new work style is the ability to create a comfortable working environment, no matter where one is.

Let us consider telecommuting (remote working). There are various types of noise in the home, such as from the air conditioner, traffic outside, and sometimes chimes to announce deliveries. If there are family members present, there may be their voices and sounds from the television (TV). Such noise includes sounds that telecommuters do not want to hear. However, in some situations, chimes and a baby crying may be sounds one wants to hear. When a conference call is held at home, we do not want the noise generated near us to reach the other party. We also do not want those near us to hear the voice from the other party. If we can create an ultimate private sound space in which telecommuters can hear only what they want to hear, e.g., voices from their telecommunication partners, they will be able to work comfortably at home.

The goal of NTT Media Intelligence Laboratories is to create the ultimate private sound space, which we call a Personalized Sound Zone (PSZ), by accurately collecting sound information and understanding the situation. NTT Media Intelligence Laboratories has accumulated a large amount of knowledge on sound-collection technology and is currently working on further developing sound situational understanding and control technologies. The following describes the three major technologies (**Fig. 1**).

2. Technology to understand surroundings from sound (event-detection and scene-identification techniques)

People want to hear different sounds depending on the situation. For example, a user may want to hear his/her dog bark when at home, but may not want to hear the barking of other dogs when away from home. In such a case, if we can detect the sound of a dog barking while the user is away from home, we can judge that the sound as undesired.

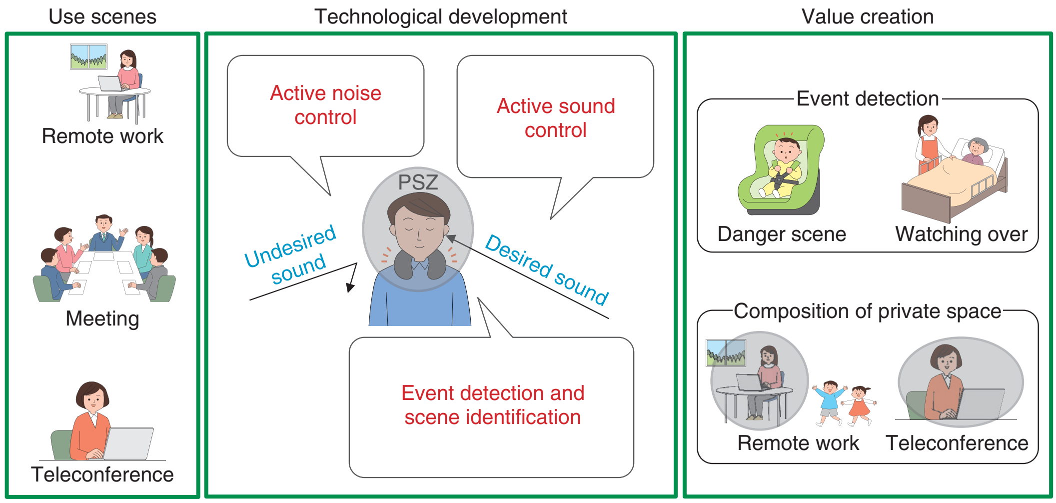


Fig. 1. Concept of Personalized Sound Zone (PSZ).

Thus, to achieve a PSZ, it is important to selectively convey sounds and situations to the user according to the situation, rather than simply suppressing all ambient sounds. To do so, it is necessary to be aware of the environment surrounding the user. For this purpose, we are developing an event-detection technique for simultaneous estimation of information, such as "when," "what," and "where," and a scene-identification technique for estimating the meaning of information such as "in what situation" and "why."

One of the difficulties with event detection is that the sound changes in various ways depending on the surrounding environment, even if the sound reaches the same location. For example, deep neural networks (DNNs) have recently become the mainstream method for determining where sounds are generated, but due to the diversity of the environment, even DNNs cannot be 100% effective. We are working on improving estimation accuracy by using the spatial symmetry of the sound field and combining it with physical-quantity estimation methods [1–3]. We are also investigating a method for satisfying the requirement of detecting only specific events with fast computation and low complexity [4].

The goal with the scene-identification technique is to estimate the information about the user's situation as higher-level information than events and sound-source locations. For example, this technique will estimate not only the sound of a car running but also the user's situation; therefore, it can judge whether the sound should be suppressed or alert the user according to his/her situation. We are currently developing a sound-description-generation technique [5] for describing sound signals in natural language as an elemental technology of the scene-identification technique.

3. Active-sound-control technology to make the sound heard only by listeners who want to hear

Listeners have been using earphones and headphones to listen to sound without affecting their surroundings. However, there are many problems, such as the inconvenience of wearing/carrying a device, possibility of fatigue and hearing loss due to prolonged use, and difficulty in detecting the situation and possible danger around when wearing/carrying such a device. Therefore, if localized sound-zone generation is possible so that only the target listener can hear without using earphones or headphones, these problems can be solved (**Fig. 2**). NTT Media Intelligence Laboratories is engaged in research and development (R&D) of both software and hardware to develop active-sound-control technology for generating a localized sound zone. The following sections discuss issues and approaches regarding this technology.

3.1 Efforts to improve software performance

Although localized sound-zone generation requires multiple loudspeakers, its control is implemented using software and uses signal processing called active sound control. This technology requires more

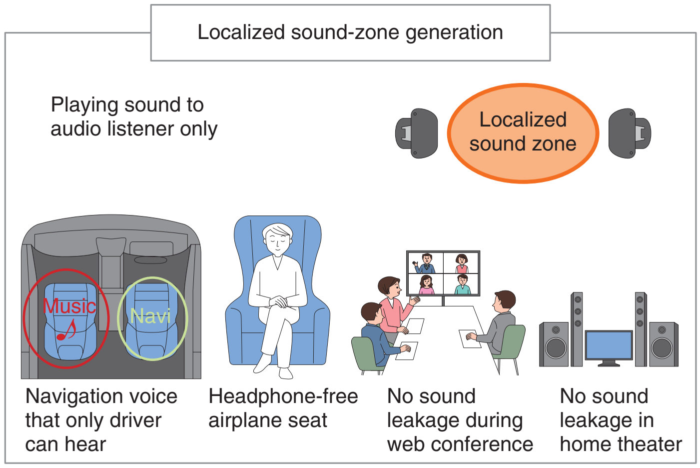


Fig. 2. Application range of active-sound-control technology.

loudspeakers when the upper frequency range is set higher. Also, there may be restrictions on the placement of each speaker. PSZs are designed for individual users, and the number and degree of freedom of loudspeaker placement are severely restricted, unlike general problem settings. For example, at home, the space where the speakers can be installed is limited, such as around a computer desk. The aim with PSZs is to achieve localized sound-zone generation with a small number of loudspeakers and limited space. Active-sound-control technology identifies all the conditions necessary for filter design and optimizes the overall filter to satisfy all conditions simultaneously.

3.2 Efforts to improve hardware performance

In addition to researching signal processing, we are also investigating the loudspeaker arrangement that minimizes sound leakage with a limited number of loudspeakers and installation locations as well as hardware that has a greater attenuation of sound level as it becomes further away from the loudspeakers. We are also engaged in R&D of small speakers for bass reproduction. Bass is important to maintain high sound quality. The physical size of the loudspeaker body is necessary to reproduce bass at a sufficient volume. However, as described above, it is not practical to use a large loudspeaker because of the space constraint of PSZs. We aim to improve the hardware so that the bass limit of small speakers can be lowered more than ever before.

4. Active-noise-control technology to eliminate undesired sounds

A PSZ uses the event-detection and scene-identification techniques to discriminate between incoming sounds to create a space where undesired sounds are not heard. Currently, noise cancellation in earphones and other widely used devices is easy to achieve because the space to muffle the sound is small and fixed. However, wearing earphones for a long period causes discomfort, such as earache. If we can develop technology to muffle undesired sounds using a device that does not need to be worn, it will be more convenient and expand the range of applications (**Fig. 3**).

Active-noise-control technology that muffles sound in a space consists of a control loudspeaker that generates the control sound, error microphone that detects the error signal at the control point, reference microphone that references the noise signal, and controller that uses an adaptive algorithm to produce the control sound. If the error signal detected with the error microphone is small, undesired sounds are reduced.

The more loudspeakers there are that produce controlled sounds, the more points can be controlled and the easier it is for active noise control to eliminate undesired sounds. However, when considering use in the home, a small number of speakers is desirable. Considering the housing conditions in Japan, these loudspeakers and microphones will be located close to each other. Therefore, problems may occur, such as degradation in performance due to the control sound going around the reference microphones, which was

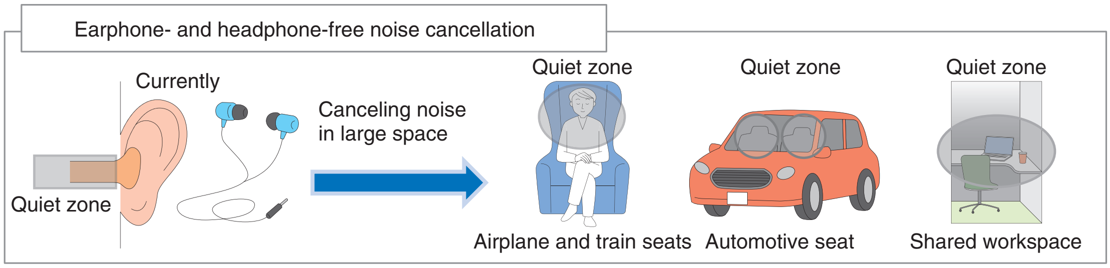


Fig. 3. Application area of active-noise-control technology.

not envisioned in the previous active-noise-control technology.

Undesired sound depends on the individual and environment. For example, people prefer not to hear the sound of a car when they are indoors, but a pedestrian feels safer if he/she can hear car noise. After detecting the sound with the event-detection and scene-identification techniques, technology is also needed to determine the necessity of sound according to the situation, for example, whether the sound is necessary at the moment. To achieve a space where the listener can hear only what the listener wants to hear, multiple technologies are needed, and they must be coordinated at a high level.

5. Future prospects

This article outlined PSZs and described the current status of its elemental technologies, such as technology to understand surroundings from sound (event-detection and scene-identification techniques), active-sound-control technology, and active-noise-control technology. There are still many technical issues to be addressed, and NTT Media Intelligence Laboratories will continue to conduct R&D on this topic. We will also collaborate with other organizations both inside and outside the group to achieve PSZs.

References

[1] L. Mazzon, Y. Koizumi, M. Yasuda, and N. Harada, "First Order Ambisonics Domain Spatial Augmentation for DNN-based Direction of Arrival Estimation," Proc. of the 4th Workshop on Detection and Classification of Acoustic Scenes and Events (DCASE 2019), New York, USA, Oct. 2019.

[2] M. Yasuda, Y. Koizumi, S. Saito, H. Uematsu, and K. Imoto, "Sound Event Localization Based on Sound Intensity Vector Refined by DNN-based Denoising and Source Separation," Proc. of the 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2020), May 2020.

[3] R. Sato, K. Niwa, and K. Kobayashi, "Ambisonic Domain DNN Preserving Physical Symmetry and Its Application to Sound Event Detection and Direction of Arrival Estimation," The 2020 Autumn meeting of the Acoustical Society of Japan, Sept. 2020 (in Japanese).

[4] S. Murata, S. Saito, K. Kobayashi, and A. Nakagawa, "A Study of Lightweight Sound Event Detection Method Based on Decision," The 2020 Autumn meeting of the Acoustical Society of Japan, Sept. 2020 (in Japanese).

[5] Y. Koizumi, R. Masumura, K. Nishida, M. Yasuda, and S. Saito, "A Transformer-based Audio Captioning Model with Keyword Estimation," Proc. of INTERSPEECH 2020, Oct. 2020.

Masahiro Fukui
Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E., M.E., and Ph.D. in information science from Ritsumeikan University, Shiga, in 2002, Nara Institute of Science and Technology in 2004, and Ritsumeikan University in 2018. Since joining NTT in 2004, he has been engaged in research on acoustic echo cancellers and speech coding. He received the best paper award of the ICCE conference and the technical development award from the Acoustic Society of Japan (ASJ) in 2014. He is a member of the Institute of Electrical and Electronics Engineers (IEEE) and the ASJ.

Shoichiro Saito
Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. from the University of Tokyo in 2005 and 2007. Since joining NTT in 2007, he has been engaging in research and development of acoustic signal processing systems, including acoustic echo cancellers, hands-free telecommunication, and anomaly detection in sound. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE), IEEE, and ASJ.

Kazunori Kobayashi
Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E., M.E., and Ph.D. in electrical and electronic system engineering from Nagaoka University of Technology, Niigata, in 1997, 1999, and 2003. Since joining NTT in 1999, he has been engaged in research on microphone arrays, acoustic echo cancellers, and hands-free systems. He is a member of IEICE and ASJ.

Saxe: Text-to-Speech Synthesis Engine Applicable to Diverse Use Cases

Yusuke Ijima, Nozomi Kobayashi, Hiroko Yabushita, and Takashi Nakamura

Abstract

Technological advances such as deep learning and changes in the social structure where artificial intelligence supports and substitutes human activities have recently been accompanied by changes in use cases requiring speech-synthesis technology and in the functions and performance they require. Key issues in implementing these new use cases are understanding according to context, reproduction of speaker characteristics, and support for diverse operation environments. NTT Media Intelligence Laboratories is developing a speech-synthesis engine (development codename "Saxe") based on deep neural networks to address these issues. This article provides a technological overview of Saxe along with several application examples and touches upon future developments.

Keywords: speech synthesis, deep learning, diversity

1. Introduction

Speech synthesis is technology for generating speech from input text, therefore it is sometimes called text-to-speech synthesis. NTT has a long history of research and development in the field of speech synthesis and has developed a variety of speech-synthesis technologies that have found widespread use in services with the purpose of conveying information accurately. These include the telephone services web171 (disaster emergency message board) and 177 (weather forecast telephone service) and IVR (interactive voice response) systems.

There have also been a variety of technological advances typified by deep learning and changes in the social structure where artificial intelligence (AI) is increasingly being used to support and substitute human activities. These developments are being accompanied by changes in use cases requiring speech-synthesis technology and in the functions and performance they require. Uses cases that aimed to convey information accurately have required the generation of speech for standard text in a specific speaker's voice. However, use cases that support and substitute human activities require the generation of speech for all kinds of text in a user-desired speaker's voice in diverse operation environments. NTT Media Intelligence Laboratories is developing a speech-synthesis engine (development codename "Saxe") based on deep neural networks (DNNs) in response to these issues and is promoting its practical application in diverse use cases. In this article, we outline the technologies behind this engine, introduce several application examples, and touch upon future developments.

2. Technology overview

2.1 High-accuracy reading of heteronyms according to context

Speech synthesis can be broadly divided into a text-analysis section that infers the reading and accents of characters from input text and a speech-synthesis section that generates synthesized speech from the

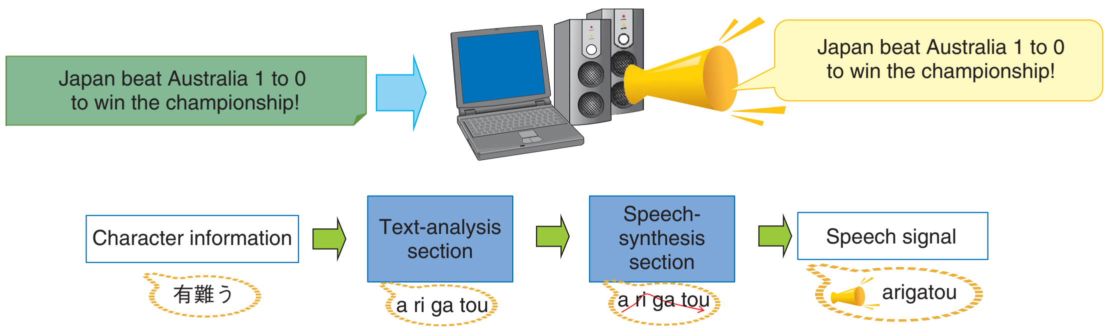


Fig. 1. Overview of speech-synthesis technology.

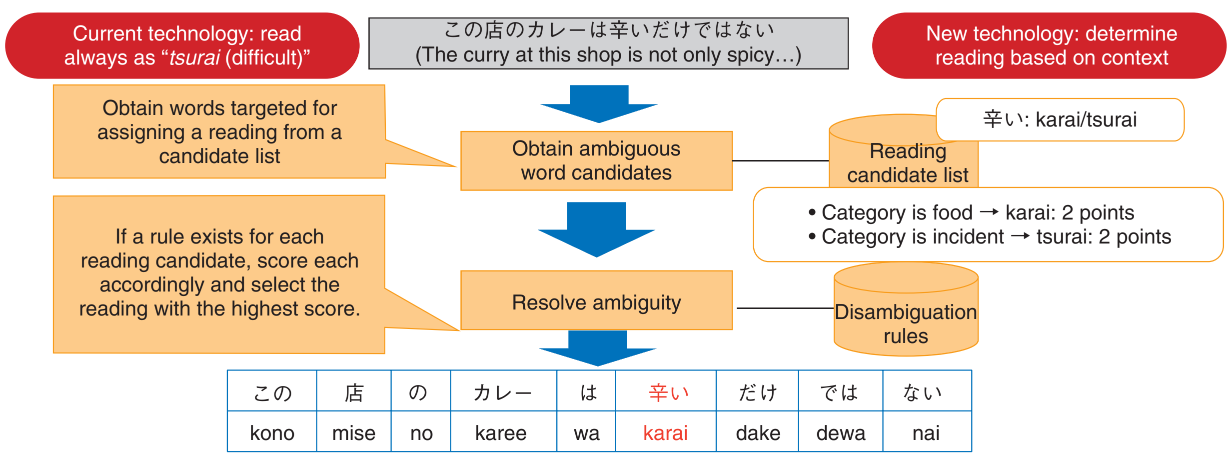


Fig. 2. Overview of word-reading disambiguation technology.

inferred reading and accents (**Fig. 1**). In the text-analysis section, an erroneous inference of the reading or accents of characters can result in the transmission of incorrect information to the listener of that synthesized speech, so there is a need for high-accuracy inference of reading and accents with respect to input text. The Japanese language, however, includes heteronyms, which are words for which reading and accents differ according to context despite being written the same. For example, the word "辛い" can be pronounced "*karai* (spicy)" or "*tsurai* (difficult)" and the word "寒気" can be pronounced "*samuke* (chills)" or "*kanki* (cold air)." For this reason, achieving high-accuracy inference of reading and accents has become a major objective.

To this end, we developed word-reading-disambiguation technology that infers the correct reading against an obvious misreading. This technology can infer the reading of a word having ambiguity through the use of dictionaries and rules that make use of linguistic knowledge. For example, a rule can be prepared beforehand that says if the word "カレー," which is pronounced "*karee* (curry)," should appear near the word "辛い," points are given to reading the latter as "*karai* (spicy)." Consequently, the word "辛い" appearing in the sentence "この店のカレーは辛いだけではない" (The curry at this shop is not only spicy…) would then be correctly read as "*karai*" (**Fig. 2**). Since it would be difficult to exhaustively write out all words considered to be associated with "*karai* (spicy)," an alternative means would be a framework that uses word categories (e.g., "food") as rules. Therefore, we can reduce the number of rules while improving comprehensiveness. This technology can infer readings with high accuracy while saving on memory.

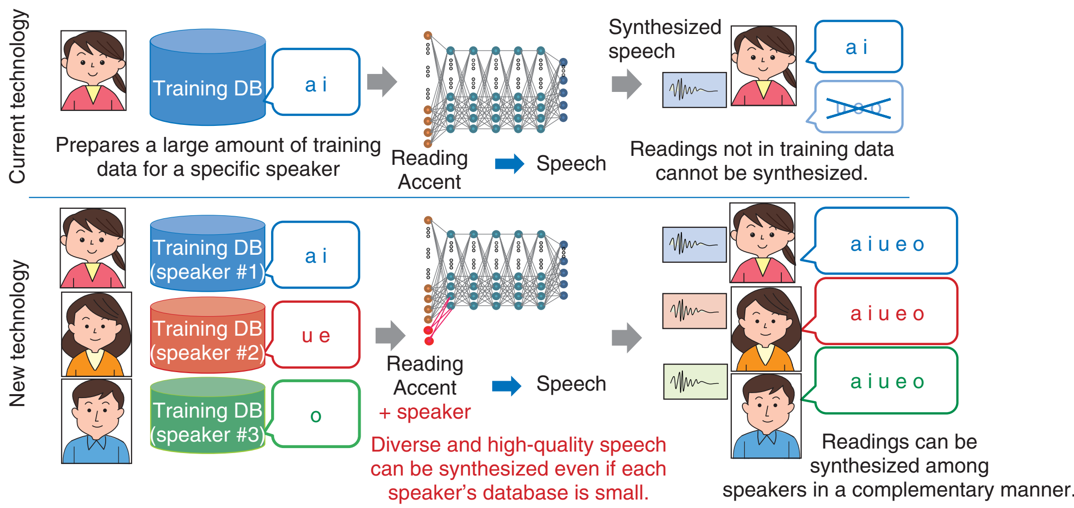


Fig. 3. DNN speech-synthesis technology for reproducing diverse speaker characteristics.

2.2 DNN speech-synthesis technology for reproducing diverse speaker characteristics at low cost

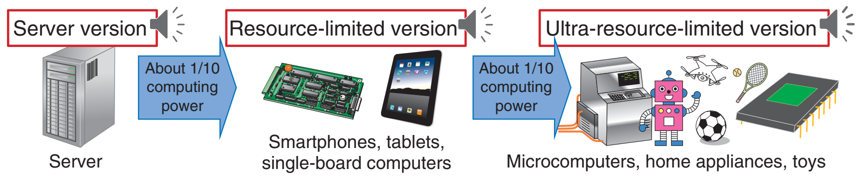
In contrast to the text-analysis section that is required to infer the reading and accents of input text with high accuracy, the speech-synthesis section is required to reproduce with high accuracy the voice of a desired speaker according to the wishes of the customer or other factors. However, achieving high-quality speech synthesis for a desired speaker requires a large amount of speech data uttered by that speaker. For example, in Cralinet [1], a unit-selection speech-synthesis system, up to 20 hours of speech data are needed to generate high-quality synthesized speech. As a result, achieving speech synthesis of a variety of speaker characteristics in an interactive speech system requires speech recordings and database (DB) construction, which can be a major issue due to the costs involved.

In response to this issue, we have achieved high-quality speech synthesis of a desired speaker from 20–30 minutes of speech data (about 2 hours of speech recordings) by using a previously constructed multi-speaker speech DB together with a DNN. A key feature of this system is that the speech data of multiple speakers can be modeled with a single DNN (**Fig. 3**). The information needed to generate speech (readings and accents) is learned from previously prepared multi-speaker speech data, and features of the desired speaker, such as voice quality and speaking style, are learned from the speech data of that desired speaker. Therefore, we have achieved high-quality speech synthesis even with a relatively small amount of speech data of the desired speaker [2]. Additionally, by combining this system with generative adversarial networks that have been found to be effective in image generation and other tasks, we have achieved improvements in the quality of synthesized speech and in the reproducibility of a speaker's voice [3].

2.3 DNN speech-synthesis technology for diverse environments

Depending on the environment in which speech synthesis is actually being used, a variety of constraints (e.g., no network access, need for high-speed response) require that the system be operated not on a computer server having abundant computing resources (central processing unit (CPU), read only memory (ROM), random access memory (RAM), etc.) but rather on devices such as smartphones or robots that are very limited in such resources. To address this issue, we developed an embedded type of DNN speech-synthesis library that can operate at practical speeds on devices with limited computing resources while maintaining the quality of synthesized speech to the extent possible. Specifically, we created a lineup of three types of libraries. In addition to the server-specific library, we developed a library for resource-limited terminals for operation on devices such as smartphones and tablets and a library for ultra-resource-limited terminals for devices such as microcomputers, home appliances, and high-grade toys that are greatly limited in computing resources

Lineup of 3 versions according to resources, specifications, application, etc.



	Server version	Resource-limited version	Ultra-resource-limited version
Features	High-quality speech synthesis, multilingual support (Japanese, English, Chinese, Korean)	Lightweight speech-synthesis library enabling high-quality speech synthesis for smartphones, etc.	Ultra-lightweight speech-synthesis library enabling operation on low-end CPUs having only integer arithmetic
CPU	x86_64	ARM Cortex-A series	ARM9, MIPS1, etc.
Memory	2GB (ROM)/4GB (RAM)	About 10 MB (ROM/RAM)	Up to 10 MB
Applications	Content creation, robots, interactive speech, entertainment, information processing, and more	Robots requiring high-speed response, use in environments with no network access	Use on inexpensive, low-end devices

Fig. 4. Speech-synthesis engine operable on diverse devices.

(**Fig. 4**).

Many microcomputers and similar devices have no built-in floating-point unit (FPU), so the issue was how to speed up DNN-inference processing. In the library for ultra-resource-limited terminals, we used fixed-point arithmetic to achieve high-speed DNN-inference processing without using floating-point operations. Additionally, by incorporating speed-enhancement measures into the text-analysis section, the library for ultra-resource-limited terminals can operate at high speeds with small memory (ROM: 7 MB or greater) even on devices equipped with no FPU.

3. Application examples

3.1 Application to news-reading speech spoken by computer-generated (CG) announcer

FutureVoice Crayon [4] from NTT TechnoCross provides the text-to-speech-synthesis technology we developed as a service. It has been used since February 2020 for synthesizing the speech of TV Asahi's "AI×CG Announcer Hanasato Yuina" [5]. This is a news program, so there is a need for rich expressive power close to that of a human announcer as well as the ability to correctly read news copy in various categories. The word-reading-disambiguation technology described above automatically assigns readings that fit the category of that news, which helps to reduce the work involved in correcting readings and accents that up to now has been performed by humans.

Note that the voice of the CG announcer in this example took form by mixing the voices of several announcers at TV Asahi. This project therefore created an original voice independent of the rights of a specific person, which demonstrates new possibilities of speech-synthesis technology.

3.2 docomo AI Agent API

NTT DOCOMO's "docomo AI Agent API" [6] is an interactive-type AI-based application service provider service that packages a speech/text user interface (UI). This service incorporates our speech-synthesis technology as its speech-synthesis engine. The application programming interface (API) includes more than 50 types of voices as preset speakers, which enables a user to implement voice UIs that fit a variety of characters and environments without having to go through the trouble of recording speech or managing rights for each UI. The DNN speech-synthesis technology described above enables the reproduction of speaker characteristics with much variation ranging from the voices of elementary school students to those of the elderly.

3.3 Disaster Mitigation Communication System

NTT DATA's Disaster Mitigation Communication System [7] is an announcement broadcasting system for transmitting administrative and disaster-prevention information from local governments to residents. It delivers this information from a transmission system inside a government building or from remote operation terminals to outdoor speaker equipment installed within the community and to tablets, smartphones, mobile phones, and other devices. The information delivered in this manner is then used as a basis for conducting speech synthesis from the outdoor speaker equipment and each device, such as a tablet or house-specific receiver, then conveying that information to local residents with synthesized speech.

4. Future developments

This article described the recent technological developments in speech-synthesis technology at NTT Media Intelligence Laboratories and presented examples of its application. Thanks to these efforts, speech-synthesis technology has reached a certain level from the viewpoint of generating synthesized speech in the voice of a desired speaker.

Nevertheless, if we were to compare the output of current speech-synthesis technology with human speech, we would find that there are still significant differences. For example, when reading the news or a script in the case of a television announcer or voice actor, the goal would be to understand, for example, the intent included in the text and to then express that text in a manner of speaking that includes emotion. However, current speech-synthesis technology is incapable of interpreting intent and can only generate synthesized speech with the same tone. As expectations are growing for means of supporting and substituting human activities, the further spread of speech-synthesis technology in society will require speech synthesis that can achieve a level of expression equivalent to or better than that of humans. Going forward, we aim to expand the range of speech-synthesis applications by pursuing technologies that can enable expression based on context, intent, and emotion and expression that take into account the attributes and receptiveness of the listener.

References

[1] K. Mano, H. Mizuno, H. Nakajima, N. Miyazaki, and A. Yoshida, "Cralinet—Text-to-Speech System Providing Natural Voice Response to Customers," NTT Technical Review, Vol. 5, No. 1, pp. 28–33, 2007. https://ntt-review.jp/archive/ntttechnical.php?contents=ntr200701028.pdf

[2] N. Hojo, Y. Ijima, and H. Mizuno, "DNN-based Speech Synthesis Using Speaker Codes," IEICE Trans. Inf. and Syst., Vol. E101-D, No. 2, pp. 462–472, 2018.

[3] H. Kanagawa and Y. Ijima, "Multi-speaker Modeling for DNN-based Speech Synthesis Incorporating Generative Adversarial Networks," Proc. of the 10th ISCA Speech Synthesis Workshop, pp. 40–44, Vienna, Austria, Sept. 2019.

[4] FutureVoice Crayon (in Japanese), https://www.futurevoice.jp/

[5] TV Asahi's "AI×CG Announcer Hanasato Yuina, https://news.tv-asahi.co.jp/news_international/articles/000175834.html

[6] docomo AI Agent API (in Japanese), https://docs.sebastien.ai/

[7] Disaster Mitigation Communication System (in Japanese), https://www.nttdata.com/jp/ja/lineup/disaster_mitigation_c/

Yusuke Ijima

Distinguished Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. in electric and electronics engineering from National Institution for Academic Degrees and University Evaluation from Yatsushiro National College of Technology, Kumamoto, in 2007, and an M.E. and Ph.D. in information processing from Tokyo Institute of Technology in 2009 and 2015. He joined NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) in 2009, where he engaged in the research and development of speech synthesis. His research interests include speech synthesis, speech recognition, and speech analysis. He received the Awaya Prize Young Researcher Award of the Acoustical Society of Japan (ASJ) in 2018. He is a member of ASJ and the International Speech Communication Association (ISCA).

Nozomi Kobayashi

Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

She received an M.E. and Dr.Eng. in information science from Nara Institute of Science and Technology in 2004 and 2007. She joined NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) in 2007. Her current research interests include natural-language processing, especially information extraction. She is a member of the Information Processing Society of Japan (IPSJ) and the Association for Natural Language Processing (NLP).

Hiroko Yabushita

Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

She received a B.S. and M.S. in information science from Ochanomizu University, Tokyo, in 2008 and 2010. She joined NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) in 2010. She engaged in the research of image recognition and received the IEEE Japan Council Women in Engineering Best Paper Award in 2013. Her current interest is in the digitization of human capabilities. She is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Takashi Nakamura

Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received an M.E. in information and communication engineering from the University of Electro-Communications, Tokyo, in 2005 and joined NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) the same year. His interests include the practical development and business development of speech synthesis. He is a member ASJ.

Speech Recognition Technologies for Creating Knowledge Resources from Communications

Yuichi Nakazawa, Takeshi Mori, Yoshikazu Yamaguchi, Yusuke Shinohara, and Noboru Miyazaki

Abstract

Speech recognition technologies have been used in an increasing number of situations to support or substitute work traditionally done by people such as call analysis in contact centers and support for creating minutes of council meetings. We are promoting research and development of speech recognition technologies because we believe that speech significantly supports human activities, especially in business. In this article, we describe the historical development of speech recognition technologies cultivated by NTT laboratories and the future contributions and roles of such technologies in human and corporate activities. It also introduces the use of non-linguistic information read from voice, such as emotion, gender, and age, which has been attracting attention.

Keywords: speech recognition, digital transformation, human-to-human communications

1. Development of speech recognition technologies

"Hey, Siri," "Ok, Google." These are the first words you speak to your voice assistant, and no doubt most people have used them. Speech recognition technologies have spread rapidly with the advent of voice assistants that control devices and provide information to people when speaking into a smartphone or artificial intelligence (AI) speaker. Various speech recognition technologies that enable such dialogue between people and computers have been put into practical use such as in automatic voice response devices (interactive voice response: IVR) in the 1980s and car navigation in the 1990s. However, with the recent introduction of deep-learning technologies, speech recognition accuracy has improved significantly, and the use of voice assistants and speech translation combined with machine translation due to increasing demand for it in the globalized world, has begun.

Nevertheless, speech is an important means of communication between people. Although the abovementioned voice assistants work with relatively short utterances, commercialization of speech recognition for long utterances (long sentences, conversations) has also been studied. Since 2000, speech recognition technologies have been used in situations in which manuscripts were at hand and targeted speech was comparatively clear, such as captioning for news programs and support for council-minutes creation. However, the application of speech recognition technology for speech occurring in natural communications between people, such as analysis of operator-customer conversations at contact centers and real-time conversational support, has recently been advancing. Thus, speech recognition technologies have been developed with improved recognition accuracy and a wider variety of targeted speech (**Fig. 1**).

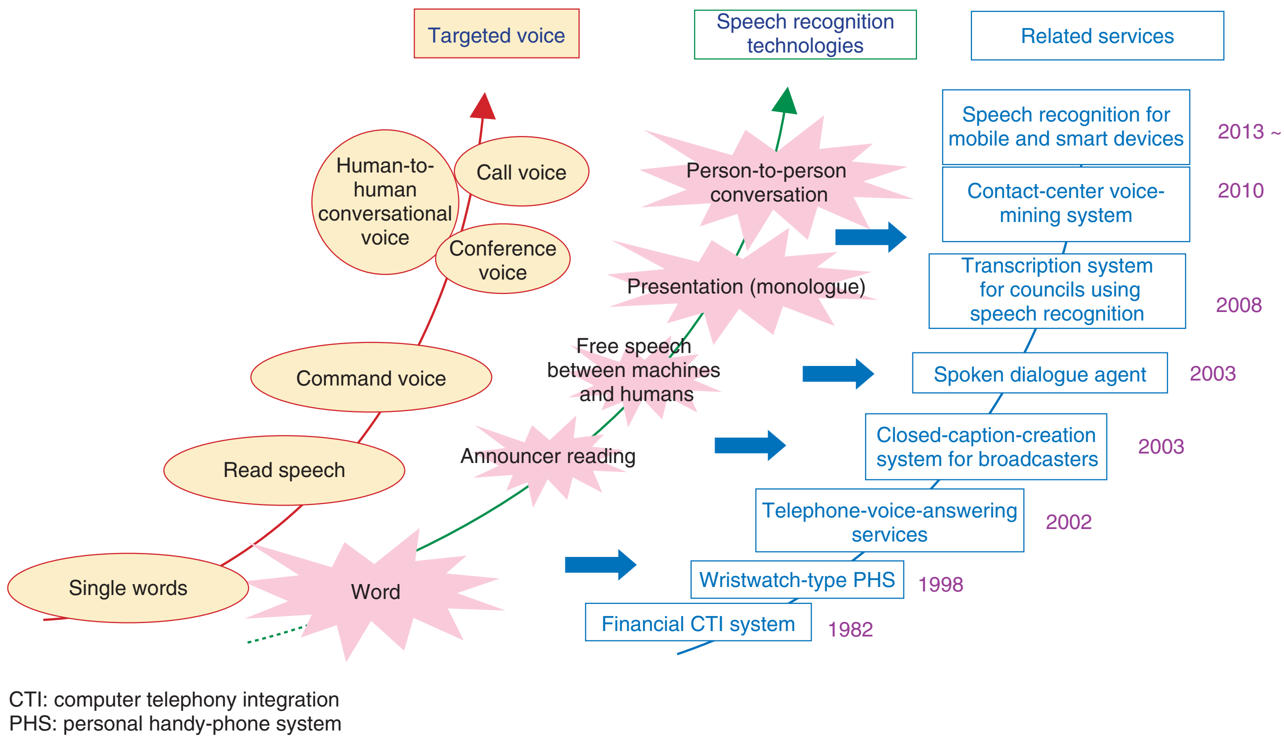


Fig. 1. NTT speech recognition technology initiatives.

2. Role of speech recognition technologies in supporting digital transformation in business

We believe that by further expanding the range of speech targeted with speech recognition technologies, such technologies will be able to play a role in driving the transformation of corporate activities. Using information technology (IT), including AI technologies, to transform business processes is known as digital transformation (DX). The significance of DX is gaining attention regardless of industry or business type. However, advancing DX requires effort such as streamlining and automating business processes using IT and creating new value through seamless collaboration between business and IT. Speech recognition technologies have demonstrated their effectiveness in streamlining and automating business processes.

We use a large amount of speech when communicating with others. Although communications tools using text, such as social networking services, email, and chat, have advanced, real-time voice communications face-to-face or via telephone are probably more often chosen to confirm or convey complex content or for decision making that requires the consensus of many people.

Compared to text communications, voice communications are advantageous because they are in real time and convey unspoken information expressed through the nuances of voice. However, if recordings or notes are not made, information communicated by voice is volatile, i.e., as soon as it has been uttered it disappears. Although an enormous amount of voice communications occurs on a daily basis in corporate activities, currently only a limited amount of voice communications is targeted for data analysis such as calls between operators and customers at contact centers. The data in most voice communications are not used.

However, a large amount of voice communications that occurs face-to-face or via telephone between sales representative and customers contain information that could be useful from a variety of perspectives such as marketing or compliance management. Information from meetings and communications among employees (e.g., short consultations), is also useful for improving business activities. Examples of such information are seeds of new business ideas, business improvement tips, or employees' mental health conditions. Preserving such information by turning it into text using speech recognition technologies to create knowledge resources for various processes, thus leading to business improvements, has the potential to contribute to the rationalization and automation of business processes. In the next section, we introduce example applications of speech recognition technologies to streamline business processes.

3. Use cases of speech recognition technologies

Compared to current speech recognition technology, research and development (R&D) in speech recognition technologies is now addressing broken utterances, which adds another layer of technical difficulty, enabling progress in business DX in an increasing number of situations. We introduce example use cases that are becoming more widespread due to improvements in speech recognition accuracy regarding broken utterances.

3.1 Conference speech recognition

Even though minutes are often taken in business meetings, many people who take minutes often feel that rather than quick notes, it requires much more work than predicted to actually record meetings. Attempting to faithfully take notes in a meeting means the minutes taker will not be able to properly participate in the meeting or its discussions or will worry whether the minutes will be accurately recorded sometime after the meeting. Also, recording all of a meeting and creating minutes while listening later takes more time than the meeting itself, so it might be a good to have an additional minutes taker participate in the meeting. Many people wish for such a device that can automatically and quickly create minutes.

Current speech recognition is not sufficiently accurate and is limited to searching speech tied to time information in the hope that important words can be recognized. However, in addition to supporting simple minutes creation, it is hoped that by enabling speech recognition for broken utterances, AI will become proficient in roles usually played by humans (facilitators), such as automatic creation of minutes in conjunction with summarization technology, linkage with issue-tracking systems by automatic extraction of action items, and facilitation of meetings.

3.2 Remote work support

Where remote operations and responses are expected to grow such as in healthcare and education, inconveniences that occur due to interactions not being face-to-face will have to be eliminated. In healthcare, for example, although it is technically possible to operate remote devices with buttons and levers as well as recording conversations, unobtrusive AI support technologies will be required so that the doctor, who only obtains a lower-than-normal amount of information about the patient through a screen, can focus on treatment. Such technologies could be used to pass a thermometer in response to “let’s take your temperature” or automatically control the lighting inside the patient’s mouth in response to “please open your mouth.” Linking with dialect-conversion technology also holds promise for smooth communications in remote healthcare in rural areas where strong dialects are spoken.

In one-to-many classes, which are fundamental in education, student levels of understanding are understood through calling on all students for their responses. In online classes, however, it is clear that cross-talk occurring in such situations makes it difficult to establish voice communications. Unobtrusive AI that will prevent device operations from interfering with students’ concentration, such as comprehending students’ speech with real-time speech-to-text recognition and executing raised hand commands for student responses of “yes,” will likely become as indispensable.

3.3 Contact centers

In contact centers, speech recognition technologies are being used for operator speech. If speech recognition results of customer speech can also be obtained with sufficient accuracy, speech recognition technologies could contribute to making business support more efficient, reducing operator tasks, increasing the number of calls that can be handled, and improving customer satisfaction to meet future increasing demand on contact centers as various services go online.

4. Use of non-linguistic information

As well as lingual information (text), non-linguistic information (gender, age, etc.) and para-linguistic information (emotions, intent, attitude) are contained in information conveyed through voice communications. Hence, there is demand to use such information to advance speech services in actual business.

We developed RexSense™, a software engine that can extract non-linguistic and para-linguistic information in speech, by studying technologies for recognizing and using non-linguistic and para-linguistic information as well as taking initiatives to recognize

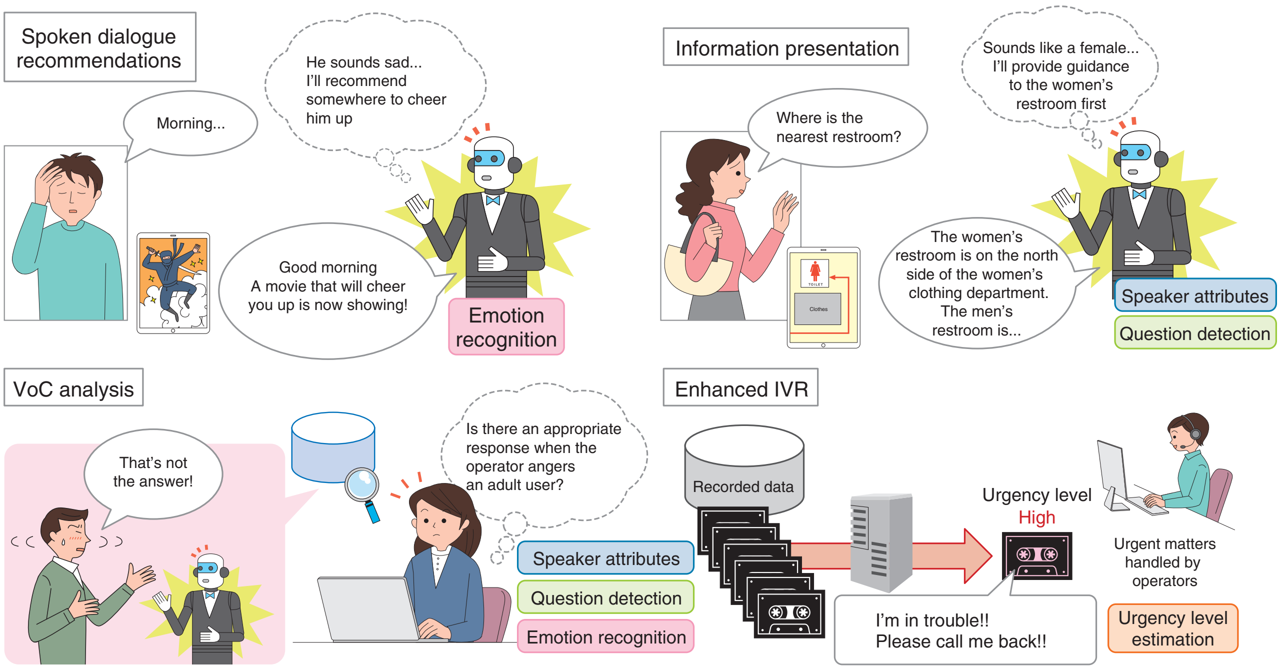


Fig. 2. RexSense™ application examples.

text information from speech with high accuracy. This software engine enables recognition and estimation of (1) speaker attributes (adult male, adult female, child), (2) emotions (joy, anger, sadness, calmness), (3) questions and non-questions, and (4) urgency, with high accuracy from voice data. We also developed the RexSense system to enable web application programming interface (Web API) services integrating this software engine with speech recognition for uses such as contact center advancement. RexSense also makes it possible to provide advanced services with robots by giving appropriate responses and recommendations according to human emotions or to provide advanced digital signage that presents more appropriate content (guidance, advertisements, etc.) based on non-linguistic information such as speaker attributes determined from voice. We expect the implementation of advanced voice of customer (VoC) analysis in contact centers, more sophisticated automated response services in IVR, and provision of more advanced audio conferencing solutions using non-linguistic and para-linguistic information (**Fig. 2**).

We also developed a customer-satisfaction-estimation technology for analyzing customer-voice characteristics and various conversation characteristics from the voices of operators in contact centers and customers to extract customer satisfaction (satisfaction and dissatisfaction) and introduced it to the contact center AI solution ForeSight Voice Mining™; commencing services in April 2019. We are developing a response-likability estimation technology for evaluating the likability of operator responses and considering putting it into service. These technologies hold promise for applications, such as call analysis (searching for good operator-response examples, analysis of customer satisfaction, etc.), operator support, operator and contact center evaluation, and operator education.

5. Future outlook

By expanding application areas from business scenarios and targeting all types of voice communications, the speech recognition technologies introduced in this article are crucial for achieving human Digital Twin Computing (DTC) [1], one element of the Innovative Optical and Wireless Network (IOWN) promoted by the NTT Group. In the cyber-physical interaction layer in the DTC architecture (**Fig. 3**), it will be necessary to collect the data required to generate digital twins by sensing real-space things (objects) and humans, and speech recognition technologies will play an important role in sensing human

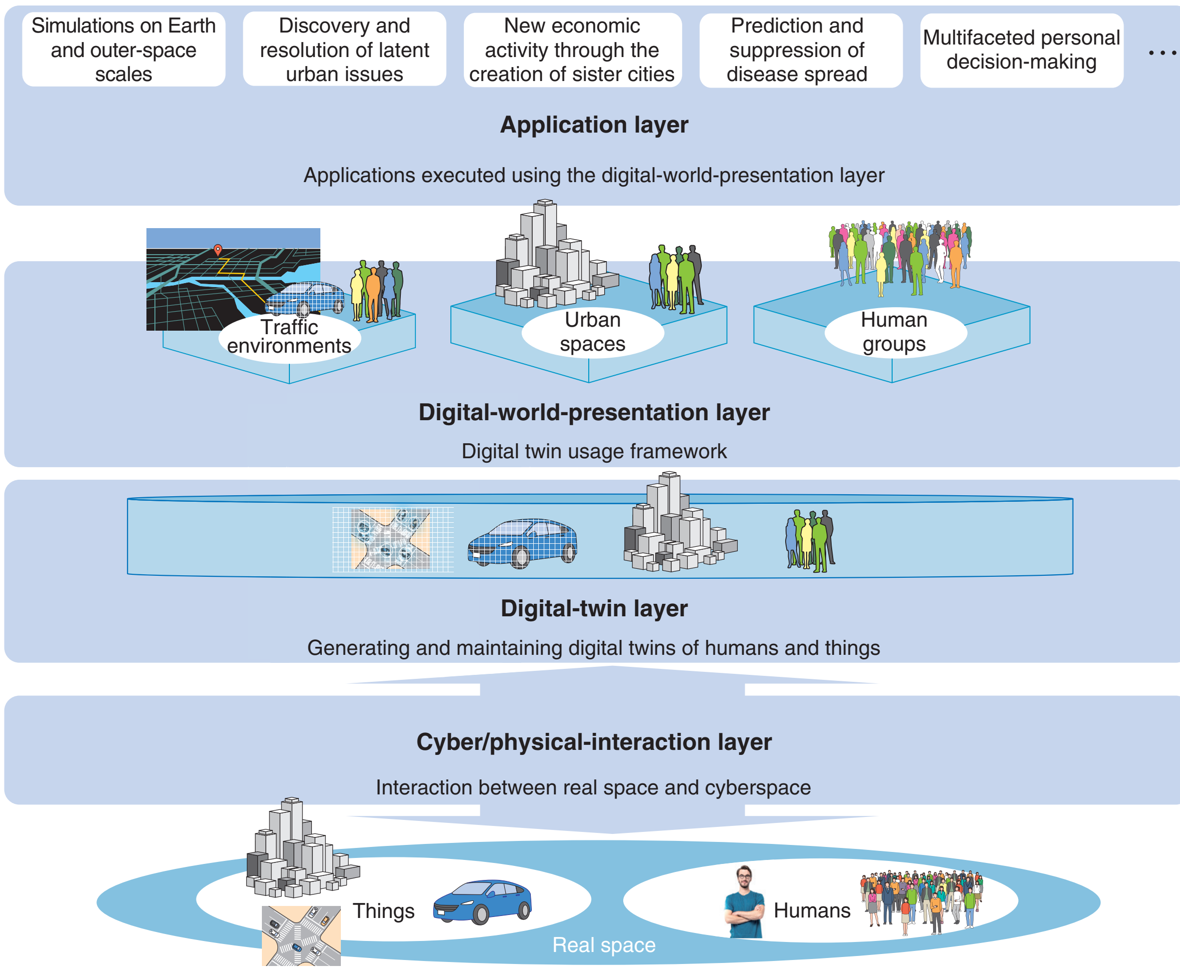


Fig. 3. Digital Twin Computing architecture.

thoughts.

Societies will become more convenient, richer, and safer as human DTC and DX of corporate activities progress. We will contribute to the creation of such a society through R&D of speech recognition technologies for human-to-human communications.

Reference

[1] I. Toshima, S. Kobashikawa, H. Noto, T. Kurahashi, K. Hirota, and S. Ozawa, "Challenges Facing Human Digital Twin Computing and Its Future Prospects," NTT Technical Review, Vol. 18, No. 9, pp.19–24, Sept. 2020.
https://ntt-review.jp/archive/ntttechnical.php?contents=ntr202009fa2.html

Yuichi Nakazawa

Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. from Keio University, Kanagawa, in 2001 and 2003. He joined NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) in 2003 and has been involved in R&D of automatic speech recognition technologies. He is a member of Acoustical Society of Japan (ASJ).

Takeshi Mori

Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. from Tokyo Institute of Technology in 1994 and 1996 and a D.E. from University of Tsukuba, Ibaraki, in 2007. Since joining NTT in 1996, he has been engaged in research on speech and audio processing algorithms. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), ASJ, and the Institute of Electronics, Information and Communication Engineers (IEICE).

Yoshikazu Yamaguchi

Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. in electrical engineering from Osaka Prefecture University in 1993 and 1995. He joined NTT Human Interface Laboratories (now NTT Media Intelligence Laboratories) in 1995 and has been involved in R&D of automatic speech-recognition technologies.

Yusuke Shinohara

Senior Research Engineer, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.S. and M.S. in mechanoinformatics from the University of Tokyo in 2002 and 2004. From 2004 to 2017, he was a research scientist at Toshiba Corporate Research and Development Center, Kanagawa. Since joining NTT in 2017, he has been involved in R&D of speech recognition, in particular, acoustic modeling with various types of deep neural networks. His research interests include speech recognition, computer vision, and machine training. He received the IEICE Pattern Recognition and Media Understanding (PRMU) Young Researcher Award in 2004 and the ASJ Awaya Prize Young Researcher Award in 2013. He has served as a committee member (2013–2015, 2017-present) and secretary (2015–2017) for the Information Processing Society of Japan Special Interest Group (IPSJ SIG) on Spoken Language Processing. He is a member of IEEE, IEICE, ASJ, and IPSJ.

Noboru Miyazaki

Senior Research Engineer, Supervisor, Cognitive Information Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. from Tokyo Institute of Technology in 1995 and 1997. Since joining NTT in 1997, he has been engaged in research on spoken-dialogue systems, speech synthesis, and speech-recognition technologies. He received the IEICE Inose Award in 2001. He is a member of ASJ, IEICE, and the Japanese Society for Artificial Intelligence (JSAI).

Knowledge and Language-processing Technology that Supports and Substitutes Customer-contact Work

Kyosuke Nishida, Kuniko Saito, Tetsuo Amakasu, Kazuyuki Iso, and Shuichi Nishioka

Abstract

At NTT Media Intelligence Laboratories, using natural-language-processing technology cultivated over many years as one of our core competences, we are researching and developing knowledge and language-processing technology that contributes to improving the productivity of contact centers and offices. In this article, a language model, document-summarization technology, and interview-support technology are introduced.

Keywords: AI, document summarization, response analysis

1. Introduction

At NTT Media Intelligence Laboratories, we have researched and developed knowledge and language-processing technology for contact centers for analyzing documents (such as work manuals and frequently asked questions) and presenting appropriate documents to agents handling customers. Not only in contact centers but also in offices, it is necessary to further improve productivity of agents and employees. In response to this necessity, we are developing technology for understanding and generating large-scale documents and various responses. We first explain a language model for handling documents and document-summarization technology using this language model then describe interview-support technology for analyzing responses between a customer and agent.

2. Development of natural-language understanding by using a language model (BERT)

It has been considered difficult for artificial intelligence (AI) to understand human language. However, with the advent of Bidirectional Encoder Representations from Transformers (BERT) [1], announced by Google in October 2018, the research and development (R&D) of natural-language understanding has undergone a large paradigm shift. For example, regarding the task called *machine reading* [2], which requires reading comprehension to understand the content of text and answer questions, it has been reported that AI using BERT greatly exceeded the response score of a human. Performance of natural-language-processing tasks other than machine reading has also improved significantly, and language models are attracting attention as a basic technology for giving AI the ability of language comprehension.

A language model estimates the plausibility of sentences (**Fig. 1**). For example, many people may feel that “cake” is more natural than “egg” as the word “X” in the sentence “Today is my birthday, so I ate [X].” Also, it seems natural that the two sentences “It’s a fine day today.” and “It’s a good day to do the laundry.” appear consecutively. BERT learns in advance such fill-in-the-blank problems (word prediction) and relationships between two consecutive sentences (prediction of the next sentence) on the basis of a large set of texts such as all Wikipedia

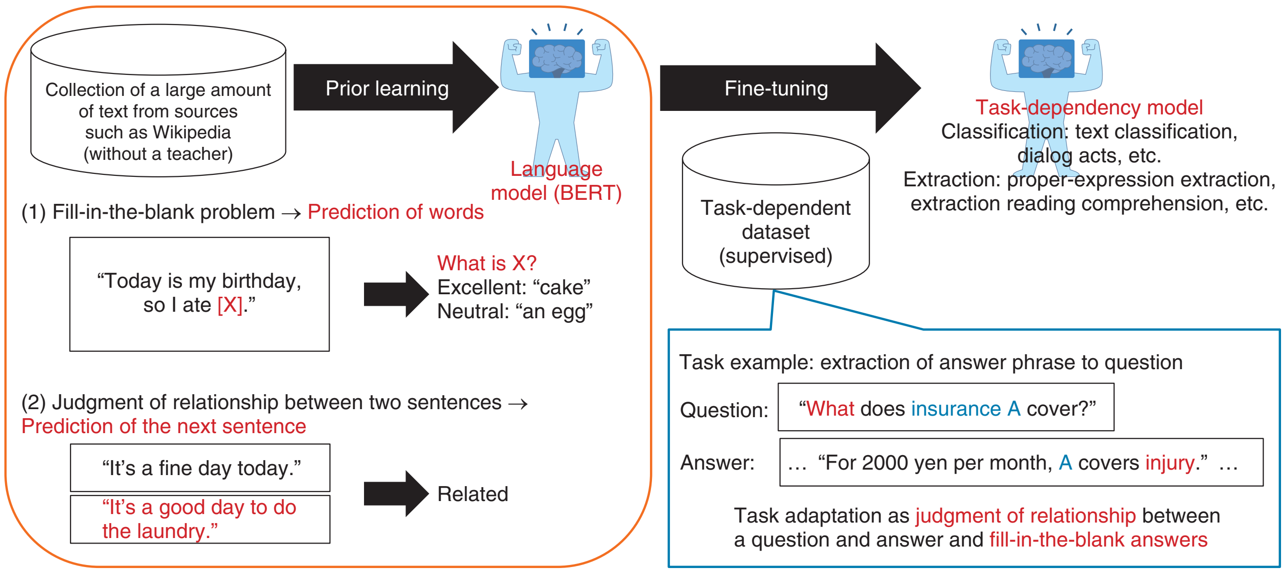


Fig. 1. Language model BERT.

sentences. By learning (fine-tuning) with various task-dependent datasets, the language model (BERT) obtained in this manner can be applied to various tasks (such as classifying texts by genre and extracting phrases that will answer questions) and achieve high performance even when a large amount of learning data is not available for the applied task.

BERT has had a great impact on natural-language processing, and language models are still being researched and constructed all over the world. At NTT Media Intelligence Laboratories, while collecting a large amount of Japanese text data and creating a Japanese-language BERT, we are researching technologies that use language models for tasks such as summarizing documents [3, 4], retrieving documents [5], and answering questions [6, 7]. In each case, high performance is achieved by not only simply applying BERT but also using the knowledge we have gained through our previous research on natural-language processing and deep learning. Moreover, by investigating the characteristics and internal operation of BERT [8], we are researching with the aim of constructing our own language model that remedies the shortcomings of BERT.

3. Document-summarization technology for summarizing documents by specifying length

Document-summarization technology is introduced as a representative example of a technology that uses the language model described above. Document summarization has been grappled with for many years. For customer-contact work such as at contact centers, if AI returns a long sentence as a result of searching previous calls and answering a customer's question, it will be difficult for the customer to read such a long sentence; accordingly, it is desirable to appropriately adjust the length of the sentence.

Given the above-mentioned issue, NTT Media Intelligence Laboratories developed a document-summarization technology that can control sentence length by using a neural network [3]. This technology consists of a combination of an extractive model (which identifies important points in a sentence) and abstractive model (which generates a summary sentence from the original sentence). The extractive model learns on the basis of the language model. By controlling the number of important words output by the extractive model according to the specified sentence length and by generating a summary sentence based on both the important words and the original sentence, it was possible to establish a technology that can control sentence length and summarize with high accuracy.

This document-summarization technology is used as the core engine of the COTOHA™ API summarization function developed by NTT Communications [9]. COTOHA Summarize offers a service for outputting a summary text when a document is input and provides our customers a free tool for generating summary texts of sites browsed with a web browser. We plan to further develop this technology for the NTT Group.

We will continue to advance R&D to improve this summarization technology, which includes not only document summarization but also summarization that targets dialogue and summarization that allows the viewpoint and keywords of the summarization to be specified externally. To further strengthen the competitiveness of our language model, we plan to continue technological development while incorporating the results of our latest research on language processing. Such work will include scaling up the model, constructing a generative language model for generating more-natural sentences, and establishing summarization technology based on that model [4].

4. Technology for using knowledge obtained from contact-center calls

4.1 Challenges in supporting customer contacts

NTT Media Intelligence Laboratories developed an automatic knowledge assistance system [10]. This is a technology that automatically retrieves and presents documents according to the content of a conversation with a customer. By supporting the agent with knowledge and responding promptly to customers with appropriate information, it became possible to improve relationships with customers.

However, contact centers are required to play a new role as the business style changes due to the digital transformation (DX) of offices and the spread of the novel coronavirus. One of these roles is as a sales method called *inside sales*. Inside sales is explained as follows. In contrast to the conventional sales method, namely, a full-time salesperson conducts face-to-face sales to understand customer needs and conclude business negotiations, inside sales is based on (i) maintaining continuous communication with the other party while gathering information (such as understanding needs) that triggers business negotiations via telephone or web conference and (ii) dispatching a salesperson when the possibility of receiving an order or contract increases. The number of contact centers that handle customers by using inside sales is increasing. The challenges facing contact centers playing this role are as follows.

- Further improvement of agent or productivity: When creating a report after finishing a call, it is necessary to support the extraction of important information in the response from business negotiations in which the flow of conversation is complicated and the topics are diverse.
- Improvement of productivity of sales-information analysis: From the standpoint of the agents' supervisor, it is necessary to understand and analyze the likelihood of making contracts, the tendency of customer needs, and other conversational tendencies included in business negotiations between each agent and customers. Planning and agent operations are then improved on the basis of the analyzed information. It is thus necessary to support such analysis and improvement work.

4.2 Interview-support technology

To address the above-mentioned challenges, we developed interview-support technology. This technology consists of two elements (**Fig. 2**). The first specifies a series of responses to the customer for each section of the topic. The second extracts important utterances such as questions, answers, and explanations from the sections specified by the former.

A characteristic of a business conversation that draws out issues and requests concerning a customer is that the topic continuously changes according to the answers given by the customer. Since conventional technology cannot cope with such a dynamic situation, we developed a means of firmly determining the point at which the topic changes. This is the first feature of the interview-support technology. Sections of the same topic are specified by machine learning using characteristic expressions at the point of change of the topic appearing in a conversation, and the type of topic is acquired by machine learning using characteristic expressions in the topic. The second feature of this technology is that the utterances of questions and answers of agents and customers for each topic are extracted by machine learning using characteristic expressions included in important utterances such as questions and answers.

By using the text of a call that is divided into sections for each topic by using the interview-support technology, after the call, the agent can quickly identify the points at which he/she heard about the customer's needs and budget and create a report based on that information. Moreover, by collecting such information from many calls, it becomes possible to support the analysis of sales information. We aim to apply this technology to responses via voice but also via online chat (text messages), which has been increasing in popularity.

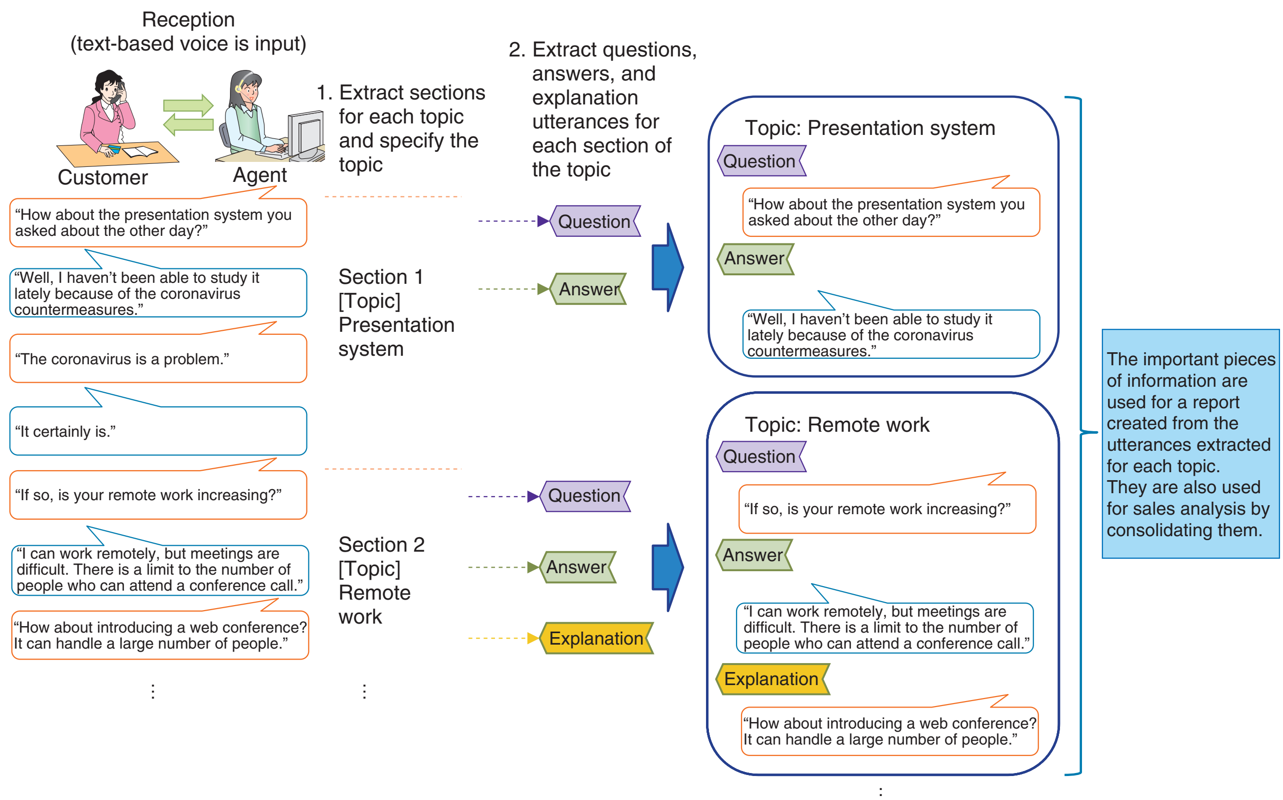


Fig. 2. Interview-support technology.

5. Concluding remarks

We introduced our language model, document-summarization technology, and interview-support technology. Going forward, we plan to work on two technologies for understanding and generating large-scale documents and various responses: (i) a technology that can read various document layouts and search for necessary information at high speed and high accuracy and (ii) a technology that supports strategic conversations by grasping the details of a conversation and extracting needs that customers may have difficulty noticing themselves.

References

[1] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proc. of NAACL-HLT 2019, Vol. 1, pp. 4171–4186, Minneapolis, Minnesota, USA, June 2019.

[2] K. Nishida, I. Saito, A. Otsuka, K. Nishida, N. Nomoto, and H. Asano, "Toward Natural Language Understanding by Machine Reading Comprehension," NTT Technical Review, Vol. 17, No. 9, pp. 9–14, Sept. 2019.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201909fa3.html

[3] I. Saito, K. Nishida, K. Nishida, A. Otsuka, H. Asano, J. Tomita, H. Shindo, and Y. Matsumoto, "Length-controllable Abstractive Summarization by Guiding with Summary Prototype," arXiv, 2001.07331, Jan. 2020.

[4] I. Saito, K. Nishida, K. Nishida, and J. Tomita, "Abstractive Summarization with Combination of Pre-trained Sequence-to-Sequence and Saliency Models," arXiv, 2003.13028, Mar. 2020.

[5] T. Hasegawa, K. Nishida, S. Kaku, and J. Tomita, "Sparse Contextual Sentence Representation for Fast Information Retrieval," Proc. of the 34th Japanese Society of Artificial Intelligence (JSAI 2020), Mar. 2020 (in Japanese).

[6] K. Nishida, K. Nishida, I. Saito, H. Asano, and J. Tomita, "Unsupervised Domain Adaptation of Language Models for Reading Comprehension," Proc. of the 12th Conference on Language Resources and Evaluation (LREC 2020), pp. 5392–5399, May 2020.

[7] K. Nishida, K. Nishida, I. Saito, H. Asano, and J. Tomita, "Explainable Machine Reading Comprehension," Proc. of the 26th Annual Meeting of the Association for Natural Language Processing (NLP 2020), pp. 1–19, Mar. 2020 (in Japanese).

[8] Y. Ohsugi, I. Saito, K. Nishida, H. Asano, and J. Tomita, "How Do Masked Language Models Perform When the Input Sequence Length Changes?", Proc. of JSAI 2020, Mar. 2020.

[9] COTOHA™ API, https://api.ce-cotoha.com/contents/about.html

[10] T. Hasegawa, Y. Sekiguchi, S. Yamada, and M. Tamoto, "Automatic Knowledge Assistance System Supporting Operator Responses," NTT Technical Review, Vol. 17, No. 9, pp. 15–18, 2019.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201909fa4.html

Kyosuke Nishida
Distinguished Researcher, Knowledge Media Project, NTT Media Intelligence Laboratories.
He received a B.E., MIS, and Ph.D. in information science and technology from Hokkaido University in 2004, 2006, and 2008 and joined NTT in 2009. He was a chief engineer at NTT Resonant from 2013 to 2015. He received the Yamashita SIG Research Award from the Information Processing Society of Japan (IPSJ) in 2015 and the Kambayashi Young Researcher Award from the Database Society of Japan (DBSJ) in 2017. His current interests include natural-language processing and AI. He is a member of the Association for Computing Machinery (ACM), IPSJ, the Association for Natural Language Processing (NLP), the Japanese Society for Artificial Intelligence (JSAI), the Institute of Electronics, Information and Communication Engineers (IEICE), and DBSJ.

Kuniko Saito
Senior Research Engineer, Supervisor, NTT Media Intelligence Laboratories.
She received a B.E. and M.E. in chemistry from the University of Tokyo in 1996 and 1998. She joined NTT Information and Communication Systems Laboratories in 1998. Her research focuses on natural-language processing and AI. She is a member of NLP and IPSJ.

Tetsuo Amakasu
Senior Research Engineer, Supervisor, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.
He received a B.E. and M.E. from Tohoku University, Miyagi in 1997 and 1999. Since joining NTT in 1999, he has been researching spoken dialogue systems, spoken language processing, and speech analytics. He joined NTT Communications in 2008 and was involved in the development of IP phone services on smartphones. He is a member of the Acoustical Society of Japan and JSAI.

Kazuyuki Iso
Senior Research Engineer, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.
He received a B.E. from Gunma National College of Technology in 1998, and an M.S. and Ph.D. in knowledge science from Japan Advanced Institute of Science and Technology, Ishikawa, in 2000 and 2020. He joined NTT Cyber Space Laboratories in 2000. He is currently a senior research engineer at NTT Media intelligence Laboratories. His research interests include communication systems and computer-supported cooperative work. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Virtual Reality Society of Japan (VRSJ), IPSJ, and IEICE.

Shuichi Nishioka
Senior Research Engineer, Supervisor, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.
He received a B.E. and Ph.D. in engineering from Yokohama National University, Kanagawa, in 1995 and 2005. He joined NTT in 1995. His current interests include data engineering and AI. He is a member of IPSJ and DBSJ.

Spatial-information Processing Technology for Establishing a 4D Digital Platform

Yasuhiro Yao, Kana Kurata, Naoki Ito, Shingo Ando, Jun Shimamura, Mayuko Watanabe, Ryuichi Tanida, and Hideaki Kimata

Abstract

The four-dimensional (4D) digital platform integrates various sensing data about humans, things, and environments in real time into high-precision spatial information, enabling fusion with various industries' platforms and future predictions. The platform achieves these by matching and integrating 4D information such as latitude, longitude, altitude, and time with high accuracy on an advanced geospatial-information database. For maintaining an advanced geospatial-information database with high accuracy and abundant semantic information, this article introduces two spatial-information processing technologies: (i) real-world digitalization to detect features from images and sparse/low-density 3D point clouds and (ii) 4D-point-cloud coding that efficiently stores and uses 3D data including time variations.

Keywords: 4D digital platform, real-world digitalization, 4D-point-cloud coding

1. What is a 4D digital platform?

The four-dimensional (4D) digital platform collects various sensing data concerning people, things, and environments in real time, and matches and integrates 4D information such as latitude, longitude, altitude, and time with high accuracy. In other words, it is a platform that enables data fusion with various industrial platforms as well as future prediction (**Fig. 1**). By combining the 4D digital platform and various data collected from Internet of Things, the platform makes it possible to determine the exact position of various mobile objects in geographic space, and various future predictions based on that data will become possible. We thus believe that it can be used in various applications such as smoothing road traffic, optimizing the use of urban assets, and maintaining social infrastructures.

Regarding the elemental technologies that make up the 4D digital platform, namely, spatial-information processing technologies necessary for constructing an advanced geospatial-information database with high accuracy and abundant semantic information, we are promoting the research and development of (i) real-world digitalization technologies for detecting features from images and sparse and low-density 3D point clouds and (ii) 4D-point-cloud coding technology for efficiently storing and using 3D data including time changes. We introduce both technologies and discuss the status of our efforts concerning them.

2. Real-world digitalization technologies

To construct an advanced geospatial-information database, it is necessary to maintain highly accurate 3D spatial information centered on roads; however, satisfying this requires enormous cost and effort. Specifically, a costly dedicated vehicle equipped with a sensing device called laser-imaging detection and ranging (LiDAR) and a manual process for generating

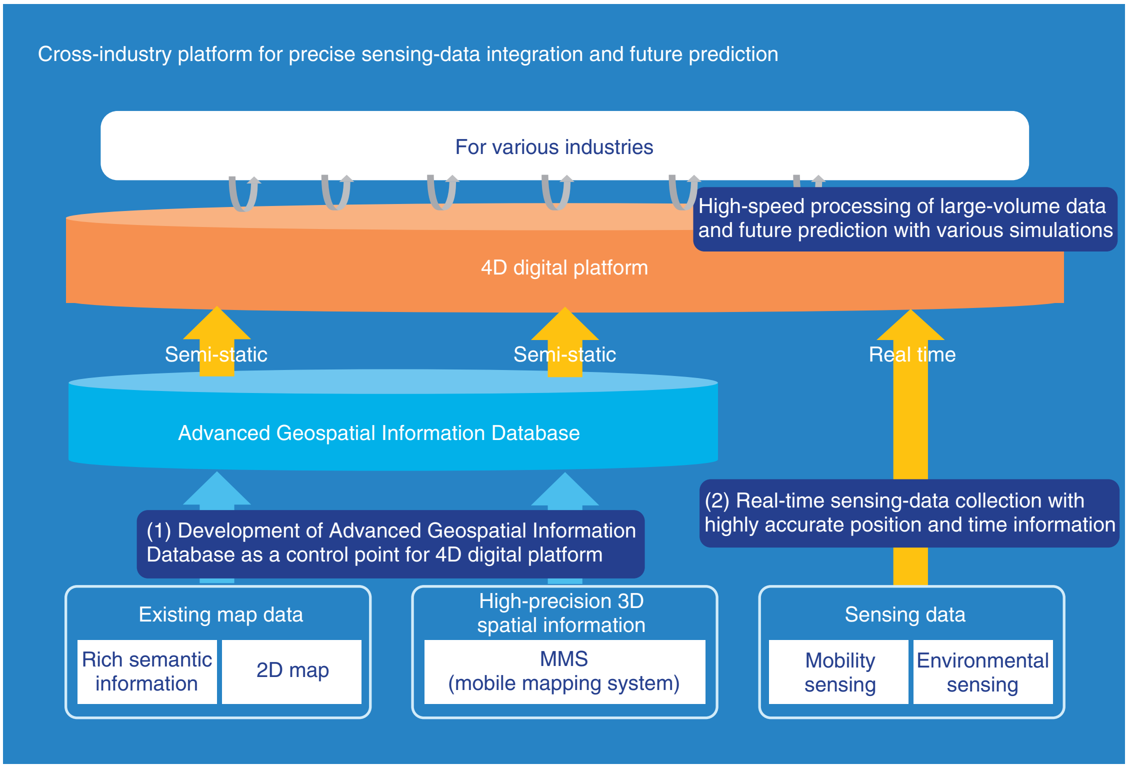


Fig. 1. Conceptual diagram of 4D digital platform.

maps are required.

To efficiently collect high-precision 3D spatial information, we are researching and developing real-world digitalization technologies that automatically and highly accurately detect various natural and artificial objects near roads from a combination of sparse and low-density 3D point clouds (measured with low-cost LiDAR) and images taken with a camera (**Fig. 2**).

Our real-world digitalization technologies include: (i) 3D-data high-resolution enhancement technology for generating high-resolution 3D data from sparse 3D point clouds based on images and videos; (ii) 3D-scene understanding and state-estimation technology for recognizing entire scenes (including complex shapes) and understanding states; and (iii) 3D-moving-object detection technology for using the sensing data obtained from moving-body sensors (in addition to 3D point clouds and images) to determine the position and orientation of vehicles and people. We also discuss our latest research achievements concerning our 3D-data high-resolution enhancement and 3D-scene understanding and state-estimation technologies.

2.1 3D-data high-resolution enhancement technology

The 3D-data high-resolution enhancement technology increases the resolution of 3D data, namely, a 3D point cloud with texture, from a combination of sparse and low-density 3D point clouds (measured by low-cost LiDAR) and images taken with a camera. Three-dimensional measurement with low-cost LiDAR provides sparse measurement results, and although 3D measurement is possible regardless of distance, the measurement results include noise. On the contrary, images taken with a camera contain dense data; however, stereoscopic 3D measurement using multiple images is not highly accurate for distant objects. We believe it is possible to generate 3D data with the same measurement accuracy as LiDAR and the same density as images taken with a camera while removing noise by processing the information from both LiDAR and camera data in an integrated manner.

We are working in stepwise on the research and development of the 3D-data high-resolution enhancement technology. Under the assumption that data are measured while a vehicle with in-vehicle sensors is

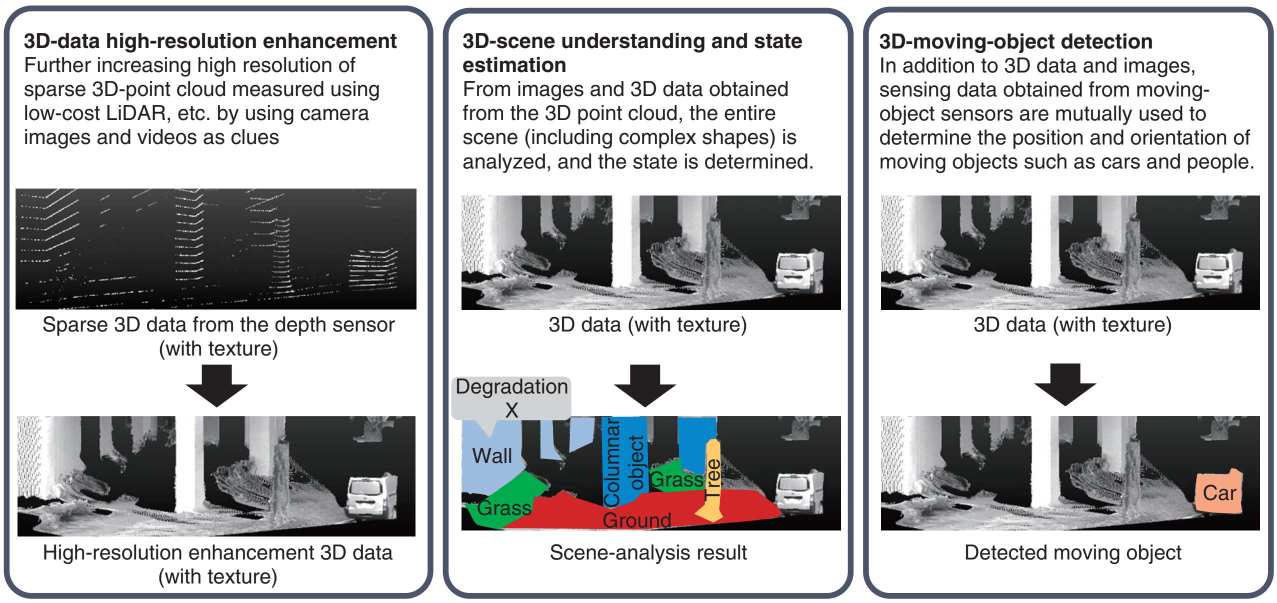


Fig. 2. Real-world digitalization technologies.

moving, the amount of information that we will integrate increases as follows for improving accuracy: (i) one image and one frame of LiDAR measurement data, (ii) multiple images and one frame of LiDAR measurement data, and (iii) multiple images and multiple frames of LiDAR measurement data that are continuous in a time series. Note that "one frame" of LiDAR measurement means data for one measurement over 360 degrees. Although it depends on the LiDAR equipment, LiDAR measures 360 degrees of the surroundings while rotating about 10 times per second.

We now introduce the 3D-data high-resolution enhancement technology, which derives high-density 3D data in real time without supervision from a single image and a sparse 3D point cloud measured by LiDAR.

First, a 3D point cloud measured by LiDAR is projected onto an image to generate an image that holds depth information called a depth map. The depth map created in this manner is referred to as a sparse depth map with many pixels that do not have depth values (**Fig. 3**).

This sparse depth map is processed using the input image as a clue to generate a dense depth map in which all pixels have a depth value. Such a method can be called *depth completion*. Even though depth completion had been developed, the conventional method generates a dense depth map for pixels that do not have depth by smoothly connecting the observed depth values [1]. This conventional depth completion is effective for complementing continuous observations between sparse observations; however, it smoothly interpolates the depth between different objects (Fig. 3).

To solve this problem, we previously proposed a method for smoothly connecting the observed depth values while adding a constraint that the change in depth becomes discontinuous when it straddles an object [2]. Compared with the conventional depth completion, our method not only improves accuracy but also obtains natural results when the depth map is visualized as 3D data (Fig. 3). In the future, we will further improve accuracy by increasing the amount of information to be integrated in the manner described above.

2.2 3D-scene understanding and state-estimation technology

The 3D-scene understanding and state-estimation technology is used for analyzing an entire scene (including complex shapes) and understanding its state. The aim with this technology is to automatically identify object areas such as buildings and roads from 3D data obtained from LiDAR and cameras and estimating states such as position and orientation.

The construction of an advanced geospatial information database faces two technical challenges: (i)

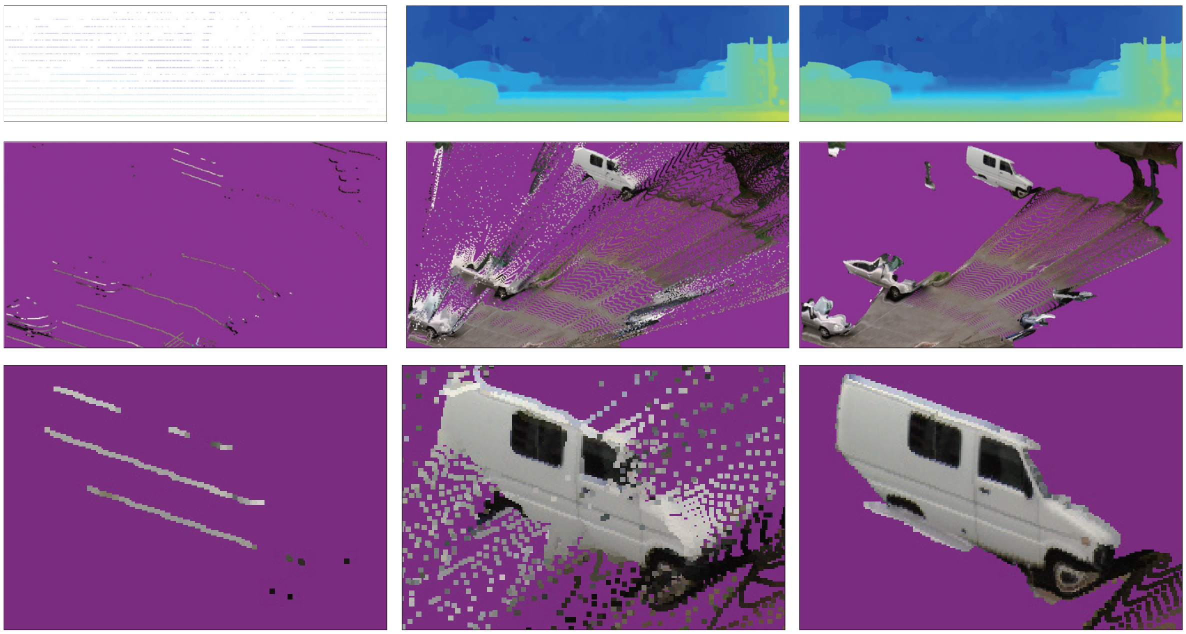


Fig. 3. 3D-data high-resolution enhancement technology using one image and one frame of LiDAR measurement.

being able to identify various natural and artificial objects near roads and (ii) keeping high classification accuracy while using data splitting and sampling to efficiently process large-scale (large-area and high-density) 3D data. To address the first challenge, we have been researching deep-learning-based point-cloud classification methods for making it easier to identify various objects. To address the second challenge, the problem that the identification accuracy decreases due to data splitting and sampling in deep-learning based methods needs to be solved.

We are working on solving this tradeoff between efficiency and classification accuracy. We experimentally identified that random sampling, often used to improve processing efficiency, deteriorates classification accuracy. Therefore, we are developing a deep-learning-based object classification model with sampling method that takes into account local shape around each 3D point of point clouds.

Specifically, high classification accuracy is achieved by executing the deep-learning-based classification process with a sampling method that preferentially leaves the points that have high distinguishability with respect to other points and do not change due to rotation or translation of the object instead of random sampling [3]. This research is just beginning; accordingly, we would like to improve the performance of object identification through technical improvement and application in real environments.

3. 4D-point-cloud coding technology

In the real world, each person has a different purpose, and we use objects that have substance and take actions that suit our purposes. In this world, things change with the passage of time. There are various scales for purposes, objects, and units concerning the people involved, and the scales differ according to the action. We expect that the ability to acquire and reuse the states of things in the real world, which have different spatial and temporal scales, can provide a diversity of value to people. Aiming to use point clouds for various purposes, we are researching and developing 4D-point-cloud coding technology.

The method called LASzip has been applied for compressing a 3D point cloud. And the ISO/IEC*1 international standardization is currently promoting

*1 ISO/IEC: The International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC)

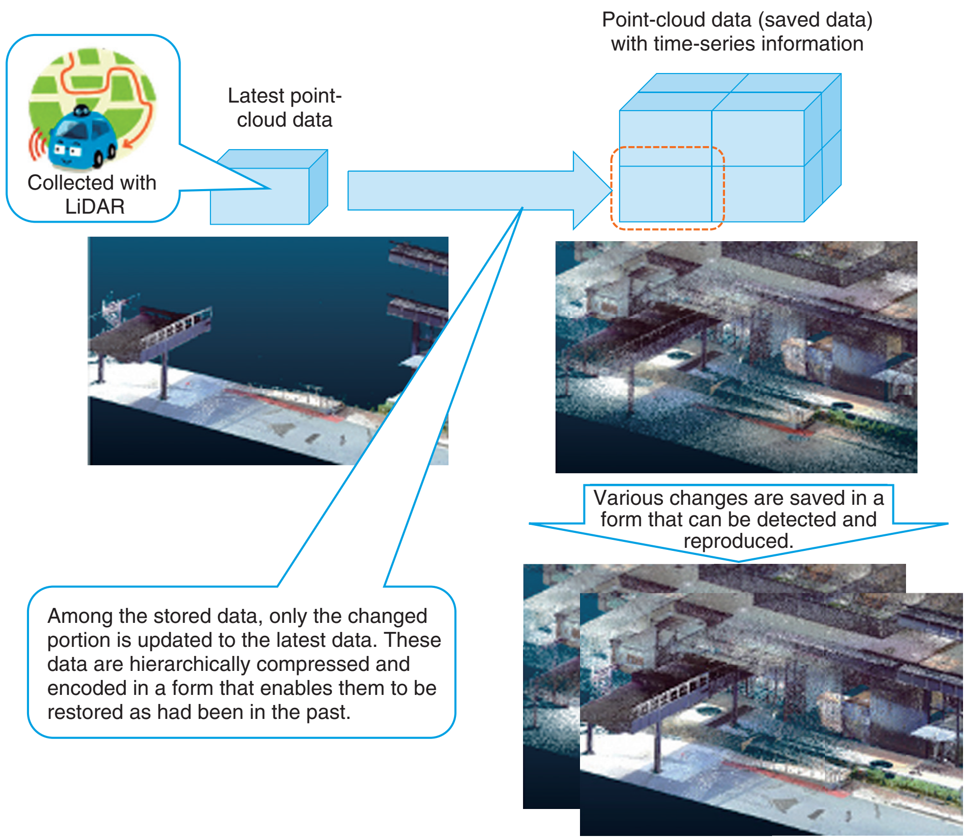


Fig. 4. Overview of 4D-point-cloud coding technology.

international standardization of a point-cloud coding method under the name MPEG[*2] geometry-based point cloud compression (G-PCC). Neither method has a mechanism for preserving temporal changes, so they are insufficient for our purpose. We are researching and developing a representation and compression-coding method for point-cloud data that is based on our knowledge of video coding that expresses changes over time as differences in the latest point-cloud data.

The 4D-point-cloud coding technology is illustrated in **Fig. 4**. Since some of the spatial information can be obtained when acquiring point-cloud data, the entire space to be expressed is divided into a grid. When the grid is made, it has a structure of multiple sizes (spatial layers) in the form of recursive inclusion, just like a matryoshka doll. The grid of each minimum unit can possess point-cloud data, and the grid of the point-cloud data is expressed by a grid of the intermediate hierarchy. As a result, point-cloud data in the entire space can be compactly represented by a hierarchical grid. Also, if it is desirable to replace some data with the latest point cloud data, point-cloud data partially included in the grid is replaced. While the point-cloud data are updated and the entire space is re-encoded, some of the past data are compressed and encoded as a difference in point-cloud data. As a result, the latest data can be decoded and expressed at any time, and the functionality that partially traces back to the past can be achieved. The functionality of compression coding as a difference in point-cloud data can be applied to quickly detect an object that was not there in the past. It is assumed that MPEG G-PCC will be used for compression encoding of point-cloud coordinate data.

We believe that the following use cases can be applied by compressing, encoding, and storing a point cloud by using this method. By acquiring point clouds from vehicles traveling on the same roads in a city every day, it becomes possible to obtain information

*2 MPEG: The Moving Picture Experts Group (MPEG) is a working group of ISO/IEC that develops standards for coded representation of digital audio and video and related data.

in real time about things that do not usually exist. Moreover, when a change is found, it becomes possible to reproduce the state before the change occurred retroactively in the past and simulate a secular change.

To apply these use cases, 4D-point-cloud coding technology, which can be used to store the point cloud (including temporal changes), will be indispensable. We will continue to promote research and development of 4D-point-cloud coding to make the real world more convenient.

4. Future developments

For constructing an advanced geospatial information database with high accuracy and abundant semantic information, the following technologies are being developed: (i) real-world digitalization for detecting features with high accuracy from images and sparse and low-density 3D point clouds and (ii) 4D-point-cloud coding that efficiently stores and uses 3D data (including temporal changes).

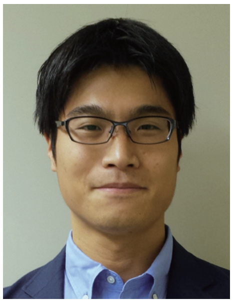
We will continue to study the means of establishing each of these technologies and evaluate their performances with actual data through demonstration experiments, and in doing so, we will contribute to establishing a 4D digital platform.

References

[1] D. Ferstl, C. Reinbacher, R. Ranftl, M. Ruether, and H. Bischof, "Image Guided Depth Upsampling Using Anisotropic Total Generalized Variation," ICCV 2013, pp. 993–1000, Sydney, Australia, Dec. 2013.

[2] Y. Yao, M. Roxas, R. Ishikawa, S. Ando, J. Shimamura, and T. Oishi, "Discontinuous and Smooth Depth Completion with Binary Anisotropic Diffusion Tensor," IEEE Robotics and Automation Letters, Vol. 5, No. 4, pp. 5128–5135, Oct. 2020.

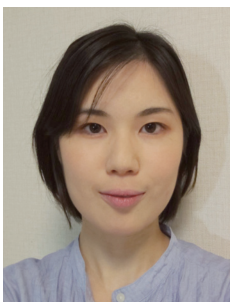
[3] K. Kurata, Y. Yao, S. Ando, and J. Shimamura, "Point Cloud Classification by Using Sampling Method for Complex Shapes," IPSJ SIG Technical Reports Computer Vision and Image Media, 2020-CVIM-220, pp. 1–6, Jan. 2020.



Yasuhiro Yao

Research Engineer, Universe Data Handling Laboratory, NTT Media Intelligence Laboratories.

He received a B.S. in pharmaceutical sciences from the University of Tokyo in 2007 and an M.E. in quantum engineering and system sciences from the University of Tokyo in 2010. He joined NTT in 2010, where he has been engaged in research and practical application development in computer vision and pattern recognition.



Kana Kurata

Staff, Universe Data Handling Laboratory, NTT Media Intelligence Laboratories.

She received a B.S. in earth and planetary sciences from Nagoya University in 2016 and an M.E. in environmental studies from Nagoya University in 2018. She joined NTT in 2018, where she has been engaged in research in computer vision and pattern recognition.

Naoki Ito

Senior Research Engineer, Universe Data Handling Laboratory, NTT Media Intelligence Laboratories.

He received an M.E. from Toyohashi University of Technology, Aichi, in 2001. He joined NTT the same year and engaged in research on character recognition. He moved to NTT EAST in 2004 and engaged in the development of security systems. He then moved to NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) in 2008. He has been working on the development of real-world digitalization technologies.

Shingo Ando

Senior Research Engineer, Supervisor, Visual Media Project, NTT Media Intelligence Laboratories.

He received a B.E. in electrical engineering from Keio University, Kanagawa, in 1998 and a Ph.D. in engineering from Keio University, Kanagawa, in 2003. He joined NTT in 2003, where he has been engaged in research and practical application development in the image processing, pattern recognition, and digital watermarks. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Institute of Image Information and Television Engineers.

Jun Shimamura

Senior Research Engineer, Supervisor, Scene Analysis Technology Group Leader, NTT Media Intelligence Laboratories.

He received a B.E. in engineering science from Osaka University in 1998 and an M.E. and Ph.D. from Nara Institute of Science and Technology in 2000 and 2006. He joined NTT Cyber Space Laboratories (now NTT Media Intelligence Laboratories) in 2000. His research interests include computer vision and mixed reality. He is a member of IEICE and the Institute of Image Information and Television Engineers.

Mayuko Watanabe

Research Engineer, Signal Modeling Technology Group, NTT Media Intelligence Laboratories.

She received a B.S. and M.S. in mathematics from Tohoku University, Miyagi, in 2007 and 2009. She joined NTT in 2009, where she has been engaged in research and practical application development in video coding.

Ryuichi Tanida

Research Engineer, Visual Media Coding Group, Visual Media Project, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. in electrical engineering from Kyoto University in 2001 and 2003. He joined NTT to study video coding to develop a real-time H.264 software encoder application. He moved to NTT Advanced Technology Corporation in 2009 to manage various software development projects related to video compression, distribution, and TV phones. He has been at NTT Media Intelligence Laboratories since 2012 as a research engineer in the visual media coding group. His current research focuses on the development of HEVC (High Efficiency Video Coding) software encoders. He is a member of IEICE.

Hideaki Kimata

Senior Research Engineer, Supervisor, Signal Modeling Technology Group Leader, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. in applied physics, and a Ph.D. in electrical engineering from Nagoya University in 1993, 1995, and 2006. He joined NTT in 1995 and has been involved in the research and development of video coding, realistic communication, computer vision, and video recognition based on machine learning (deep learning). He is a chair of the IEICE Technical Committee on Image Engineering.

History of 5G Initiatives

Yoshihisa Kishiyama and Satoshi Suyama

Abstract

In March 2020, NTT DOCOMO began providing commercial fifth-generation mobile communications systems (5G) services, which we have been studying since about 2010. Since then, we have steadily released products supporting 5G, including six smartphone models and one data communications device (a Wi-Fi router), and have introduced seven services utilizing the high speed and high capacity of 5G, such as "Shintaikan Live CONNECT," which enables multi-angle (multi-viewpoint) and virtual-reality live viewing.

This article describes the history of NTT DOCOMO's development work and co-creation with partner companies and organizations in actualizing 5G, from when we started studying 5G through to provision of these services.

Keywords: 5G concepts, 5G trials, 5G use cases

1. Introduction

In 2010, when NTT DOCOMO began its Long Term Evolution (LTE) services, we started investigating fifth-generation mobile communications systems (5G) with the goal of implementing them within ten years. We studied the basic concepts and radio access technologies for 5G, prototyped 5G simulators, and in 2014, we published a 5G white paper [1]. That same year, we also began 5G trials in collaboration with individual hardware vendors from around the world, and evaluation of 5G frequencies and radio access technologies. We also expanded collaboration in 5G trials with various partner enterprises around the world and developed many use cases for features of 5G through joint trials with these partners. In February 2018, we started the NTT DOCOMO 5G Open Partner Program to promote co-creation of solutions with partner organizations. This article reviews these initiatives at NTT DOCOMO, from proposing basic 5G concepts to co-creation of solutions through trial activities.

2. Preparing for publication of the 5G white paper (2010 to 2014)

In about 2010, looking forward to 2020, mobile communication traffic was expected to increase rapidly, by a factor of 1000 in ten years, so to meet such requirements, we proposed a "cube" concept (improve spectral efficiency*1 × expand frequency bandwidth × increase network density) as a direction for technical advancement [2]. In around 2011, we proposed specific technical candidates for implementing this cube concept, such as a radio access technology combining the existing low-frequency bands with high-frequency bands of 6 GHz or greater [3]. Then, around 2012, requirements that are now features of 5G in addition to high speed and capacity came to light, including low latency and many-terminal connectivity, and the current definition of 5G combining extensions to LTE and a new radio access technology (New RAT)*2 was proposed [4]. By about 2013, the technical concepts described in the 5G white paper published in 2014 were largely complete, including services and applications anticipated for the 5G era [5]. In this way, the basic concepts and radio access technologies for 5G advanced, building consensus with major companies around the world, and also including proposals from other companies that shared requirements and other concepts, such as exploiting higher frequency bands [6, 7].

*1 Spectral efficiency: The number of data bits that can be transmitted per unit time and unit frequency.

*2 New RAT: A new radio interface standard that is not backward compatible with 4G LTE. Also referred to as 5G New Radio.

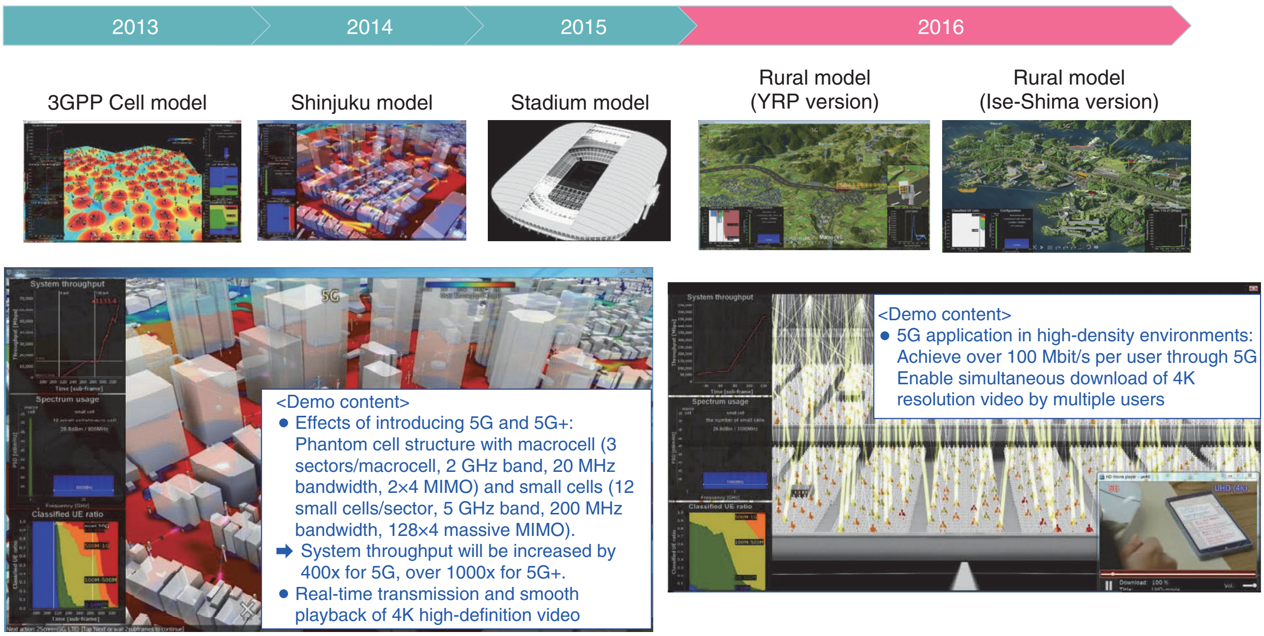


Fig. 1.   History of 5G Simulator.

Initially, NTT DOCOMO used the name "future radio access" externally, but the name "5G" was first used in October 2013 at CEATEC. At this first exhibit of 5G, a demonstration was given using a simulator to visualize the basic 5G concepts and radio access technologies, which received the Minister for Internal Affairs and Communications CEATEC Award 2013 [8]. As shown in **Fig. 1**, development of the 5G simulator continued after that, with updates such as the stadium model. Around the time when the "5G" was first used, the focus for the 5G radio access technology was on using high-frequency bands efficiently, with massive multiple-input multiple-output (Massive MIMO)*3, in 5G research projects such as Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS)*4 in Europe. NTT DOCOMO participated in METIS, leading proposals for technologies such as Phantom cells*5 (C/U (control and user plane) separation) [9] and non-orthogonal multiple access (NOMA)*6 [10]. These radio technologies and simulator prototyping were also published in the "DOCOMO 5G White Paper," as the latest work at the time [1].

3. 5G trials and development of use cases (2014 to 2020)

3.1 NTT DOCOMO initiatives with partner enterprises

To verify 5G frequencies and key radio access technologies, NTT DOCOMO conducted 5G trials collaborating individually with major global vendors. In May 2014, we announced joint testing with six companies [11], and in July 2015, we expanded this to 13 companies [12]. Through joint testing with each of these companies, we verified 5G radio access technologies such as Massive MIMO in frequency bands

*3 Massive MIMO: Large-scale MIMO using a very large number of antenna elements. Massive MIMO is expected to be useful for 5G because antenna elements can be miniaturized when using high-frequency bands, and many elements can be placed within the same area.

*4 METIS: A European Union's research project laying the foundation of 5G wireless technology. Ran from November 2012 to April 2015. Participants included communication vendors, mobile carriers, and universities. A successor project, METIS-II ran from July 2015 to June 2017.

*5 Phantom cell: The name of the advanced small-cell system proposed by NTT DOCOMO, with the basic concept of C/U plane separation. Reflected under Dual Connectivity in 5G standardization.

*6 NOMA: A method for improving efficiency of multiple access by permitting resources to overlap among users when allocating resources such as time, frequencies, or codes to multiple users.

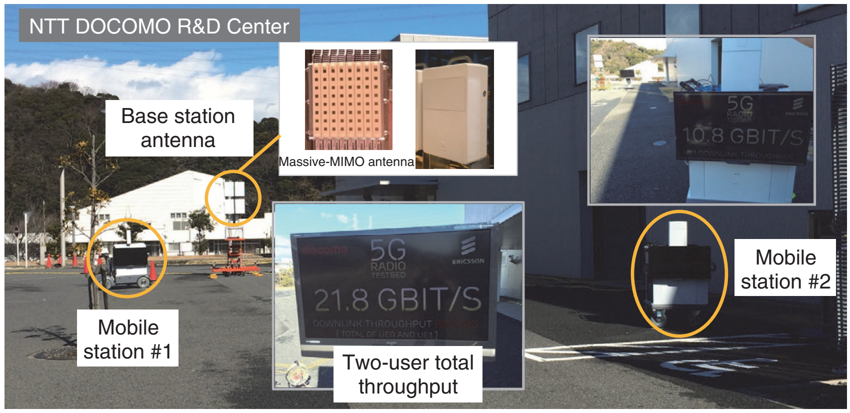


Fig. 2. 20 Gbit/s field trial.

Table 1. Major results in 5G radio technology trials (press releases).

Press release date (year/month/day)	Test results	Trial partner
2015/3/3	Successful reception of data at over 4.5 Gbit/s using the 15 GHz band	Ericsson
2015/3/3	Successful reception of data at over 2 Gbit/s using the 70 GHz band	Nokia
2015/11/26	Successful communication tests at over 2 Gbit/s in real environments with commercial equipment using the 70 GHz band	Nokia
2015/11/26	Successful data communication receiving over 2.5 Gbit/s in a high-speed mobile environment traveling at approximately 60 km/h using the 28 GHz band	Samsung
2016/2/22	Successful multi-user 5G communication trial achieving capacity over 20 Gbit/s in an outdoor environment using the 15 GHz band	Ericsson
2016/5/24	Successful real-time 5G transmission of 8K video using the 70 GHz band	Nokia
2016/11/16	Successful 5G wireless data transmission at 2.5 Gbit/s while moving at high-speeds of 150 km/h using the 28 GHz band	Samsung
2016/11/16	High-speed, high-capacity communication achieved between 23 terminals distributed in an outdoor environment, totaling 11.29 Gbit/s, using the 4.5 GHz band	Huawei
2017/11/2	Successful field trial of 5G ultra-reliable, low-latency communication (URLLC) using the 4.5 GHz band	Huawei
2017/11/2	Successful trial increasing spectral efficiency using the world’s first smartphone-sized NOMA chip set	MediaTek
2017/11/6	Successful 5G trials for implementing connected cars in the Odaiba area using the 28 GHz band	Toyota, Ericsson, Intel
2018/4/23	Successful 5G wireless communication trials in an ultra-high-speed environment of 300 km/h using the 28 GHz band	NEC
2018/5/23	Successful 5G wireless communication with multiple mobile terminals, reducing interference using cooperative beamforming among 4.5 GHz base stations	NEC
2018/7/25	Successful 5G communication for connected cars using “vehicle glass mounted antenna” and using the 28 GHz band	AGC, Ericsson
2018/11/22	Successful field trial of 27 Gbit/s communication using the 28 GHz band	Mitsubishi Electric
2019/5/29	Successful communication using a glass antenna for 5G terminals using the 28 GHz band	AGC, Ericsson

up to the 70 GHz band, and in February 2016, we successfully conducted the world’s first multi-user 5G communication tests, achieving total throughput exceeding 20 Gbit/s in an outdoor environment (**Fig. 2**). Results of other major 5G radio technology trials published in press releases are shown in **Table 1**. By March 2018, we had given 170 conference presentations based on these 5G trials [13].

Starting around 2016, we have developed various use cases utilizing features of 5G in collaboration with partner enterprises in various industries [14]. To date, we have conducted 430 service trials (as of June 2020), such as building 5G trial sites as 5G test environments in the Tokyo Rinkai Fukutoshin district (Odaiba/Aomi areas) and the Tokyo Skytree Town neighborhood.

3.2 MIC 5G field trial initiatives (2017 to 2019)

Over three years starting in FY2017, the Ministry of Internal Affairs and Communications (MIC) has conducted the 5G Field Trials, toward creation of new markets and new services and applications, with participants from various application fields [15]. The content of 5G field trials conducted by NTT DOCOMO throughout Japan with collaboration partners over the three years is shown in **Table 2**. Trial group GI conducted field trials in low-speed mobile environments, while trial group GII conducted field trials in high-speed mobile environments (60 km/h and greater). In 5G field trials in FY2017, GI conducted trials of 10 Gbit/s ultrahigh-speed communication in densely populated areas and service and application trials using 4.5-GHz and 28-GHz bands in the application fields of entertainment, smart city, and medicine [16]. NTT DOCOMO also participated in 2-Gbit/s high-speed communication trials while moving at high speed (GII), which were conducted mainly by NTT Communications, with entertainment field trials communicating using the 28-GHz band while traveling at the high speed of 90 km/h [16].

In 5G field trials in FY2018, we conducted trials at ultrahigh-speeds (GI) averaging 4 to 8 Gbit/s in an outdoor environment to verify both maximum and average 5G performance, and in office/workplace environments in addition to the three application fields tested in FY2017 [17]. We also conducted trials of high-speed communication averaging 1 Gbit/s for fast-moving objects traveling at 60 to 120 km/h (GII), and in the field of transportation in addition to entertainment [17].

In January 2019, MIC held the 5G Utilization Ideas Contest, with the objective of uncovering unique ideas that can resolve various issues that arise in outlying areas [18]. 5G field trials in FY2019 emphasized user models that use 5G to solve regional issues or contribute to regional revitalization, encompassing results from the idea contest described above, and from earlier technology trials. This involved trials with various new collaboration partners in various application fields [19].

4. Creating solutions with partners

To expand initiatives for creating new user scenarios with a wide range of partners, NTT DOCOMO started the DOCOMO 5G Open Partner Program in February 2018 (**Fig. 3**). This program provided information regarding 5G technologies and specifications to partner enterprises and organizations, and 5G Partner Workshops as a place for them to exchange ideas with each other. As of the end of June 2020, there were 3440 partners participating.

NTT DOCOMO also operates the DOCOMO 5G Open Lab as a permanent 5G technology test bed, providing partners with test equipment free of charge, such as 5G test base stations and video transmitters connected to mobile test stations. By participating in this program, partner enterprises and organizations can build and test services using 5G early, before commercial services begin, improving the quality of their own services and creating new services using the features of 5G, which are high speed, high capacity, low latency, and multi-terminal connectivity. As of the end of June 2020, DOCOMO 5G Open Lab had 11 locations in Japan and internationally. DOCOMO 5G Open Lab is also providing the DOCOMO Open Innovation Cloud trial environment connected directly to cloud computing infrastructure and is conducting technology verification.

NTT DOCOMO also began the 5G pre-commercial service in September 2019. It uses the same network equipment and frequency bands as 5G commercial services, and represents the start of services substantially the same as connecting to a 5G commercial service, enabling business creation to begin in earnest, and allowing users to experience an environment equivalent to a 5G commercial service. The 5G pre-commercial service is for the over-3000 partners participating in the DOCOMO 5G Open Partner Program, and by lending out terminals that support the 5G pre-commercial service together with provision of 5G Open Lab, over 200 of field trials have been conducted throughout Japan, toward creating industries using 5G and solving social issues. From these, seven services and 22 solutions produced through co-creation with all of our partners, such as Shintaikan Live CONNECT, were announced at the start of 5G services, focused on solving social issues such as advancing industry and reform work practices. In the future, we intend to provide new kinds of value related to work practices, utilizing the 5G features of high speed, high capacity, and low latency, and focusing on areas such as remote work support and

Table 2. Location and details of 5G field trials conducted, 2017 to 2019.

Trial group	Application field	Use case	Location	Fiscal year 2017	Fiscal year 2018	Fiscal year 2019
GI	Entertainment	(1) Live music VR experience using 5G	Sumida Ward, Tokyo	✓		
		(2) MR communication using 5G	Sumida Ward, Tokyo	✓		
		(3) 8K video multichannel MMT transmission using 5G	Sumida Ward, Tokyo	✓		
		(4) 4K low-power digital signage using 5G	Sumida Ward, Tokyo	✓		
		(5) High-definition video transmission in a shopping-mall environment	Sumida Ward, Tokyo	✓		
		(6) Live viewing of a sports event using 5G	Yokohama City, Kanagawa Prefecture	✓		
		(7) Ultra-high-definition video transmission to steam locomotive "SL Taiju" using 5G	Nikko City, Tochigi Prefecture		✓	
		(8) Tourism promotion using 5G and 8K video	Kyoto City, Kyoto Prefecture		✓	
		(9) Live video distribution for wheelchair basketball using 5G	Kamakura City, Kanagawa Prefecture		✓	
		(10) Remote museum visits using 5G and VR technology	Katsuyama City, Fukui Prefecture		✓	
		(11) Live viewing of tourism events using 5G	Aizuwakamatsu City, Fukushima Prefecture		✓	
		(12) Remote live support for sporting events using 5G	Kawasaki City, Kanagawa Prefecture		✓	
		(13) Promoting tourism in Okinawa Prefecture with 5G and Body Sharing technology	Nago City Region, Okinawa Prefecture			✓
		(14) Real-time cloud editing/relay solution using 5G	Sendai City, Miyagi Prefecture			✓
	Smart city	(1) Facilities monitoring using 5G	Koto Ward, Tokyo	✓		
		(2) Wide area surveillance using elevated cameras and 5G	Sumida Ward, Tokyo	✓		
		(3) Car security using 5G	Kamiyama Town, Tokushima Prefecture		✓	
		(4) Rural monitoring services using 5G	Aizuwakamatsu City, Fukushima Prefecture		✓	
	Medicine	(1) Remote medical examination and treatment using 5G	Wakayama City, Wakayama Prefecture Hidakagawa Town, Wakayama Prefecture	✓	✓	✓
		(2) Advanced urgent conveyance solutions using 5G	Maebashi City, Gunma Prefecture		✓	✓
		(3) Monitoring and behavior understanding using highly accurate face recognition authentication, sensors, and 5G	Hiroshima City, Hiroshima Prefecture			✓
	Workplace	(1) Mobile satellite offices using 5G	Kamiyama Town, Tokushima Prefecture		✓	
		(2) Ensuring work safety for highly skilled workers using 5G	Imabari City, Ehime Prefecture			✓
	Smart life	(1) Transmission of traditional arts using 5G (remote education)	Nakatsugawa City, Gifu Prefecture			✓
		(2) Lifestyle support through visualization of sound using 5G	Nakatsugawa City, Gifu Prefecture			✓
GII	Entertainment	(1) Transmission of high-definition video to high-speed moving objects using 5G	Tochigi, Kanuma Cities, Tochigi Prefecture; Kasukabe City, Saitama Prefecture	✓	✓	
		(2) Guidance at a golf course using 5G	Nagano City, Nagano Prefecture			✓
	Transportation	(1) Monitoring for safe operation of high-speed trains using 5G	Takatsuki City, Osaka Prefecture		✓	
		(2) Driving assistance during heavy fog using 5G	Oita City Region, Oita Prefecture			✓
		(3) Support ensuring subway safety using 5G	Osaka City, Osaka Prefecture			✓
	Smart city	(1) Protection against snow damage using 5G (improving efficiency of snow removal)	Eiheiji Town, Fukui Prefecture			✓

MMT: MPEG Media Transport   MR: mixed reality   VR: virtual reality

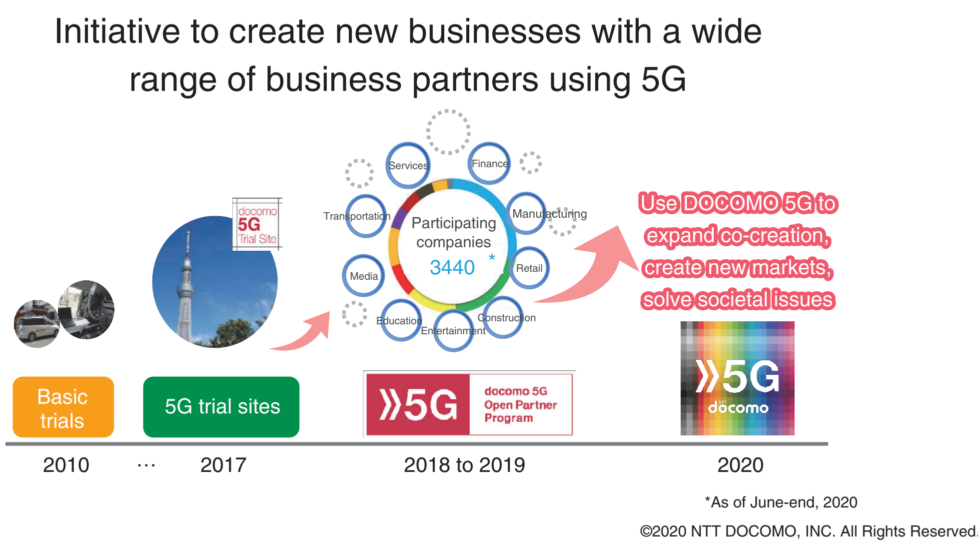


Fig. 3. NTT DOCOMO 5G Open Partner Program.

high-definition video transmission.

5. Conclusion

In this article, we described the path leading to the start of 5G services by NTT DOCOMO. In the future, NTT DOCOMO will continue to study technologies and conduct research and development, with the goal of further advancing 5G.

This article is the reproduction of the special article published in NTT DOCOMO Technical Journal (Vol. 22, No. 2, Oct. 2020).

References

[1] NTT DOCOMO, "DOCOMO 5G White Paper," Sept. 2014. https://www.nttdocomo.co.jp/english/corporate/technology/whitepaper_5g/index.html

[2] Y. Kishiyama, "LTE Enhancements and Future Radio Access," APCC 2010 Seminar on Future Wireless Technologies, Auckland, New Zealand, Nov. 2010.

[3] Y. Kishiyama, "LTE Enhancements and Future Radio Access towards Next Decade," 1st NGMN Innovation Day, Stockholm, Sweden, Sept. 2011.

[4] Y. Kishiyama, "Future Radio Access Challenges," WWRF Wireless World 2020 Workshop, Berlin, Germany, Oct. 2012.

[5] NTT DOCOMO, "Future Radio Access for 5G," ARIB 2020 and Beyond Workshop, Tokyo, Japan, Nov. 2013.

[6] Ericsson, "White Paper - 5G Radio Access," June 2013.

[7] Nokia Solutions and Networks, "White Paper – Looking ahead to 5G," Dec. 2013.

[8] NTT DOCOMO Press Release: "'Next-generation Mobile Communications System (5G)' Receives Minister for Internal Affairs and Communications Award at the CEATEC AWARDS 2013," Oct. 2013 (in Japanese).

[9] H. Ishii, Y. Kishiyama, and H. Takahashi, "A Novel Architecture for LTE-B: C-plane/U-plane Split and Phantom Cell Concept," IEEE Globecom, Anaheim, CA, USA, Dec. 2012.

[10] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal Multiple Access (NOMA) for Cellular Future Radio Access," IEEE VTC-Spring, Dresden, Germany, June 2013.

[11] NTT DOCOMO Press Release: "DOCOMO to Conduct 5G Experimental Trials with World-leading Mobile Technology Vendors," May 2014.
https://www.nttdocomo.co.jp/english/info/media_center/pr/2014/0508_00.html

[12] NTT DOCOMO Press Release: "DOCOMO to Collaborate on 5G with Five Additional Vendors—Expands 5G collaborations to 13 world-leading vendors—," July 2015.
https://www.nttdocomo.co.jp/english/info/media_center/pr/2015/0722_00.html

[13] NTT DOCOMO, "Overview of 5G Trials in Collaboration with World-leading Vendors" (in Japanese),
https://www.nttdocomo.co.jp/corporate/technology/rd/tech/5g/5g_trial/

[14] T. Nakamura, "Views of the Future Pioneered by 5G—A World Converging the Strengths of Partners—," NTT DOCOMO Technical Journal 25th Anniversary, pp. 53–59, Dec. 2018.

[15] MIC Press Release: "5G Integration Trials Begin," May 2017 (in Japanese).
https://www.soumu.go.jp/menu_news/s-news/01kiban14_02000297.html

[16] NTT DOCOMO Press Release: "Results of FY2017 Ministry of Internal Affairs and Communication '5G Comprehensive Demonstration Test'," Mar. 2018 (in Japanese).
https://www.nttdocomo.co.jp/binary/pdf/corporate/technology/rd/topics/2017/topics_180326_03.pdf

[17] NTT DOCOMO Press Release: "FY2018 MIC '5G Integration Testing' Results: Verification of Various 5G Use Cases Contributing to Regional Revitalization and Japan's First Application of 5G for Emergency Medical Treatment," Mar. 2019 (in Japanese).
https://www.nttdocomo.co.jp/binary/pdf/info/news_release/topics_190319_01.pdf

[18] MIC Press Release materials: "Holding the '5G Utilization Idea Contest'," Dec. 2018 (in Japanese).

https://www.soumu.go.jp/menu_news/s-news/01kiban14_02000362.html
[19] NTT DOCOMO Press Release: "FY2019 MIC '5G Integration Testing' Results: Verification of 12 5G Use Cases Contributing to Regional Revitalization," Mar. 2020 (in Japanese).
https://www.nttdocomo.co.jp/binary/pdf/info/news_release/topics_200316_00.pdf

Trademark notes

Yoshihisa Kishiyama
Manager, 6G Laboratories, NTT DOCOMO, INC.

He received a B.E., M.E., and Ph.D. from Hokkaido University in 1998, 2000, and 2010. Since joining NTT DOCOMO in 2000, he has been involved in the research and development of 4G and 5G radio access technologies, including concept development, standardization, and trials. He is currently working on "5G evolution & 6G" and is the main writer of "DOCOMO 6G White Paper." He has more than 400 granted patents in the area of mobile communications. In 2012, he received the ITU Association of Japan Award for his global contributions to LTE.

Satoshi Suyama
Manager, 6G Laboratories, NTT DOCOMO, INC.

He received a B.E., M.E., and Dr. Eng. from Tokyo Institute of Technology in 1999, 2001, and 2010. From 2001 to 2013, he was an Assistant Professor at Tokyo Institute of Technology. Since joining NTT DOCOMO in 2013, he has been involved in research and development of 5G/6G radio access technologies. He received the Best Paper Prize from the European Wireless Technology Conference (EuWiT) in 2009, the Best Paper Award from the International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) in 2016, and the Paper Award from the Institute of Electronics, Information and Communication Engineers (IEICE) in 2012.

NTT DOCOMO Contributions to 5G Standardization

Satoshi Nagata, Atsushi Minokuchi, Anil Umesh, Yuta Oguma, and Shinji Takeda

Abstract

At the 3rd Generation Partnership Project (3GPP), the initial version of the fifth-generation mobile communications systems (5G) standard was completed in June 2018, and at the Open Radio Access Network (O-RAN) Alliance, specifications for achieving a 5G open radio access network were released in March 2019. NTT DOCOMO is an active contributor to standardization efforts at 3GPP and O-RAN and the commercial 5G service that it launched in March 2020 conforms to 3GPP and O-RAN specifications. This article provides an overview of NTT DOCOMO's contributions to the formulation of these specifications.

Keywords: 3GPP, 5G standard, O-RAN

1. Introduction

At the 3rd Generation Partnership Project (3GPP) Technical Specification Group (TSG) #84 plenary meeting held in June 2018, 3GPP Release 15 specifications were declared frozen marking the completion of the initial fifth-generation mobile communications systems (5G) standard. The 5G standard is a new radio access specification called New Radio (NR). It features a new core network*1 called the 5G core network (5GC) and includes functions for linking with the existing Long Term Evolution (LTE)/Evolved Packet Core (EPC)*2 system.

At 3GPP, 5G standardization envisioned the following three usage scenarios based on the ITU-RM.2083 vision recommendation issued by the International Telecommunication Union - Radio Communication Sector (ITU-R) in September 2015: (i) enhanced mobile broadband (eMBB), (ii) massive machine type communications (mMTC), and (iii) ultra-reliable and low-latency communications (URLLC). Studies for Release 14 commenced in March 2016 with a study item (SI) established on requirements and elemental technologies. This was followed by studies for Release 15 from March 2017 with a work item (WI) established on detailed specifications.

NTT DOCOMO proactively contributed to the formulation of 5G specifications at 3GPP over the two-year, three-month period from the beginning of Release 14 to the completion of Release 15. This contribution included technical proposals, compilation of specifications, and Chairman and Vice Chairman roles at meetings. During this time, the number of contributions to 5G specifications submitted by NTT DOCOMO came to approximately 3700 items, which ranked ninth among worldwide companies and first as a mobile operator [1]. NTT DOCOMO was also evaluated as first among mobile operators in terms of the number of 5G candidate essential-patent applications [2].

In parallel with these 3GPP standardization activities, NTT DOCOMO established the Open Radio Access Network (O-RAN) Alliance with AT&T, China Mobile, Deutsche Telekom, and Orange in February 2018. Aiming to construct a highly extendible,

*1 Core network: A network consisting of packet transfer equipment, subscriber information management equipment, etc. Mobile terminals communicate with the core network via the radio access network.

*2 EPC: A core network that can accommodate diverse radio access systems including LTE.

open, and intelligent next-generation radio access network[*3] as typified by 5G, the O-RAN Alliance has been working to (i) promote open interfaces with high interoperability, (ii) apply virtualization and white box[*4] schemes to radio network equipment, and (iii) leverage artificial intelligence (AI) and big data.

As a result of these efforts, 3GPP Release 15 specifications were completed in June 2018 and O-RAN fronthaul[*5] specifications were completed in March 2019 thereby opening the door to development for 5G commercial services.

This article describes the activities involved in the formulation of 5G specifications and NTT DOCOMO's contributions to this effort.

2. Activities toward early formulation of 5G specifications at 3GPP and contributions

In the formulation of 5G specifications at 3GPP, studies on NR, the new radio access system, proceeded with the aim of standardizing specifications targeting use cases for achieving parts of eMBB and URLLC in Release 15 and standardizing remaining specifications including URLLC and mMTC in the next stage consisting of Release 16 and beyond. Here, it was decided to formulate specifications in a systematic manner through stepwise standardization of 5G that encompasses a wide array of requirements and use cases. For example, standardization activities would first tackle specifications for the case of combining and operating 4G base stations and the 4G core network with 5G base stations in a scheme called "non-standalone" operation and then take up all specifications in the next stage including "standalone" operation that enables operation with only 5G base stations and the 5G core network.

NTT DOCOMO made major contributions not only to promoting 5G standardization overall but also to early formulation of 5G specifications. Specifically, while guaranteeing the standardization of non-standalone operation and standalone operation in the same release (Release 15) by June 2018 as originally scheduled, NTT DOCOMO and 47 other companies jointly submitted to 3GPP a document stating that specifications for non-standalone operation would be completed by December 2017 thereby contributing to early formulation. This early completion of specifications helped to accelerate the worldwide deployment of commercial 5G.

3. Activities and contributions in 3GPP TSG Service and System Aspects/Core Network and Terminals

At 3GPP, TSGs carry out technical studies under the Project Coordination Group that oversees all projects. These TSGs consist of three groups: TSG Service and System Aspects (SA), TSG Core Network and Terminals (CT), and TSG Radio Access Network (RAN). The first of these, TSG SA, studies use cases and derives system requirements, decides on architecture and derives requirements for individual functions, and decides on the flow of information between functional units and the operation of individual functions. It also prescribes specifications for each of the above items. TSG CT, meanwhile, prescribes protocols between terminals and the core network and between functional units within the core network plus detailed specifications for an external application programming interface (API) for third-party users and for a universal subscriber identity module (USIM)[*6]. Among its work in prescribing 5G specifications, NTT DOCOMO has held Vice Chairman posts in TSG CT and SA Working Group 3 (SA-WG3) and has contributed greatly to establishing overall policies on prescribing specifications, drawing up schedules, and promoting technical studies.

In prescribing 5G specifications, TSG SA/CT expanded functions for accommodating NR as a non-standalone system in EPC and studied 5GC as new core-network specifications. Thinking that a suitable amount of time would be needed to expand the 5G coverage area, NTT DOCOMO emphasized the launching of 5G services in a non-standalone system within SI studies for Release 14. However, at the beginning of these studies, it was the opinion of both vendors and mobile communications operators that studies on a new core network should be unified, so no approval for our proposal could be obtained. However, thanks to repeated explanations and progress in TSG RAN technical studies (described below), approval was gradually obtained, and later, priority

*3 Radio access network: The network situated between the core network and mobile terminals consisting of base stations and other equipment for controlling the radio layer.

*4 White box: Hardware whose internal configuration and processing are open in contrast to black box hardware.

*5 Fronthaul: The circuit between radio equipment and the baseband processor of base-station equipment achieved by optical fiber, etc.

*6 USIM: An integrated circuit card used to store information for identifying a subscriber having a contract with a mobile phone company.

was placed on drafting specifications for a non-standalone system as 5G at the RAN/SA #72 plenary meeting. Despite this drawn-out chain of events, many operators around the world have since adopted this non-standalone system as the main network configuration in the initial period of 5G deployment.

In addition, specifications for a new core-network concept were established as 5GC for application to the 5G service rollout period.

NTT DOCOMO contributed to advancing specifications for both accommodating NR in EPC (non-standalone system) and for 5GC.

3.1 SA-WG1

SA-WG1, whose responsibilities are to study use cases and derive system requirements, studied only 5GC without relation to the non-standalone system. Here, NTT DOCOMO hypothesized that customer needs and desires would be varied with a long-tail*7 distribution that would require not a uniform network but rather a divided network (the network slice*8 concept), which it promoted. NTT DOCOMO also took up issues fundamental to operators and took the lead, in particular, in drawing up specifications for access control.

3.2 SA-WG2

SA-WG2 is responsible for determining architecture, deriving requirements for each functional unit, and deciding the flow of information between functional units and the operation of individual functions. This WG studied both the non-standalone system and 5GC.

Some European communications operators treated the non-standalone system as part of the flow toward higher LTE radio speeds. With this being the case, policies were established on operating this system without any changes to the core network so that many operators could deploy it and on inserting functions enabling advanced control for operators willing to make changes to the core network. Additionally, for individual technical items, policies were established on studies regarding encryption and serving gateway (S-GW)*9 selection in SA-WG3 and CT-WG described below.

In 5GC, several companies declared their candidacy for the position of Rapporteur thereby delaying the drafting of the work item description (WID), but NTT DOCOMO took on the role of Editor at the request of WG members and got the WID completed. After that, NTT DOCOMO also contributed to studies on network slicing and EPC-5GC interoperability.

3.3 SA-WG3

SA-WG3 has the responsibility of making comprehensive studies on security issues from use cases to protocol. In this WG, studies were made on both of the systems described above. NTT DOCOMO served as Vice Chairman and led technical studies. In the non-standalone system, it was decided that security applied to the secondary radio access technology (Secondary RAT)*10 radio interval would operate by reading LTE security capabilities into NR security capabilities without having to make changes to the mobility management entity (MME)*11. NTT DOCOMO led the drafting of specifications for this scheme. In 5GC studies, NTT DOCOMO served as Rapporteur of main specifications documents for 5GC overall and led improvements to encryption and authentication in the radio portion.

3.4 CT-WG1

CT-WG1 is responsible for specifying protocol in detail between terminals and the core network and of studying certain basic functions in terms of architecture. In relation to 5GC, NTT DOCOMO promoted the drafting of specifications for the Earthquake and Tsunami Warning System, Steering of Roaming, and access control.

3.5 CT-WG4

CT-WG4 is responsible for specifying protocol in detail between functional units within the core network. In the non-standalone system, NTT DOCOMO promoted the drafting of specifications for a system to select an appropriate S-GW for accommodating high traffic.

4. Activities and contributions in 3GPP TSG RAN

TSG RAN is responsible for the formulation of all specifications describing the radio access network.

*7 Long tail: Here, a state in which a distribution appears to have a long tail when plotting the number of customers desiring each kind of service in descending order. In other words, a state having not a small demand for each of a variety of services.

*8 Network slice: One format for achieving next-generation networks in the 5G era. Architecture that optimally divides the core network in units of services corresponding to use cases, business models, etc.

*9 S-GW: The area packet gateway accommodating the 3GPP access system.

*10 Secondary RAT: Radio access technologies such as NR, LTE, 3G, GSM, and Wi-Fi used by secondary nodes.

*11 MME: A logical node that accommodates a base station (eNodeB) and provides mobility management and other functions.

NTT DOCOMO led technical discussions in the WGs under TSG RAN and contributed greatly to the formulation of 5G RAN specifications by serving as TSG RAN Vice Chairman.

4.1 RAN-WG1

RAN-WG1, which carries out studies in relation to the radio interface on the physical layer, studied, in particular, radio access systems and multiple transmit/receive antenna technologies. For example, in NR, it defined orthogonal frequency division multiplexing (OFDM)*12 based on multiple subcarrier*13 intervals with the aim of supporting a variety of use cases and a wide frequency band from low-frequency bands used in existing cellular systems to high-frequency bands including the millimeter-wave frequency band.

Here, NTT DOCOMO made a number of technical proposals for band-expansion technology, antenna technology, and initial-stage access technology and made significant contributions to completing specifications by coordinating progress management and concerned individuals as Rapporteur for formulating 5G specifications, serving as RAN-WG1 Chairman, and taking leadership roles in various elemental technologies.

4.2 RAN-WG2

RAN-WG2, which has the responsibility of studying radio interface architecture and protocol, drafted a specification to enable NR communications to be carried on top of existing LTE communications through Dual Connectivity*14 between LTE and NR. This specification is based on the specification for LTE Dual Connectivity that was promoted by NTT DOCOMO as Rapporteur and standardized in Release 12. It enables high speeds and low latency by NR while making use of the existing LTE network by adopting a split bearer*15 [3] that terminates at a secondary node (SN)*16 proposed by NTT DOCOMO for 5G. NTT DOCOMO made significant contributions to completing specifications in RAN-WG2 to achieve diverse use cases and meet requirements for 5G. For example, given that functions and capabilities reported by a terminal to a base station and parameters set in a terminal by the base station can result in an enormous amount of data, NTT DOCOMO led studies on Radio Resource Control messages for transferring such data.

4.3 RAN-WG3

RAN-WG3 is responsible for studying RAN architecture and interfaces. This WG extended the X2 interface between eNB base stations and enabled eNB–gNB linkage to achieve Dual Connectivity between LTE and NR. Additionally, to enhance the flexibility of RAN rollout with improved performance and cost efficiency in mind, RAN-WG3 specified a F1 interface to enable a base station to be split into a central unit (CU), which terminates the Packet Data Convergence Protocol sublayer and above, and distributed unit, which terminates the Radio Link Control sublayer and below. It also specified an E1 interface to enable CU to be split into a CU-control plane, which terminates the control plane*17, and CU-user data plane, which terminates the user plane*18 [3].

Here, to enhance multi-vendor interoperability in these interfaces, NTT DOCOMO submitted many technical contributions by leveraging its experience in achieving multi-vendor capabilities in the commercial LTE network. It also contributed by serving as Moderator to encourage discussions in this area and as Editor of specifications.

4.4 RAN-WG4

RAN-WG4 bears the responsibility of formulating specifications for base-station and terminal radio frequency (RF) performance and radio resource management. In addition to frequency bands below 6 GHz as used in LTE/LTE-Advanced, this WG studied the use of sub-millimeter wave*19 and millimeter wave*20 frequency bands introduced for NR (defined

*12 OFDM: A multi-carrier modulation format where information signals are modulated with orthogonal subcarriers. A type of digital modulation scheme where information is split across multiple orthogonal carriers and transmitted in parallel. It can transmit data with high spectral efficiency.

*13 Subcarrier: Individual carrier for transmitting signals in multi-carrier transmission such as OFDM.

*14 Dual Connectivity: A technology that achieves wider bandwidths by connecting two base stations in a master/secondary relationship and performing transmission and reception using multiple component carriers supported by those base stations.

*15 Split bearer: In Dual Connectivity, a bearer that is transmitted and received via both master node (MN) and secondary node (SN) base stations.

*16 SN: A base station that provides a UE in Dual Connectivity with radio resources in addition to those provided by the MN. In LTE-NR Dual Connectivity, the SN is an NR base station (gNB) if the MN is an LTE base station (eNB) and an LTE base station (eNB) if the MN is an NR base station (gNB).

*17 Control plane: Control processes to transfer data, such as route control of data in use for communications.

*18 User plane: The part of the signal sent and received in communication, which contains the data sent and received by the user.

*19 Sub-millimeter wave: Radio signals of frequencies in the millimeter wave range, from approximately 10 to 30 GHz.

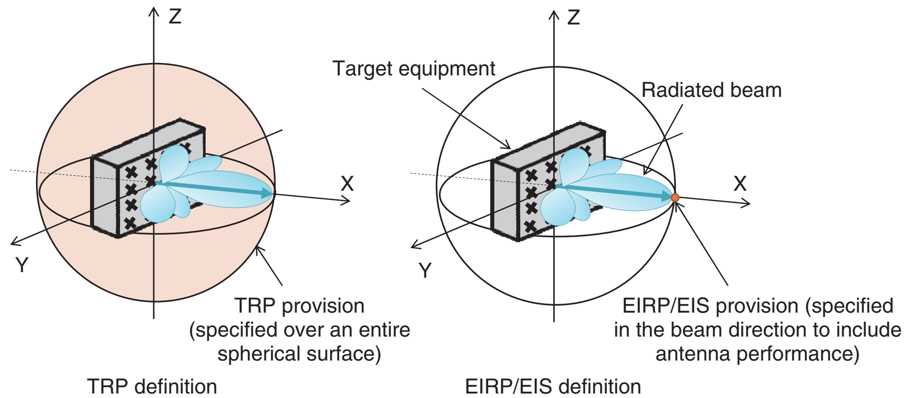


Fig. 1. Definitions used in the RF specification of NR base stations and terminals.

in Release 15 as Frequency Range 2 (FR2): 24,250–52,600 MHz frequency band). Serving a four-year term as Vice Chairman of RAN-WG4 from 2015, NTT DOCOMO actively submitted technical proposals and promoted discussions toward the formulation of frequency bands taking the 5G frequency allocation plan in Japan into account and toward the drafting of RF performance specifications for base stations, terminals, etc. The legal system surrounding 5G in Japan advanced in step with these specifications.

The use of FR2 heralds the coming of wide-band communications, but it is not without its problems from the viewpoint of configuring a RF section. Specifically, it means an increase in power loss and signal propagation loss in the RF section due to high frequencies and a contraction of area coverage brought on by a drop in power density due to wideband communications. To solve these problems, coverage must be ensured by achieving high antenna gain through an antenna array formation. However, it is difficult to implement a radio-signal transceiver and antennas at high density in a small space, so for FR2, a conventional RF configuration with connectors for measurement use cannot be applied. For this reason, an over the air (OTA)*21 provision was introduced in the FR2 RF specification to enable testing of the RF configuration without connectors.

As shown in **Fig. 1**, the OTA provision defines total radiated power (TRP) emitted from the equipment targeted for measurement as well as equivalent isotropic radiated power (EIRP)*22 and equivalent isotropic sensitivity (EIS)*23 in the beam direction that includes antenna characteristics.

Focusing on EIRP maximum transmission power in FR2 for a terminal, the adopted provision uses a cumulative distribution of each EIRP value obtained when manipulating the beam direction over an entire spherical surface centered about the terminal (**Fig. 2**). The purpose of introducing this provision was to statistically guarantee that the beam is correctly oriented in the intended direction (the direction of the base station performing communications) and within the necessary range. The value that must be minimally satisfied by at least one of the measured EIRP values is defined as the Min peak value and the value at *X*% of the cumulative distribution, that is, the value for which the (100–*X*)% area in spherical surface space must be guaranteed, is defined as spherical coverage. Thinking that spherical coverage would be essential to ensuring FR2 area performance as an operator, NTT DOCOMO expanded discussions while gaining the approval of various operators that spherical coverage would preferentially become a high-performance provision among various types of provisions with technical feasibility considered.

*20 Millimeter wave: Radio signals of frequencies in the range from 30 GHz to 300 GHz.

*21 OTA: A method for measuring radio characteristics transmitted or received from a base station or terminal, by positioning it opposite to a measurement antenna. Configurations have been defined for NR base stations and terminals that have no antenna connectors, so this method has been established for regulating such devices.

*22 EIRP: The transmission power at the reference point in radio radiation space.

*23 EIS: The received power at the radiated requirement reference point in radio reception space.

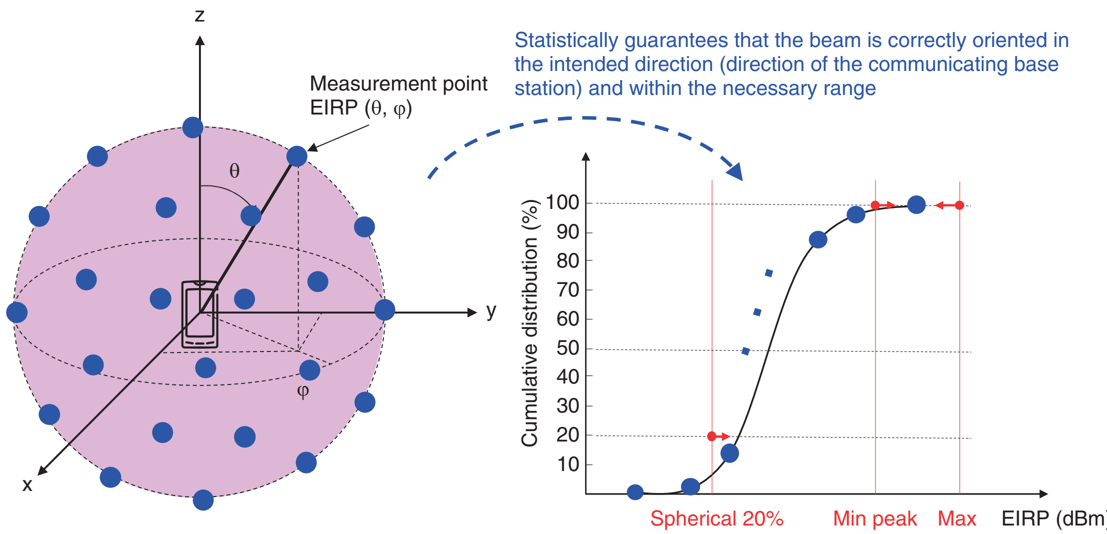


Fig. 2. OTA EIRP evaluation technique using a cumulative distribution.

5. Activities and contributions in O-RAN

NTT DOCOMO joined up with AT&T, China Mobile, Deutsche Telekom, and Orange in February 2018 to launch the O-RAN Alliance with the aim of constructing a highly extendible, open, and intelligent RAN for the 5G era. Today, many operators and vendors are participating as members (about 200 subscribed members as of June 26, 2020) and a wide range of studies are being conducted in WGs in such areas as open interfaces to enable interoperability in RAN, virtualization, use of AI and big data, white box hardware, and open source software. Many specifications have already been released. As a result of these activities, the O-RAN Alliance has been attracting attention and generating many expectations.

Among these activities, many operators and vendors have announced their intention to adopt O-RAN open fronthaul specifications released in March 2019 to facilitate multi-vendor interoperability in the fronthaul interface between the digital processing section (centrally located) and radio section (remotely located) of a base station that up to now has been vendor-specific on a global basis [4]. This declaration is expected to have a major impact on the industry and is attracting much attention as a result.

In actuality, NTT DOCOMO had already achieved an original multi-vendor RAN since the time of the LTE generation using a common fronthaul interface in cooperation with partner vendors. Making use of the know-how gained there, NTT DOCOMO served as SI Rapporteur for base-station functional split carried out for 3GPP Release 14/15 from March 2016 to December 2017. NTT DOCOMO also participated in the xRAN Forum from February 2018 prior to its integration with the O-RAN Alliance and promoted the drafting of fronthaul specifications there in cooperation with approving operators and vendors. Now, as well, NTT DOCOMO is actively promoting fronthaul standardization such as by serving as a co-chair of WG4 that manages open fronthaul specifications. In addition, NTT DOCOMO has achieved multi-vendor interoperability using O-RAN open fronthaul specifications in a commercial 5G network as the world's first, and going forward, it seeks to achieve a global ecosystem of multi-vendor RAN in 5G. Multi-vendor interoperability in RAN expands the portfolio of base-station solutions that can be used and enables quick, flexible, and cost-efficient network construction. These features will become all the more important in 5G having many and varied requirements.

At the O-RAN Alliance, NTT DOCOMO has also served as a co-chair of WG5 that studies the conversion of 3GPP interfaces such as X2 and F1 to open interfaces and contributed to the release of profile*24 specifications to ensure multi-vendor interoperability in those interfaces. From here on, NTT DOCOMO

*24 Profile: Refers to the method of using inter-equipment messages as defined by a standardization agreement to ensure interoperability between equipment according to the application.

also plans to promote other O-RAN Alliance studies in such areas as virtualization and the use of AI and big data.

6. Conclusion

This article described activities and NTT DOCOMO contributions at 3GPP and the ORAN Alliance toward 5G standardization.

Today, at 3GPP, specifications for Release 16 toward advanced 5G have been completed and technical studies for Release 17 that aim to specify more usage scenarios are progressing. Meanwhile, at the O-RAN Alliance, in addition to further studies on open interfaces, technical studies on the virtualization of the radio network and use of big data are moving forward.

At NTT DOCOMO, we view 5G as the technical foundation for wireless communications over the next 20 years. To achieve further advances, we will continue making proactive contributions to standardization activities at 3GPP and the O-RAN Alliance.

This article is the reproduction of the special article published in NTT DOCOMO Technical Journal (Vol. 22, No. 2, Oct. 2020).

References

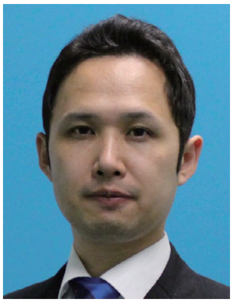
[1] Cyber Creative Institute Co. Ltd., "Application Trend of ETSI Standard Essential Patent (5G-SEP) Candidates Contributing to Realization of 5G and Proposal Trend of Contributions for Standards," 2019 (in Japanese).

[2] Cyber Creative Institute Co. Ltd., "Application Trend of 5G-SEP Candidates and 5G-SEP Declarations Contributing to Realization of 5G and Proposal Trend of Contributions for Standards," 2020 (in Japanese).

[3] T. Uchino, K. Kai, T. Toeda, and H. Takahashi, "Specifications of NR Higher Layer in 5G," NTT DOCOMO Technical Journal, Vol. 20, No. 3, pp. 62–78, Jan. 2019.
https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol20_3/vol20_3_008en.pdf

[4] NTT DOCOMO Press Release: "Mobile Industry Leaders to Develop O-RAN Fronthaul-compliant Products and Drive Multi-vendor Radio Access Networks," Feb. 2019.
https://www.nttdocomo.co.jp/english/info/media_center/pr/2019/0222_00.html

Trademark notes



Satoshi Nagata

Manager, 6G Laboratories, NTT DOCOMO, INC.

He received a B.E. and M.E. from Tokyo Institute of Technology in 2001 and 2003. In 2003, he joined NTT DOCOMO, where he engaged in research and development (R&D) of wireless access technologies for LTE, LTE-Advanced, and 5G. He is currently a manager working for 5G and 3GPP standardization. He has been contributing to 3GPP over 10 years. He served as a vice chairman of 3GPP TSG-RAN WG1 from November 2011 to August 2013, and a chairman from August 2013 to August 2017. He has been a vice chairman of 3GPP TSG-RAN since March 2017.

Atsushi Minokuchi

Assistant Manager, Core Network Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from The University of Tokyo in 1992 and 1996. Since joining NTT DOCOMO in 1996, he has been involved in R&D of 2G to 5G core network. He contributed to ITU-T SSG, SG19, and SG13 from April 2001 to March 2010, and since 2011 he has been involved in 3GPP SA and CT. From March 2015 to March 2019, he served as a vice chairman of 3GPP TSG-CT.

Anil Umesh

Manager, Radio Access Network Development Department, NTT DOCOMO, INC.

He received a B.S. in engineering from Case Western Reserve University, USA, in 2002. He joined NTT DOCOMO in 2002 and has since been involved in R&D activities related to wireless access technologies for HSPA, LTE, LTE-Advanced, and 5G. He has been the O-RAN Alliance Working Group 4 co-chair since the start of O-RAN Alliance in 2018 and has chaired open fronthaul topics.

Yuta Oguma

Research Engineer, Communication Device Development Department, NTT DOCOMO, INC.

He received a B.E. in electrical and electronic engineering in 2015 and an M.E. in informatics in 2017 from Kyoto University. Since joining NTT DOCOMO in 2017, he has been involved in R&D activities related to radio frequency wireless technologies for LTE, LTE-Advanced, and 5G, and contributed to the 3GPP TSG-RAN WG4 as a delegate from NTT DOCOMO.

Shinji Takeda

Manager, Communication Device Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. from Kyoto University in 1999 and 2001. Since he joined NTT DOCOMO in 2001, he has been involved in the research activities of wireless technology related to Wi-Fi, WiMAX and Celullar. He has also been involved in the development of commercial services of LTE and VoLTE since 2008. He has been managing the development of 5G technologies in communication devices since July 2020. Prior to the current assignment, he managed the standardization activities in the R&D strategy department.

5G Network

Yuta Sagae, Shinsuke Sawamukai, Yusuke Ohwatari, Kohei Kiyoshima, Keiichi Kanbara, and Jo Takahashi

Abstract

NTT DOCOMO launched its fifth-generation mobile communications system (5G) commercial service in March 2020 achieving early provision and stable quality by making maximum use of functions and coverage provided by LTE (Long Term Evolution). This article describes the development of radio base-station equipment and core network equipment for providing the 5G commercial service.

Keywords: 5G/NR, sub-6/mmW, non-standalone

1. Introduction

Featuring high speed and large capacity, low latency, and massive connectivity, the fifth-generation mobile communications system (5G) is generating high expectations. In addition to providing richer content as in video services, it is anticipated that 5G will help find solutions to social problems that have so far been difficult to resolve and create new industries at the same time.

After launching its 5G pre-commercial service in September 2019, NTT DOCOMO launched its 5G commercial service in March 2020. As of July 6, 2020, the number of 5G subscribers had already exceeded 170,000. The demand for 5G is expected to increase from here on, so the plan is to enhance the terminal lineup and expand 5G coverage in stages.

In this article, we describe the development of radio base-station equipment and core network equipment to provide the 5G commercial service.

2. Overview of 5G frequency bands

In 5G, there is a need for a network that can accommodate a variety of use cases and usage scenarios. From a technical perspective, this means adopting new radio technology, that is, 5G New Radio (NR), and supplementing existing frequency bands with those of even higher frequencies [1].

Three frequency bands have been allocated in Japan for 5G: the 3.7 GHz band (3.6–4.1 GHz), 4.5 GHz band (4.5–4.6 GHz), and the 28 GHz band (27.0–29.5 GHz, among which the 28.2–29.1 GHz interval is unallocated) (**Table 1**). The first two frequency bands are called the "sub-6" bands and the third one is called the millimeter wave (mmW) band.

The sub-6 bands each feature a 100 MHz bandwidth/operator within Japan's 5G frequencies. Compared with Long Term Evolution (LTE), they enable wideband usage and can achieve the same coverage as the LTE 3.5 GHz band. The mmW band, meanwhile, is significantly different from the frequencies that have come to be used by LTE, and expectations are high for the spot-like rollout of services based on transmission speeds of several Gbit/s through ultra-wideband allocation. In general, however, the higher is the frequency the more difficult is wave propagation, so there will be a need to construct a heterogeneous network that combines the mmW band with existing LTE frequencies and the sub-6 bands. In addition, antenna downsizing is simple in the case of high-frequency bands, so the rollout of networks using massive multiple-input multiple-output (Massive MIMO) as advanced MIMO technology is anticipated.

It must also be noted that existing systems are operating within Japan's 5G frequency bands, so coexistence and segregation are important. In particular, since satellite system operators are also using the 3.7 GHz band, this band will be rolled out for 5G while making inter-system interference adjustments with those satellite systems. As a consequence, the 4.5

Table 1. Features of Japan's 5G frequencies.

	3.7 GHz band	4.5 GHz band	28 GHz band
Allocated frequencies	3.6–4.1 GHz (500 MHz bandwidth)	4.5–4.6 GHz (100 MHz bandwidth)	27.0–28.2 GHz 29.1–29.5 GHz (1.6 GHz bandwidth)
Allocated bandwidth	100 MHz bandwidth/operator	100 MHz bandwidth/operator	400 MHz bandwidth/operator
Use of Massive MIMO	Used in MIMO multiplexing		Coverage extension by beamforming
Use with other systems	Satellite systems	Airplane radar altimeter	Satellite systems
Overseas trends	China, Korea, Europe, United States	Planned for future use by China	United States, Korea

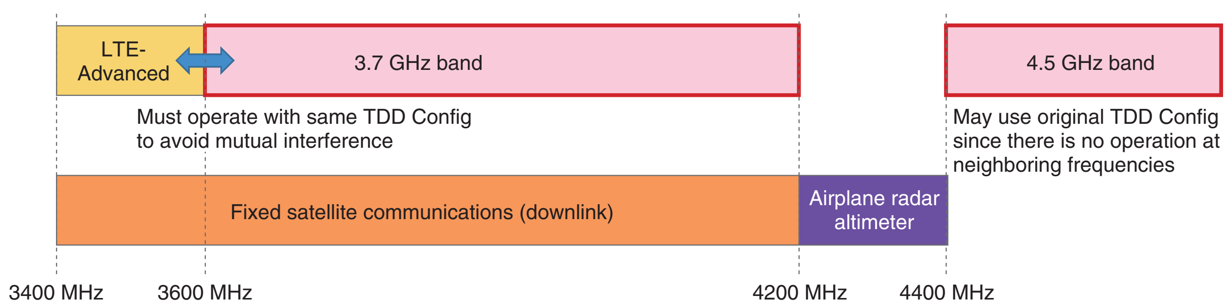


Fig. 1. Frequency allocation in 3.7 GHz band and relationship with LTE 3.5 GHz band.

GHz band has been attracting attention as a frequency band that could play an effective role in achieving an early 5G rollout (**Fig. 1**). Additionally, looking at overseas trends, the use of the 3.7 GHz band and 28 GHz band has already begun in various countries, and from the perspective of future terminal support, roaming support, etc., these bands should be effective in rolling out high-speed and large-capacity services reflecting the unique features of 5G. The plan is to execute a commercial rollout that will make maximum use of the strengths of each of these frequency bands based on the features of those frequencies, coexistence with existing operators, and global trends.

As a result of approval received from the Ministry of Internal Affairs and Communications (MIC) for base-station installation plans [2], NTT DOCOMO has been allocated the frequency intervals 3.6–3.7 GHz and 4.5–4.6 GHz for a total of 200 MHz in the sub-6 bands and 27.4–27.8 GHz for a total of 400 MHz in the mmW band. In this way, NTT DOCOMO can provide a downlink peak rate of 3.4 Gbit/s in the sub-6 bands and of 4.1 Gbit/s in the mmW band in combination with the existing LTE system (**Table 2**).

The sub-6 bands and 28 GHz band have been allocated as time division duplex (TDD)*1 bands, but from the viewpoint of spectrum efficiency, the same downlink/uplink resource allocation (TDD Config)*2 must be used between neighboring cellular operators. In Japan, the TDD Config shown in **Fig. 2** based on the TDD Config agreed upon at the 3rd Generation Partnership Project (3GPP) is being used for 5G operation taking data traffic and 5G service forecasts into account. In particular, to mitigate interference between the 3.7 GHz band and the neighboring LTE 3.5 GHz band, this TDD Config has been designed to use the same timing for downlink/uplink resources (Fig. 2). Additionally, in the 4.5 GHz band acquired solely by NTT DOCOMO, the plan is to establish an original TDD Config tailored to use cases with

*1 TDD: A bidirectional transmit/receive system. This system achieves bidirectional communication by allocating different time slots to uplink and downlink transmissions on the same frequency.

*2 TDD Config: Parameters determining how to allot uplink and downlink slots. Prescribed by 3GPP specifications.

Table 2. Technical features.

	Sub-6 bands	mmW band
Method of using Massive MIMO	MIMO multiplexing	Coverage extension by beamforming
Coverage approach	Same coverage as LTE 3.5 GHz band	High throughput by wideband use Spot-like rollout
Peak rate (3GPP standard values including LTE)	Downlink 3.4 Gbit/s, uplink 182 Mbit/s	Downlink 4.1 Gbit/s, uplink 480 Mbit/s
Number of MIMO layers	Downlink 4×4, uplink SISO	Downlink/uplink 2×2 MIMO
Modulation scheme	Downlink 256QAM, uplink 64QAM	Downlink/uplink 64QAM

QAM: quadrature amplitude modulation
SISO: single input single output

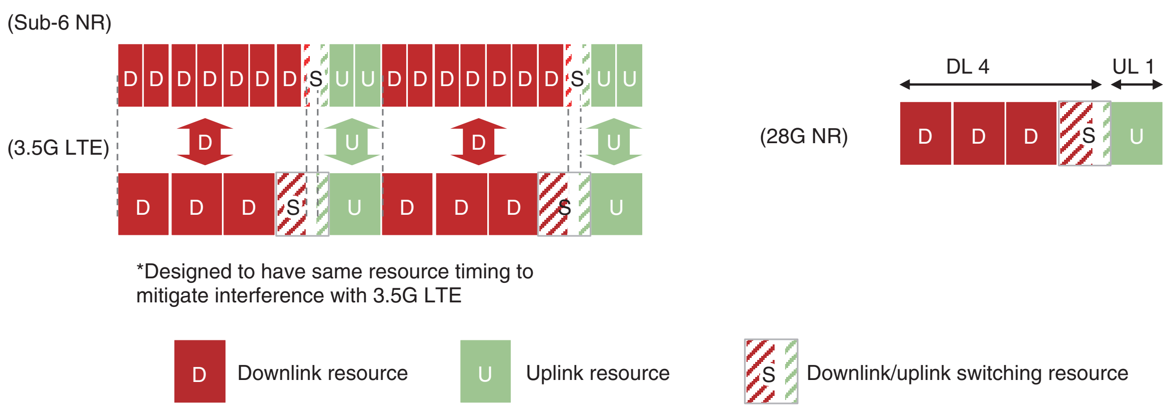


Fig. 2. TDD Config used in Japan's 5G frequencies.

spectrum efficiency in mind.

3. 5G commercial development

3.1 5G base-station equipment development

The 5G network configuration is shown in **Fig. 3**. With a view to 5G service development, NTT DOCOMO developed a central unit (CU) that consolidates the base band (BB) signal processing section supporting 5G, extended existing BB processing equipment known as high-density base-station digital processing equipment (BDE), and developed a 5G radio unit (RU) having signal transmit/receive functions. Furthermore, to have a single CU accommodate many RUs, NTT DOCOMO developed a 5G version of the fronthaul multiplexer (FHM) [4] deployed in LTE. Each of these three types of equipment is described below.

3.1.1 CU

(1) Development concept

With the aim of achieving a smooth rollout of 5G services, NTT DOCOMO developed a CU that enables area construction without having to replace existing equipment while minimizing the construction period and facility investment. This was accomplished by making maximum use of the existing high-density BDE that performs BB signal processing, replacing some of the cards of the high-density BDE, and upgrading the software to support 5G.

(2) CU basic specifications

An external view of this CU is shown in **Photo 1**. This equipment has the features described below.

As described above, this equipment enables 5G-supporting functions by replacing some of the cards of the existing high-density BDE. In addition, future software upgrades will load both software supporting conventional 3G/LTE/LTE-Advanced and software supporting 5G. This will enable the construction of a network supporting three generations of mobile communications from 3G to 5G with a single CU.

The existing LTE-Advanced system employs

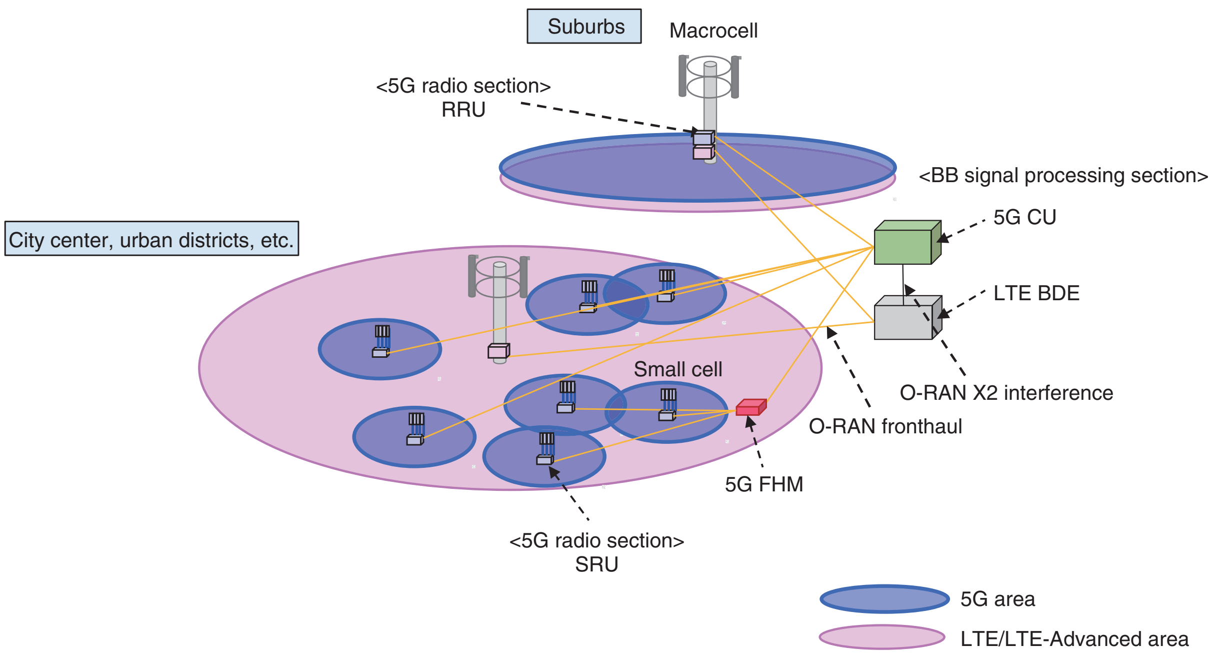


Fig. 3. 5G network configuration.

Photo 1. CU external view.

advanced centralized radio access network (C-RAN)*3 architecture [4] proposed by NTT DOCOMO. This architecture is also supported in 5G with the connection between CU and RUs made via the fronthaul. Standardization of this fronthaul was promoted at the Open Radio Access Network (O-RAN) Alliance jointly established in February 2018 by five operators including NTT DOCOMO. Since the launch of 5G services, the fronthaul in the NTT DOCOMO network was made to conform to these O-RAN fronthaul specifications that enable interoperability between different vendors, and any CU and RU that conform to these specifications can be interconnected regardless of vendor. The specifications for interconnecting base-station equipment also conform to these O-RAN

*3 Advanced C-RAN: A new C-RAN architecture proposed by NTT DOCOMO. Being controlled by the same base station, a radio access network makes a linkage between a macrocell (which covers a wide area) and a small cell (which covers a local area) by applying carrier aggregation.

(a) Antenna-integrated SRU (b) Antenna-separated SRU

Photo 2. SRU external views.

specifications, which means that a multi-vendor connection can be made between a CU supporting 5G and a high-density BDE supporting LTE-Advanced [5]. This enables NTT DOCOMO to deploy a CU regardless of the vendor of the existing high-density BDE and to quickly and flexibly roll out service areas where needed while making best use of existing assets. In addition, six or more fronthaul connections can be made per CU and the destination RU of each fronthaul connection can be selected. Since 5G supports wideband transmission beyond that of LTE-Advanced, the fronthaul transmission rate has been extended from the existing peak rate of 9.8 Gbit/s to a peak rate of 25 Gbit/s while achieving a CU/RU optical distance equivalent to that of the existing high-density BDE.

3.1.2 RU

(1) Development concept

To facilitate flexible area construction right from the launch of 5G services, NTT DOCOMO developed the low-power small radio unit (SRU) as the RU for small cells and developed, in particular, separate SRUs for each of the 3.7 GHz, 4.5 GHz, and 28 GHz frequency bands provided at the launch of the 5G pre-commercial service in September 2019. Furthermore, with an eye to early expansion of the 5G service area, NTT DOCOMO developed the regular power radio unit (RRU) as the RU for macrocells to enable the efficient creation of service areas in suburbs and elsewhere.

A key 5G function is beamforming*4 that aims to reduce interference with other cells and thereby improve the user's quality of experience. To support this function, NTT DOCOMO developed a unit that integrates the antenna and 5G radio section (antenna-integrated RU). It also developed a unit that separates the antenna and 5G radio section (antenna-separated RU) to enable an RU to be placed alongside existing 3G/LTE/LTE-Advanced radio equipment (RE) and facilitate flexible installation even for locations with limited space or other constraints.

(2) SRU basic specifications

As described above, NTT DOCOMO developed the SRU to enable flexible construction of 5G service areas. It developed, in particular, antenna-integrated SRUs to support each of the 3.7 GHz, 4.5 GHz, and 28 GHz frequency bands provided at the launch of the 5G pre-commercial service and antenna-separated SRUs to support each of the 3.7 GHz and 4.5 GHz frequency bands (**Photo 2**). These two types of SRUs have the following features.

The antenna-integrated RU is equipped with an

*4 Beamforming: A technology that gives directionality to a transmitted signal, increasing or decreasing the signal power in a particular direction. Analog beamforming works by controlling the phase in multiple antenna elements (RF devices) to create directionality, while digital beamforming controls phase in the baseband module.

Photo 3. RRU external view.

antenna panel to implement the beamforming function. In the 3.7 GHz and 4.5 GHz bands, specifications call for a maximum of 8 beams, and in the 28 GHz band, for a maximum of 64 beams. An area may be formed with the number of transmit/receive beams tailored to the TDD Config used by NTT DOCOMO. In addition, the number of transmit/receive branches is 4 for the 3.7 GHz and 4.5 GHz bands and 2 for the 28 GHz band, and MIMO transmission/reception can be performed with a maximum of 4 layers for the former bands and a maximum of 2 layers for the latter band.

The antenna-separated SRU is configured with only the radio as in conventional RE to save space and facilitate installation. With this type of SRU, the antenna may be installed at a different location. Moreover, compared to the antenna-integrated SRU operating in the same frequency band, the antenna-separated SRU reduces equipment volume to 6.5 L or less. The antenna-separated SRU does not support the beamforming function, but features four transmit/receive branches the same as the antenna-integrated SRU for the same frequency band.

(3) RRU basic specifications

The RRU was developed in conjunction with the 5G service rollout as high-power equipment compared with the SRU with a view to early expansion of the 5G service area (**Photo 3**). This type of equipment has the following features.

Compared with existing remote radio equipment for macrocells, the volume of RRU equipment tends to be larger to support 5G broadband, but in view of the latest electronic device trends, NTT DOCOMO took the lead in developing and deploying an antenna-separated RRU that could save space and reduce weight. Maximum transmission power is 36.3 W/100 MHz/branch[*5] taking the radius of a macrocell area into account. The RRU features four transmit/receive branches and achieves the same number of MIMO transmission/reception layers as the antenna-separated SRU.

NTT DOCOMO also plans to deploy an antenna-integrated RRU at a later date. The plan here is to construct 5G service areas in a flexible manner making best use of each of these models while taking installation location and other factors into account.

3.1.3 5G FHM

The 5G FHM is equipment having a multiplexing function for splitting and combining a maximum of 12 radio signals on the fronthaul. It was developed in conjunction with the 5G service rollout the same as RRU (**Photo 4**).

If no 5G FHM is being used, each RU is accommodated as one cell, but when using a 5G FHM, a maximum of 12 RUs can be accommodated as one cell in a CU. At the launch of 5G services, this meant that more RUs could be accommodated in a single CU when forming a service area in a location having low required radio capacity (**Fig. 4**). Additionally, since all RUs transmit and receive radio signals of the same cell, the 5G FHM can inhibit inter-RU interference and the occurrence of handover[*6] control

*5 Branch: In this article, an antenna and RF transceiver.

*6 Handover: A technology for switching base stations without interrupting communications when a terminal with a call in progress straddles two base stations while moving.

Photo 4. 5G-FHM external view.

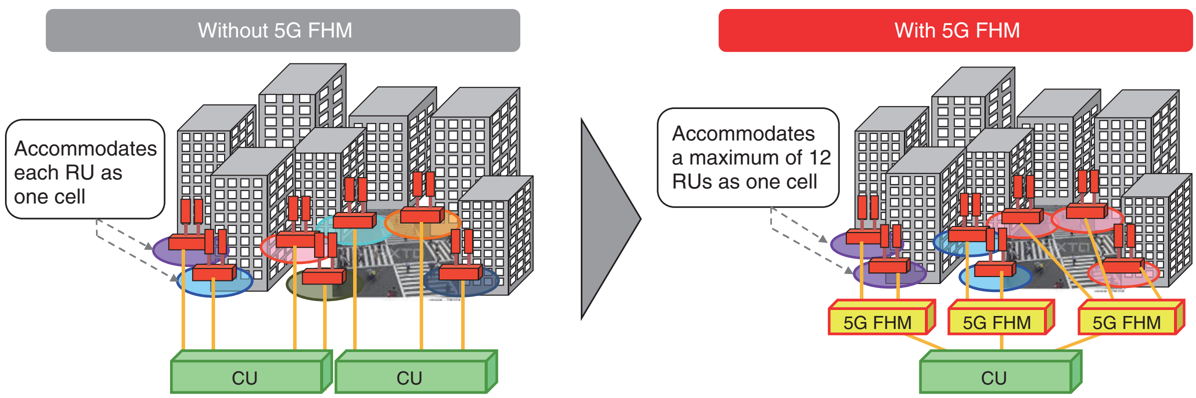


Fig. 4. Concept of 5G FHM implementation.

between RUs as in the conventional FHM. Furthermore, the 5G FHM supports all of the 5G frequency bands, that is, the 3.7 GHz, 4.5 GHz, and 28 GHz bands, which means that service areas can be constructed in a flexible manner applying each of these frequency bands as needed.

3.2 5G radio access network support

This section presents an overview of the radio access network (RAN) configuration for achieving 5G services and LTE-NR Dual Connectivity, beam management technology, and NR high-speed support as RAN technologies.

3.2.1 RAN configuration

At the launch of its 5G commercial service, NTT DOCOMO provided services through non-standalone operation, a key feature of NR. Here, "non-standalone" means an operation format that provides services while using LTE/LTE-Advanced areas as an anchor*7 instead of providing 5G areas by NR alone. As shown in **Fig. 5**, evolved NodeB (eNB)*8 in NR non-standalone operation connects to a gNB*9 base station that provides NR using an X2*10 interface. In addition, eNB and gNB connect to the Evolved Packet Core (EPC)*11 using S1 interfaces. The use of LTE as an anchor in this way through non-standalone operation made for early commercialization of 5G by enabling the use of existing network infrastructure through shared use of LTE equipment while maintaining the same level of quality with respect to connectivity that had so far been provided. In addition, the connection between eNB and gNB conforms to the O-RAN X2 specifications spearheaded by NTT DOCOMO thereby enabling interoperability between different LTE and NR operators. In short, non-standalone operation made it possible to roll out 5G areas in a speedy and flexible manner against LTE areas that had already been established.

In contrast, NR standalone operation, whose provision is scheduled for the future, will enable the provision of services through only gNB base stations. In this format, RAN will connect to the new core network, namely, the 5G core network (5GC), and gNB base stations will connect to each other using an Xn

*7 Anchor: A logical node point acting as a switching point for control signals and user bearers.
*8 eNB: A radio base station for LTE radio access.
*9 gNB: A radio base station for NR radio access.
*10 X2: A reference point between eNBs, defined by 3GPP.
*11 EPC: The core network on 3GPP mobile communication networks, mainly accommodating Evolved Universal Terrestrial Radio Access (E-UTRA).

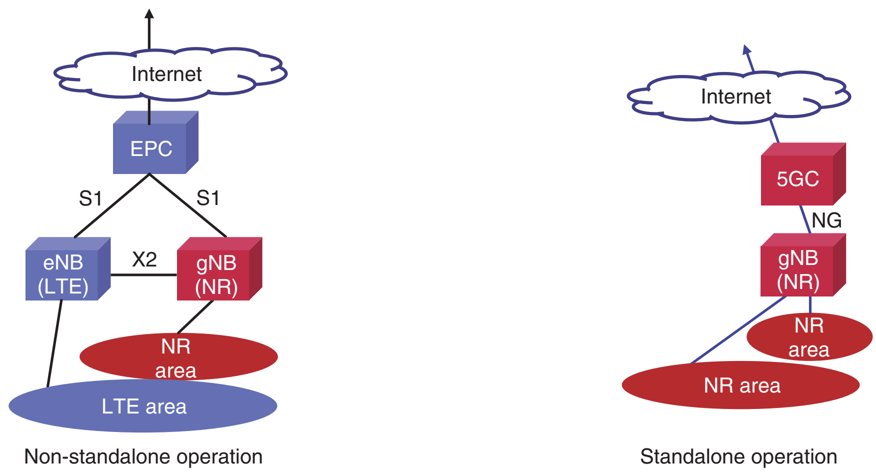


Fig. 5. Network configuration in 5G.

interface and a gNB will connect to 5GC using an NG interface.

3.2.2 LTE-NR Dual Connectivity

In the non-standalone configuration described above that combines LTE/LTE-Advanced and NR to provide services, architecture that simultaneously transmits radio signals to the user equipment (UE) by both LTE and NR is called LTE-NR Dual Connectivity. In this architecture, the simultaneous transmission of both LTE and NR types of radio resource is specified as a split bearer*12, which achieves simultaneous transmission of a maximum of five LTE carrier*13 signals and NR and enables a speedy and flexible rollout of 5G commercial services.

In LTE-NR Dual Connectivity operation in which an LTE base station acts as master node (MN), master cell group (MCG) split bearer would have to be performed on the LTE-base-station side as the bandwidth on the NR side becomes larger. It would therefore be necessary to augment the equipment on the LTE-base-station side whose capacity is limited compared with NR equipment thereby incurring equipment development and operation costs. Consequently, with the aim of minimizing the upgrading of LTE-base-station equipment and avoiding throughput limitations due to the capacity of that equipment, a secondary node (SN)-terminated split bearer and an SN-terminated MCG bearer were specified so that a branch point for user data in LTE-NR Dual Connectivity could be set up as NR equipment acting as a SN. The secondary cell group (SCG) split bearer is a method in which user data is transferred on both the MN carrier and SN carrier so that data can be transmitted to the user simultaneously from the SN and MN. This method achieves higher transmission speeds. In addition, the SN-terminated MCG bearer is a method that enables data to be transferred from the SN even if outside the NR area, which makes for more stable communications.

3.2.3 Beam management technology

New beam management technology has been adopted in the NR system. As described above, the antenna-integrated RU has a configuration enabling the forming of multiple beams (**Fig. 6**).

Specifically, in Frequency Range 1 (FR1)*14, this technology uses digital beamforming that forms beams by applying phase rotation to the BB signal, and in Frequency Range 2 (FR2)*15, it uses analog beamforming that forms beams by controlling the phase in the radio frequency (RF) signal. The technology includes processing for selecting an optimal beam when beginning an NR connection and processing for switching beams due to a change in radio quality caused by UE movement or other factors. Using an optimal transmit/receive beam according to the position of the communicating UE ensures user coverage even in a high frequency band. At the same time, it can be expected that orienting the beam

*12 Split bearer: In Dual Connectivity, a bearer that is transmitted and received via both the master and secondary base stations.

*13 Carrier: A radio signal (carrier wave) that is modulated to transmit information.

*14 FR1: A frequency band specified in 3GPP ranging from 450 to 6000 MHz.

*15 FR2: A frequency band specified in 3GPP ranging from 24,250 to 52,600 MHz.

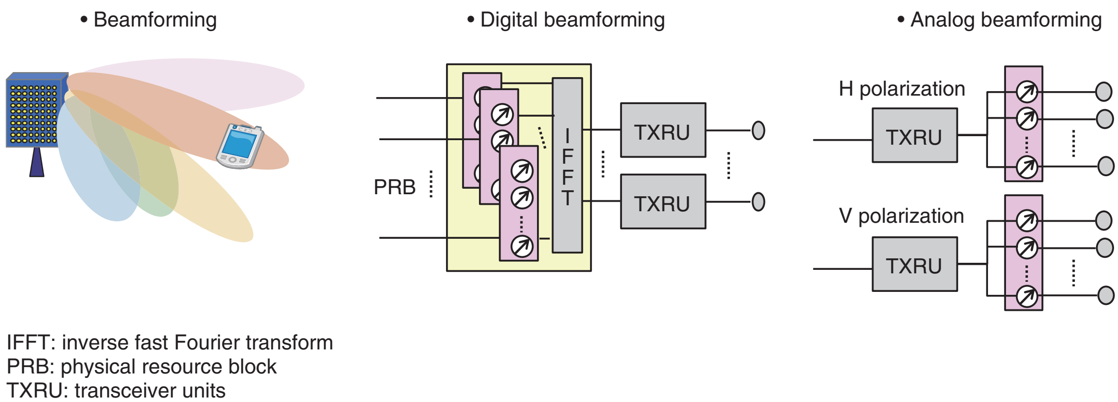


Fig. 6. Types of beamforming technologies.

toward the position of the UE will have the effect of inhibiting radio-wave radiation in unnecessary directions and suppressing inter-cell interference.

3.2.4 NR high-speed support

In 5G, even higher communications speeds can be achieved since Dual Connectivity allows NR to be simultaneously used with LTE/LTE-Advanced. Currently, in LTE/LTE-Advanced, NTT DOCOMO provides services with a downlink peak rate of 1.7 Gbit/s by combining five instances of a component carrier through carrier aggregation*16. However, in NTT DOCOMO 5G, Dual Connectivity of five LTE/LTE-Advanced component carriers and NR was achieved from the launch of 5G services enabling a downlink peak rate of 3.4 Gbit/s when using either the 3.7 GHz or 4.5 GHz NR band and a downlink peak rate of 4.1 Gbit/s when using the 28 GHz NR band.

Furthermore, in uplink communications in the 28 GHz band, in addition to the simultaneous use of two component carriers for a total bandwidth of 200 MHz through carrier aggregation, a total data rate of 480 Mbit/s has been achieved through the implementation of 2 × 2 MIMO.

The goal going forward is to increase data rates even further through the use of even wider frequency bands. In downlink communications, for example, this will be achieved through carrier aggregation between the 3.7 GHz and 4.5 GHz bands, and in uplink communications, through the carrier aggregation of four component carriers for a total bandwidth of 400 MHz in the 28 GHz band.

3.3 Development of core network equipment

Multiple migration architectures toward the provision of 5G have been proposed at 3GPP, an international standards organization. **Figure 7** shows those architectures that have been specified at 3GPP. Option 1 is the architecture under which LTE has been provided. Options 2–5 and 7 are architectures for providing 5G, and each operator may decide which option(s) to adopt. As described above, NTT DOCOMO decided to launch 5G by adopting Option 3 architecture that connects to NR and provides 5G by extending EPC that had been in commercial operation for LTE. This approach made it possible to achieve early provision of the 5G commercial service while ensuring a stable level of quality provided by LTE/LTE-Advanced that was already in place. Many operators around the world adopted Option 3 at the time of their 5G launches.

NTT DOCOMO will continue its studies toward future means of migration including the introduction of 5GC, the new core network.

3.3.1 Option 3x architecture

As described above, user data is transferred by a SCG split bearer process to keep equipment development and operation costs down. In other words, control signals are exchanged between EPC and eNB while user data is exchanged between EPC and gNB. This architecture has been specified in the standard as Option 3x. The main feature of Option 3x architecture is that it enables NR to be accommodated by extending the S1 interface with eNB and the non-access stratum (NAS) interface with UE, which eases the impact on core network equipment and achieves both stable quality and early deployment.

*16 Carrier aggregation: A technology for increasing bandwidth by simultaneously transmitting/receiving signals over multiple component carriers.

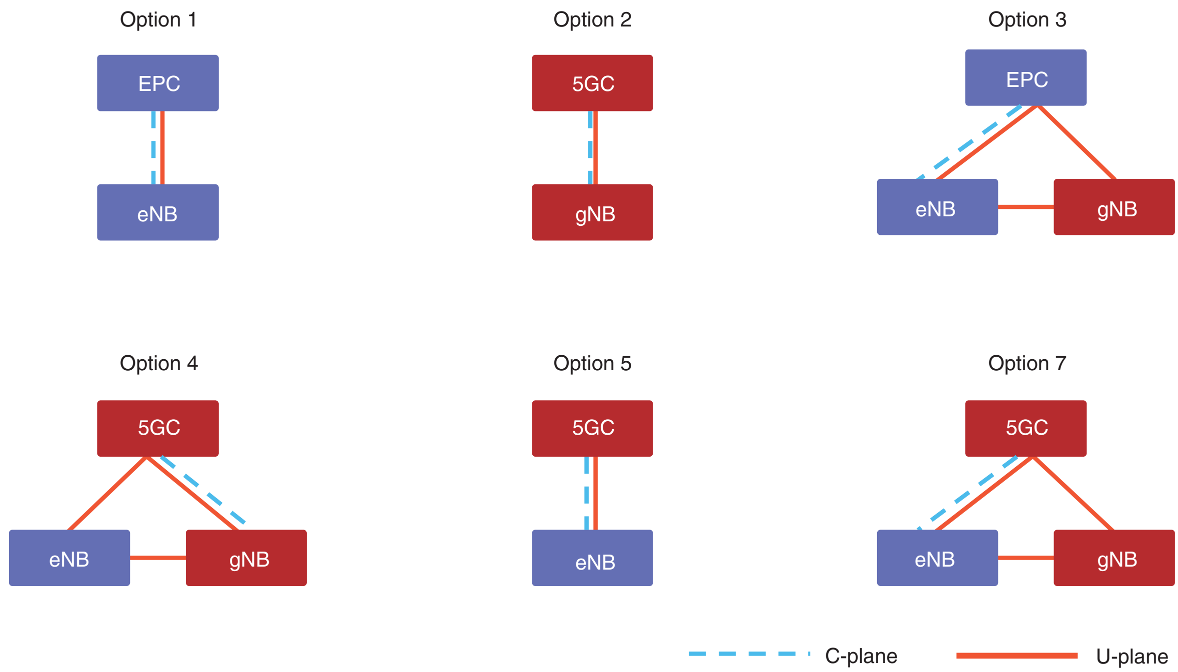


Fig. 7. Migration architectures toward 5G.

3.3.2 High-throughput support

At present, the 5G commercial service provides a maximum downlink throughput of 4.1 Gbit/s. NTT DOCOMO intends to continue its development efforts toward even higher data rates and larger capacities. However, EPC consists of various types of equipment of diverse levels of performance, and there are some types of equipment that cannot easily provide the throughput required by 5G. With this in mind, and considering that a serving gateway/packet data network gateway (S/P-GW) for performing data-transfer processing within EPC needs to be selected, NTT DOCOMO has developed a mechanism for selecting an S/P-GW that can provide 5G throughput.

In more detail, the existing method for selecting an S/P-GW via a mobility management entity is to use tracking area and access point name as keys to query the domain name system (DNS) and to then select an S/P-GW from the record included in the response. To this method, NTT DOCOMO added a value (+nc-nr) indicating 5G capability to the service parameters (network capability) included in the DNS response. For the 5G user, using this value as a basis for selecting an S/P-GW that can provide 5G throughput achieves high-speed communications.

3.3.3 Low-latency network

NTT DOCOMO began the provision of its docomo Open Innovation Cloud simultaneously with the launch of its 5G commercial service. The docomo Open Innovation Cloud provides a form of multi-access edge computing. It enables end-to-end communications latency to be shortened by deploying computing resources normally included in the core network at locations near user terminals. It also provides NTT DOCOMO's "Cloud Direct™" service that shortens network transmission latency by optimizing the transmission path between the connected terminal and the cloud platform.

4. Conclusion

This article described the development of radio base-station equipment and core network equipment at NTT DOCOMO to provide a 5G commercial service. Going forward, NTT DOCOMO will pursue co-creation opportunities through 5G with a variety of partners to contribute to the creation of a prosperous society while continuing to develop groundbreaking and advanced technologies.

This article is the reproduction of the special article published in NTT DOCOMO Technical Journal (Vol. 22, No. 2, Oct. 2020).

References

[1] MIC, "Information and Communications Council, Information and Communication Technology Sub-committee, New Generation Mobile Communication Systems Committee Report (July 31, 2018)," July 2018 (in Japanese).
https://www.soumu.go.jp/main_content/000567504.pdf

[2] MIC Press Release: "Approval to Applications for Authorization of Establishment Plans on Specified Base Stations for Spreading 5G Mobile Communications Systems," Apr. 2019.
https://www.soumu.go.jp/main_sosiki/joho_tsusin/eng/pressrelease/2019/4/10_01.html

[3] M. Fujii, S. Suwa, R. Toba, and T. Toeda, "Base-station Equipment with the Aim of Introducing 3.5-GHz band TD-LTE," NTT DOCOMO Technical Journal, Vol. 18, No. 2, pp. 8–13, Oct. 2016.
https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol18_2/vol18_2_003en.pdf

[4] T. Yoshihara, T. Toeda, M. Fujii, S. Suwa, and T. Yamada, "Radio Equipment and Antennas for Advanced C-RAN Architecture," NTT DOCOMO Technical Journal, Vol. 17, No. 2, pp. 19–24, Oct. 2015.
https://www.nttdocomo.co.jp/english/binary/pdf/corporate/technology/rd/technical_journal/bn/vol17_2/vol17_2_004en.pdf

[5] NTT DOCOMO Press Release: "DOCOMO to Commence Deployment of World's First 4G/5G Multi-vendor Radio Access Network Conforming to O-RAN Specifications," Sept. 2019.
https://www.nttdocomo.co.jp/english/info/media_center/pr/2019/0918_00.html

Trademark notes

Yuta Sagae
Manager, Corporate Strategy & Planning Department, NTT DOCOMO, INC.
He received a B.E. and M.E. in Tokyo Institute of Technology in 2003 and 2005. He joined NTT DOCOMO in 2005. He was in charge of 5G network system strategy and 5G base station equipment development.

Kohei Kiyoshima
Manager, Radio Access Network Development Department, NTT DOCOMO, INC.
He received a B.E. in electronics engineering and an M.E. in information science and technology in engineering from Kyoto University in 2006 and 2008. He joined NTT DOCOMO in 2008. He is currently in charge of 5G base-station equipment development.

Shinsuke Sawamukai
Director, Radio Network Control Group, NTT DOCOMO, INC.
He joined NTT DOCOMO in 2002 and has since been involved in research and development (R&D) activities related to the development of the base station equipment.

Keiichi Kanbara
Manager, Solution Integration Department, DOCOMO CS, INC.
He received a B.E. in electronics engineering and an M.E. in information science and technology from Hokkaido University in 2006 and 2008. He joined NTT DOCOMO in 2008. He is currently in charge of providing solutions for business utilizing 5G and IoT.

Yusuke Ohwatari
Manager, Radio Access Network Development Department, NTT DOCOMO, INC.
He received a B.E. in electronics engineering, an M.E. in information science and technology, and a Ph.D. in engineering from Hokkaido University in 2005, 2007, and 2013. He joined NTT DOCOMO in 2007. He is currently in charge of 5G base-station equipment development. His research interest includes MIMO systems for mobile communication systems.

Jo Takahashi
Chief Engineer, Packet Network Division, DOCOMO Technology, Inc.
He received a B.A. from Cornell University, USA, in 2004. Since joining DOCOMO Technology in 2005, he has been involved in R&D activities related to radio access network technologies for LTE/LTE-Advanced. After joining the current division in 2018, he has contributed to R&D for LTE/5G core network technologies.

Soft Error Countermeasure for 10G-EPON ONU

Ryota Kita, Kenta Ito, Haruka Nagoshi, Shun Morishima, Satoshi Shimazu, and Tomoaki Yoshida

Abstract

We developed a soft error countermeasure for the 10 Gigabit Ethernet passive optical network (10G-EPON) optical network unit (ONU). This countermeasure allows an ONU to detect a soft error and reboot by autonomously turning its power off/on. This feature reduces the rate of inquiries because of soft error failure in 10G-EPON ONUs. It also reduces the number of operations required to respond to failures in telecommunication services.

Keywords: ONU, soft error, reboot

1. Overview of the 10G-EPON system

The passive optical network (PON) system is a fiber-to-the-home optical access system that uses an optical splitter to branch optical signals from a single optical fiber shared by multiple users. The 10 Gigabit-Ethernet PON (10G-EPON) system is a PON system with a maximum transmission rate of 10 Gbit/s. This system consists of an optical line terminal (OLT) installed at the central office, optical network units (ONUs) installed in each user's home, an optical splitter, and optical fiber network connecting them (**Figs. 1** and **2**).

2. Soft error in the ONU

A soft error is an event in which a bit in memory is inverted by electrical noise. Operation failures caused by soft errors can be recovered by resetting the electronic device or overwriting the data [1]. Soft errors are primarily caused by cosmic rays (high energy protons). Cosmic rays collide with atomic nuclei (nitrogen and oxygen) to generate neutron rays, which then collide with silicon atomic nuclei in semiconductor devices, generating electrical noise (**Fig. 3**).

If ONU operation fails due to soft error, network service may be disconnected. An example of a soft error failure is that due to electrical noise generated in a semiconductor device, where bit inversion occurs in the main signal path control function written in the volatile memory (random access memory: RAM) of the PON-Media Access Control (MAC) large-scale integrated circuit. The main signal is thus disrupted, resulting in communication interruption. In this case, by rebooting the ONU (turning the power off/on), the information or program in RAM is refreshed from nonvolatile memory (read only memory: ROM), which recovers communication. When such a soft error failure occurs, it is necessary to restart the ONU by turning it off and on at the user's home.

Soft errors occur very infrequently in semiconductor devices, making it difficult to isolate and identify faults. For ONUs that are distributed and deployed in large numbers, it is important to improve the efficiency of responding to soft error failure.

Further advances in semiconductor miniaturization are inevitable given the demand for 10 Gbit/s-class high-speed communication. As these semiconductor devices hold less charge in a memory cell, they are more susceptible to neutron rays. Therefore, it is assumed that the occurrence rate of soft errors and frequency of troubleshooting will increase.

Error correction for soft errors includes (1) autonomous correction by using hardware, (2) autonomous correction by using a device-control program, and (3)

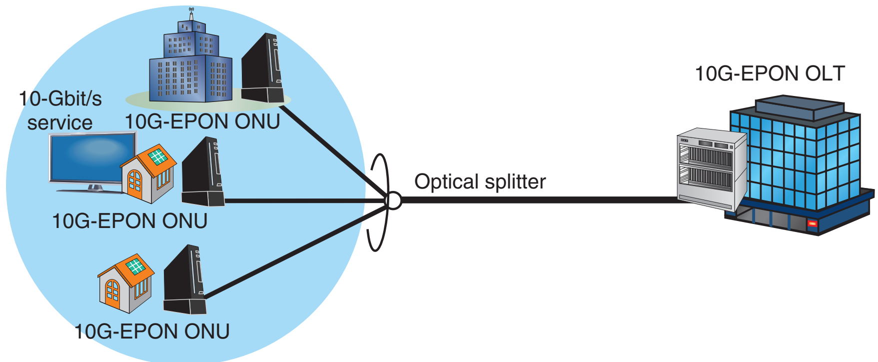


Fig. 1. 10G-EPON system.

Fig. 2. 10G-EPON ONU.

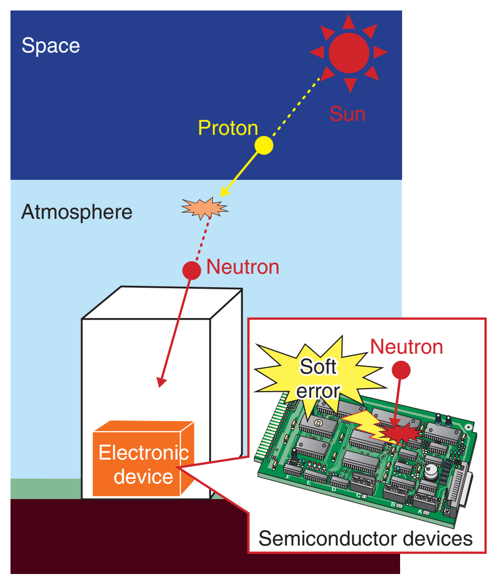


Fig. 3. Mechanism of soft error occurrence.

correction by manual intervention [2]. (1) Autonomous correction by using hardware uses components with functions such as error check code (ECC) correction. (2) Autonomous correction by using a device-control program includes methods such as system reset and device reset. The system can be reset without any special control, just like a power off/on restart. Device reset resets just the target device (components), which shortens the correction time, but requires coordination with peripheral circuits for state matching, resulting in complex control. (3) Manual correction can be done by remote control reset by the maintainer or by instructing the user to turn the power off/on. Either approach takes a long time to correct the problem, and it is necessary for the telecommunication company to respond to the problem.

For a 10G-EPON ONU, we developed a function (autonomous reset function) that corresponds to the system reset for (2); the ONU detects a soft error and restarts itself by turning its power off and on autonomously. Compared with (1) and device reset for (2), this function is economical, easy to implement, and reduces the amount of manual operation for error correction.

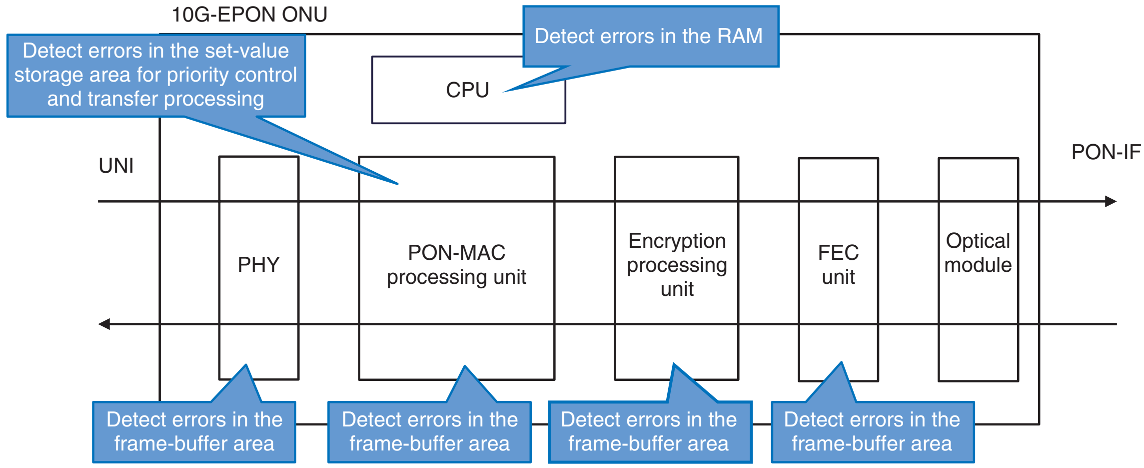


Fig. 4. 10G-EPON ONU functional block and soft error detection.

3. Soft error detection and autonomous reset-target classification

The functional block of an ONU is shown in **Fig. 4**. An uplink signal input from a user network interface (UNI) connected to user equipment, etc. is received by the physical layer (PHY), priority control and transfer processing are executed by the PON-MAC processing unit, and if enabled in the encryption unit and forward error correction (FEC) unit, the signal is encrypted and an error correction code is assigned to each block and transmitted from the optical module to the OLT. In addition, the downlink signal input from the OLT PON interface (IF) is received by the optical module, corrected by the FEC unit, decoded by the encryption processing unit, subjected to priority control and transfer processing by the PON-MAC processing unit, and transmitted from the PHY to the user equipment.

Soft error is first detected in the frame-buffer area of the RAM of each functional block of the main signal path through which the uplink and downlink signals mentioned above pass. Most soft errors detected are bits in a single frame in the frame-buffer area, which are discarded after detection and have little effect on communication. However, in rare cases, similar soft errors are continuously detected in the frame-buffer areas of the PHY, PON-MAC processing unit, encryption unit, and FEC unit. In this case, since the soft errors are continuous and extend over multiple frames, it is assumed that there is a problem with the continuity of the main signal in the ONU. Therefore, it is determined that a soft error event occurred, and the autonomous reset function is triggered.

Our countermeasure also detects errors in the RAM of the central processing unit (CPU) and the set-value storage area for priority control and transfer processing of the RAM of the PON-MAC processing unit. If a soft error is detected in the RAM of the CPU, it is assumed that CPU processing is abnormal. If a soft error is detected in the transfer-control-setting table of the PON-MAC processing unit, it is also assumed that there is a problem with the main signal transfer to the OLT. Therefore, both are judged to be soft errors, which triggers the autonomous reset function.

4. Autonomous reset function as a soft error countermeasure

Figure 5 shows the transition flow of the autonomous reset function. A semiconductor device affected by a neutron beam may suffer a physical defect as well as soft error. A physical defect is an event in which the structure of the semiconductor device degrades and malfunctions due to the effect of neutron beams; such failures (physical defect failure) cannot be recovered by restarting the semiconductor device or overwriting data. This suggests that the autonomous reset function may repeatedly turn the power supply off and on.

Our solution is a mechanism that memorizes the number of resets and, when the number of resets exceeds a threshold within a certain period, a physical-defect failure is determined and the ONU communication function is terminated along with autonomous

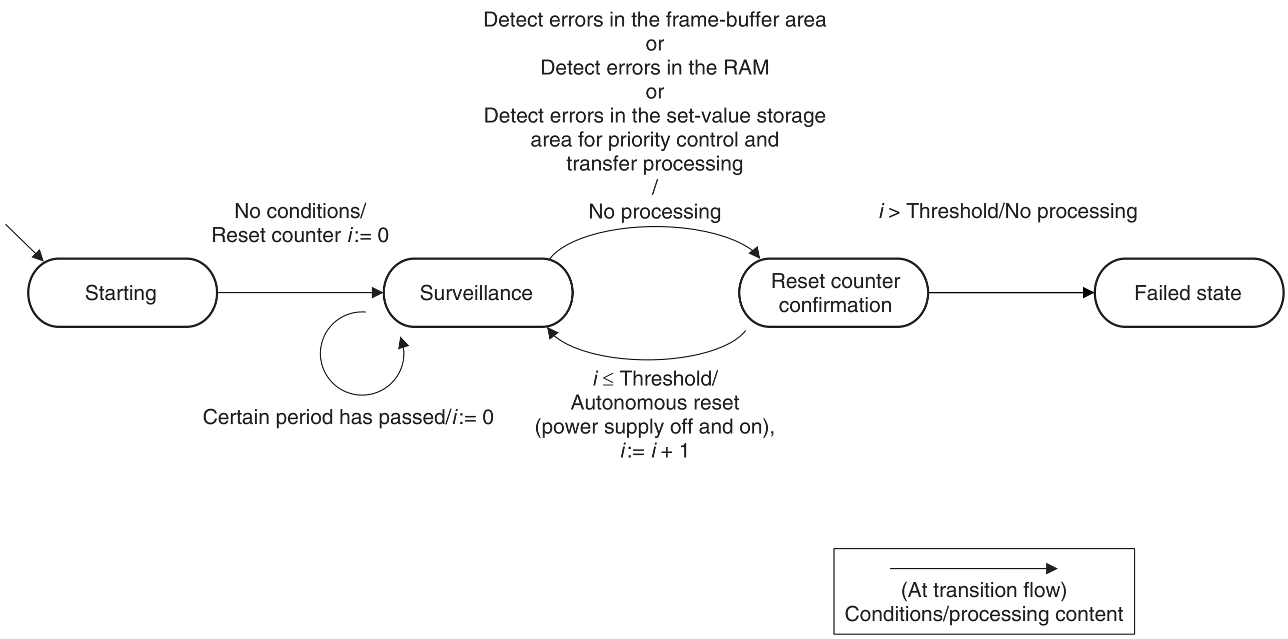


Fig. 5. Transition flow of the autonomous reset function.

reset. Specifically, reset counter i is provided to record the number of resets in the ONU. After the ONU is started, reset counter i is initialized to 0, and each autonomous reset increments the counter value by 1. When autonomous reset occurs repeatedly, reset counter i increases, and if the counter value exceeds the specified threshold, the device transitions to a failed state. This prevents endless repetition of autonomous reset and enables soft error failure to be distinguished from physical-defect failure.

5. The results of our soft error countermeasure with neutron irradiation test

Neutron irradiation tests are generally conducted to reproduce soft errors. To measure the effectiveness of our soft error countermeasure, we subjected 10G-EPON ONUs running the autonomous reset function to the neutron irradiation test [3] provided by NTT Advanced Technology based on the Telecommunication Technology Committee (TTC) standard JT-K 130 [4].

In the neutron irradiation test, the target (beryllium) is irradiated with protons accelerated using a cyclotron proton accelerator to generate neutron beams, and several 10G-EPON ONUs are irradiated simultaneously. This test uses a very high neutron intensity corresponding to about 100 million times the natural exposure rate (**Fig. 6**).

A main signal was connected to the 10G-EPON ONUs using an OLT and traffic-generator analyzer. To assess the possibility of log-file destruction due to the effect of neutron beams during the test, we collected the logs inside the 10G-EPON ONUs in real time and measured all autonomous resets. Through this test, the 10G-EPON ONUs were exposed to neutron beams equivalent to about 110,000 years per unit on average. Due to the irradiation, soft error was detected 1150 times per unit on average; among them, errors in the frame-buffer area, not the subject of autonomous reset, were confirmed 1104 times per unit, and fault recovery by autonomous reset was confirmed 46 times per unit.

Because the autonomous reset function worked as expected, it was confirmed that when the ONU exhibits a soft error fault, it can recover autonomously; thus, this is effective as a soft error countermeasure. If we assume that 10 million units will be deployed, we can expect to reduce the number of failures per year by about 4200.

6. Future plans

We developed a countermeasure against soft errors, which are expected to become more frequent with the adoption of 10G-EPON ONUs. This countermeasure

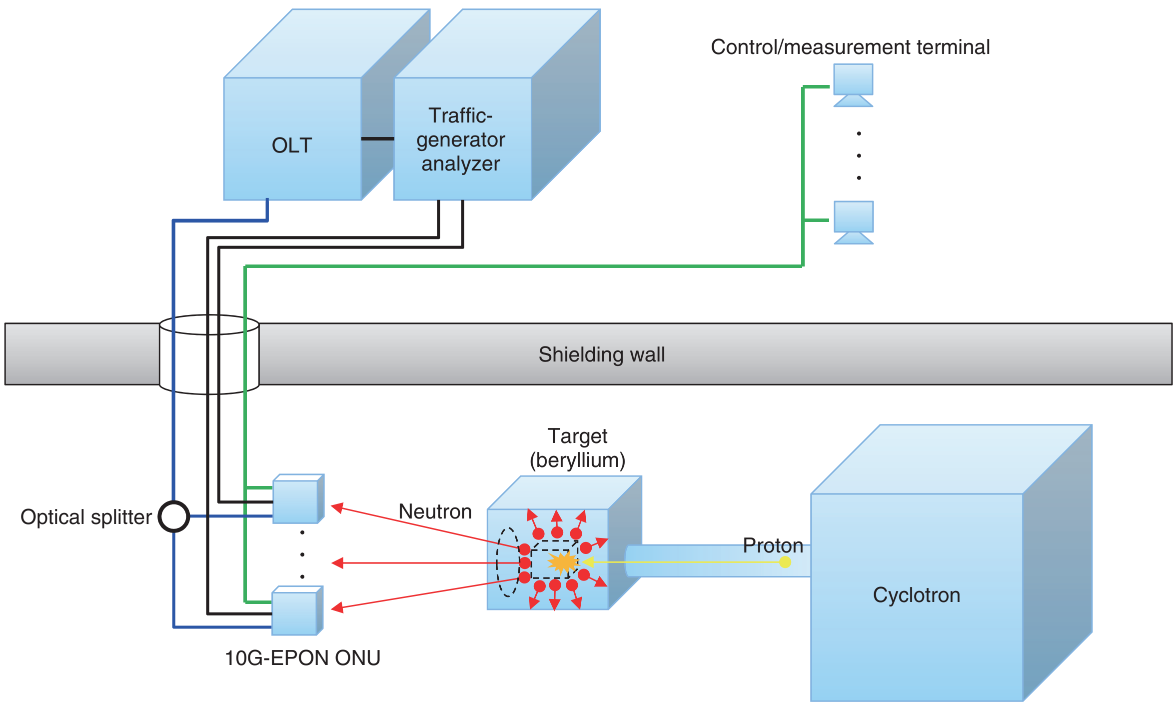


Fig. 6. Neutron irradiation test.

will contribute to the reduction in ONU failures and in manual operations required by the user or carrier personnel. In the future, we will examine the application of this countermeasure to other types of ONUs.

References

[1] TTC Standard JT-K 124: “Overview of particle radiation effects of telecommunication systems,” 2018.
[2] TTC Standard JT-K 131: “Design methodologies for telecommunication systems applying soft error measures,” 2019.
[3] NTT Advanced Technology, Soft Error Test Service, https://www.ntt-at.com/product/softerror/
[4] TTC Standard JT-K 130: “Neutron irradiation test methods for telecommunication equipment,” 2019.

Ryota Kita

Engineer, Optical Access Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E. from Ehime University in 2010 and joined NTT WEST the same year. He has been with NTT Access Network Service Systems Laboratories since 2019 and has been engaged in the research and development of Gigabit Ethernet (GE)-PON and 10G-EPON systems.

Kenta Ito

Engineer, Optical Access Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E. in engineering from Yamagata University in 2014, and an M.S. in information science from Tohoku University, Miyagi, in 2016. In 2016, he joined NTT EAST, where he engaged in the maintenance operation of the Next Generation Network. In 2018, he joined NTT, where he has been engaged in the research and development of optical access systems. His current research interests include GE-PON and 10G-EPON systems.

Haruka Nagoshi

Engineer, Optical Access Systems Project, NTT Access Network Service Systems Laboratories.

She received a B.E. from University of Teacher Education Fukuoka in 2010, and an M.S. from Yamaguchi University Graduate School of Science and Engineering in 2012. She joined NTT WEST in 2015. She has been with NTT Access Network Service Systems Laboratories since 2018 and has been engaged in the research and development of GE-PON and 10G-EPON systems. She is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

Shun Morishima

Manager, Network Department, NTT WEST.

He received a B.E. from Kobe University in 2009 and joined NTT WEST the same year. From 2017 to 2019, he was with NTT Access Network Service Systems Laboratories, where he engaged in the development of 10G-EPON systems.

Satoshi Shimazu

Senior Research Engineer, Optical Access Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E. and M.E. from Tohoku University, Miyagi, in 2006 and 2008. He joined NTT Access Network Service Systems Laboratories in 2008, where he has been engaged in the research and development of GE-PON and 10G-EPON systems.

Tomoaki Yoshida

General Manager, Optical Access Systems Project, NTT Access Network Service Systems Laboratories.

He received a B.E., M.E., and Ph.D. in communication engineering from Osaka University in 1996, 1998, and 2007. In 1998, he joined NTT Multimedia Systems Development Center, where he engaged in the development of asynchronous transfer mode optical access systems. In 1999, he moved to NTT Access Network Service Systems Laboratories. Since 2000, he has been engaged in research on next-generation optical access networks and systems. From 2013 to 2015, he was involved in the research of wavelength division multiplexing/time division multiplexing-PON and worked on the standardization of optical access systems such as NG-PON2 (next-generation passive optical network stage 2). He is a member of IEEE and IEICE.

Trends in Standardization of Mapping and Multiplexing Technologies for Optical Transport Networks

Kengo Shintaku, Shoichiro Kuwahara, and Shoukei Kobayashi

Abstract

This article introduces recent standardization activities in the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) related to optical transport networks, including fifth-generation mobile transport networks.

Keywords: OTN, OTN security, 5G mobile transport

1. Introduction

The International Telecommunication Union - Telecommunication Standardization Sector (ITU-T) Study Group 15 Question 11 (Q11/15) have addressed the standardization of framing and multiplexing technologies for optical transport networks (OTNs) used in the backbone networks of telecommunication carriers. Of particular note, the OTN specified in ITU-T Recommendation G.709 has been widely applied to optical transmission systems as key technology to achieve long-distance and large-capacity transmission as well as cost reductions. Therefore, NTT has been extensively contributing to OTN standardization activities.

Section 2 outlines the OTN, specifically, the basic and latest framing/multiplexing technologies. Section 3 explains Flexible OTN (FlexO), an enhanced interface that provides flexibility and scalability. Security has also become important in transmission networks, and Section 4 describes OTN and FlexO encryption technologies. Section 5 explains fifth-generation (5G) mobile transport networks. Finally, Section 6 introduces future activities.

2. OTN multiplexing and mapping structure

An OTN has two functions; (1) packaging various client signals into OTN frames and (2) multiplexing OTN frames into a higher-speed OTN frame (see **Fig. 1**). As shown in **Fig. 2(a)**, an optical channel transport unit-k (OTUk) frame (k = 1,2,3,4) uses the fixed frame of 4 × 4080 bytes and contains three areas; the payload area accommodates the client signal, overhead (OH) area carries maintenance and control information, and forward error correction (FEC) area transfers redundant data to correct the bit errors that may occur during transmission. The OH area in Fig. 2(a) consists of three regions, optical payload unit (OPU) OH, optical data unit (ODU) OH, and optical transport unit (OTU) OH. The OPU OH includes information that supports the adaptation of client signals, ODU OH includes information for maintenance and operational functions to support ODU connections, and OTU OH includes information for maintenance and operational functions to support OTU connections.

The OTN was originally standardized in 2001 with three bitrate classes, OTU1 (2.67 Gbit/s), OTU2 (10.7 Gbit/s), and OTU3 (43.0 Gbit/s) to accommodate

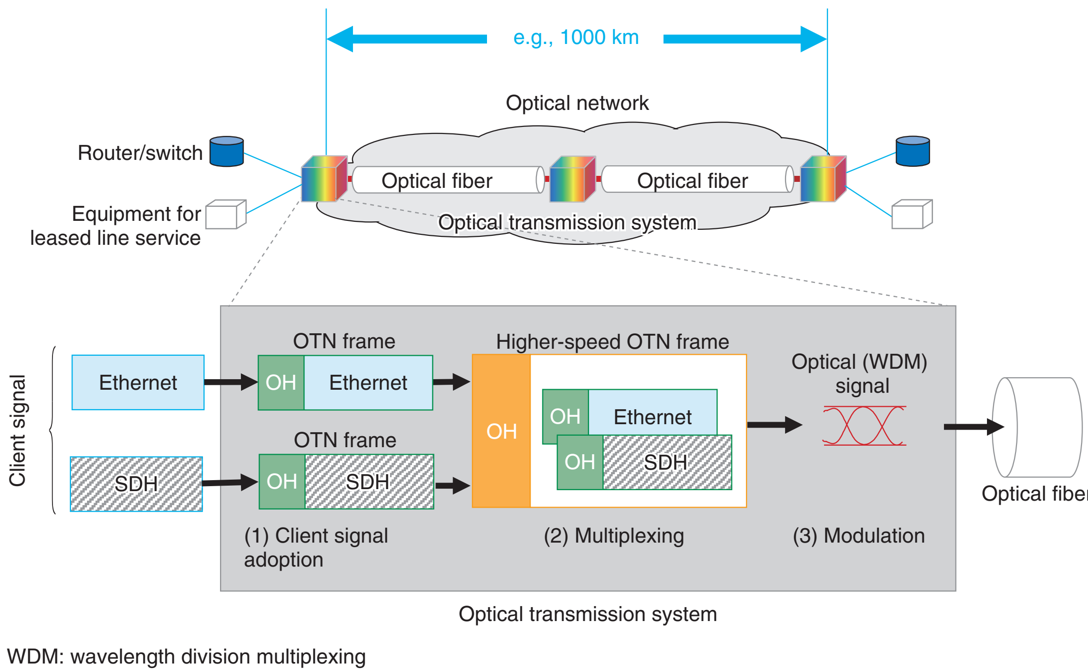


Fig. 1.   OTN.

synchronous digital hierarchy (SDH), which was the dominant client signal at that time. Next, OTU4 (112 Gbit/s) was specified in 2010 to transport 100 Gigabit Ethernet (100GbE) [1]. Subsequently, with the development of higher-speed client signals, such as the 200GbE and 400GbE specified in the Institute of Electrical and Electronics Engineers (IEEE) 802.3, Q11/15 began discussing an OTN for rates beyond 100 Gbit/s (Beyond 100G; B100G). The B100G OTN takes into account flexibility and extensibility, which are considered important features in addition to transmission capacity. Therefore, the optical channel transport unit-Cn (OTUCn) frame (The index Cn is used for n × 100G (C = 100G); n is a positive integer) was specified to have an n multiplexing structure that concatenates OTUC frames that have 100G transmission capacity, as shown in **Fig. 2(b)**. The OTUCn frame also supports multi-carrier transmission by using multiple optical signals (wavelengths) in parallel, and the optical interface can be configured flexibly by multiplexing OTUC frames according to the transmission speed per optical signal. The FEC area, which is one of the characteristic functions of the OTUk frame, is separate from the OTUCn frame and is specified not in G.709 but in G.709.1, G.709.2, and G.709.3. This is because it is necessary to specify the optimum FEC coding scheme and area according to the physical transmission speed of the optical signal. The next section describes how to transmit OTUCn frames in combination with FEC.

3. OTN enhancements

Standardization activities have recently progressed towards aggregating multiple physical interfaces to configure large-capacity links. A well-known example of aggregated virtual links is the Link Aggregation Group (LAG) of IEEE802.3ad, which bonds Ethernet physical interfaces. For example, a 400-Gbit/s aggregated link can be configured by bonding four 100GbE interfaces with this LAG. However, the bandwidth of a data flow is limited to that of the physical interface since each single data flow is transported via only one of the physical interfaces. In this 400-Gbit/s-link LAG, a single flow cannot exceed 100 Gbit/s. This limitation between the data flow and physical interface can induce traffic imbalances on some physical links. Therefore, it remains difficult to efficiently use the large-capacity link configured using the LAG. To overcome this challenge, the Optical Internetworking Forum (OIF) standardized Flexible Ethernet (FlexE) in 2016. FlexE configures large-capacity links by bonding multiple Ethernet physical interfaces similar to the

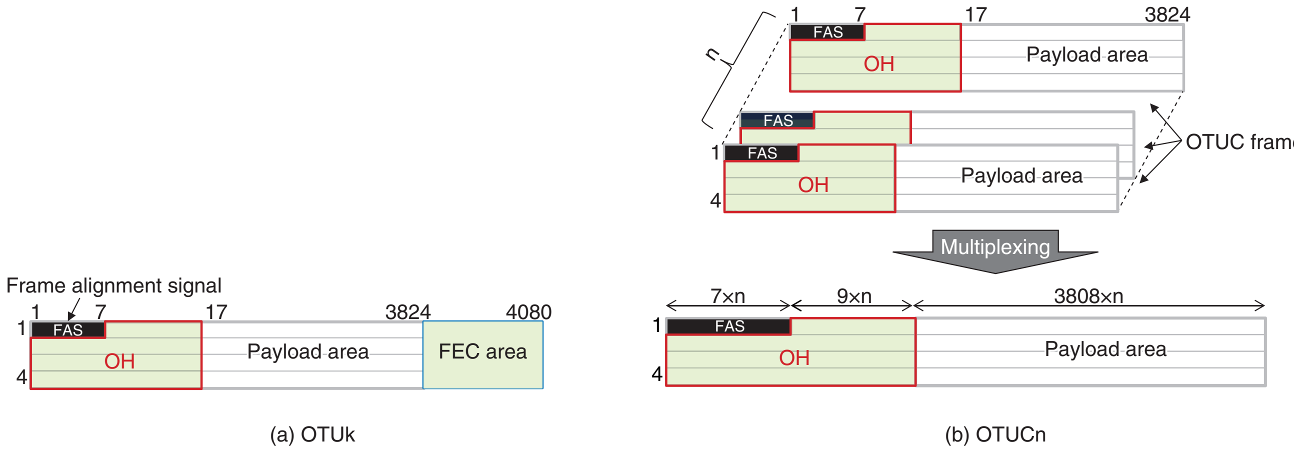


Fig. 2. OTU frame structure.

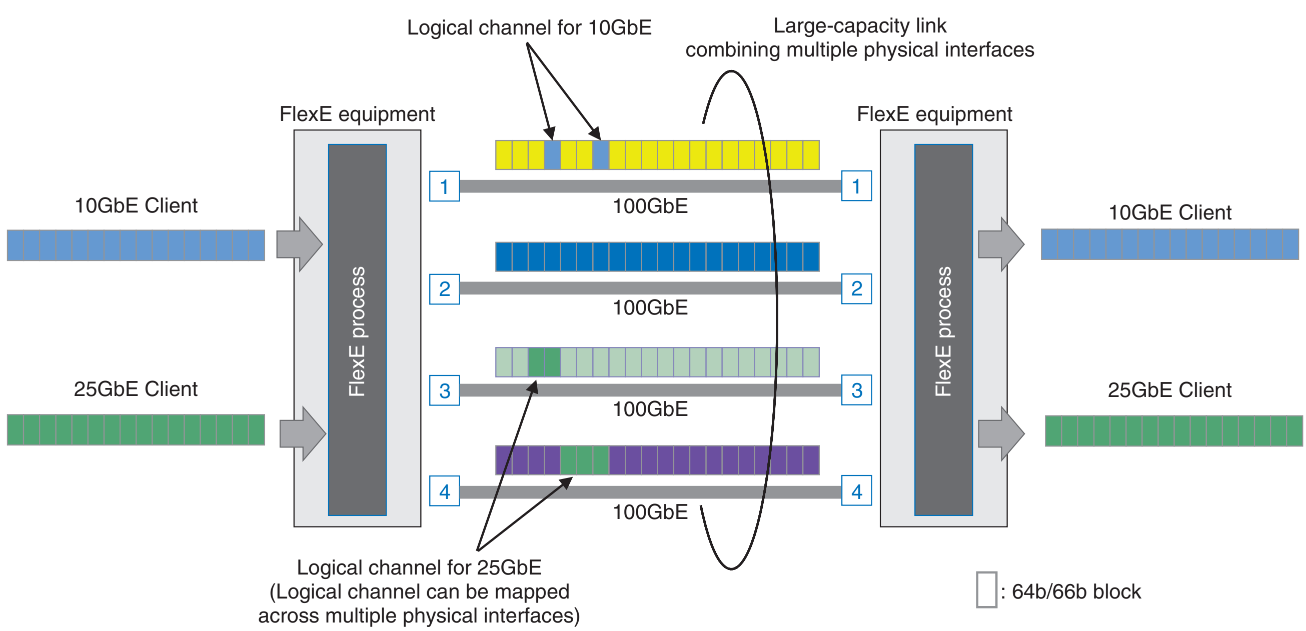


Fig. 3. FlexE.

LAG, but it can freely provide logical channels across multiple physical interfaces. The logical channel acts as an Ethernet link equivalent to a real physical interface, as shown in **Fig. 3**.

In Q11/15, FlexO, which bonds multiple physical interfaces, was specified in Recommendation G.709.1 in 2017. FlexO transfers OTUCn, making the optical transmission flexible and scalable from both logical and physical aspects. **Figure 4** shows an example of G.709.1; one OTUCn frame is mapped into n × FlexO frames, with each FlexO frame containing one OTUC frame. The n × FlexO frames are combined, interleaved, and mapped into FlexO-x-<fec>-m (m = [n/x], where [ ] is the ceiling function). For example, an OTUC4 frame can be transported using two 200G FlexO frames (m = 2, x = 2) or a single 400G FlexO frame (m = 1, x = 4). Since it is assumed that FlexO will be applied to communication between adjacent network equipment, G.709.1 adopted KP4 FEC (RS10(544,514)) as FlexO FEC (-<fec>-),

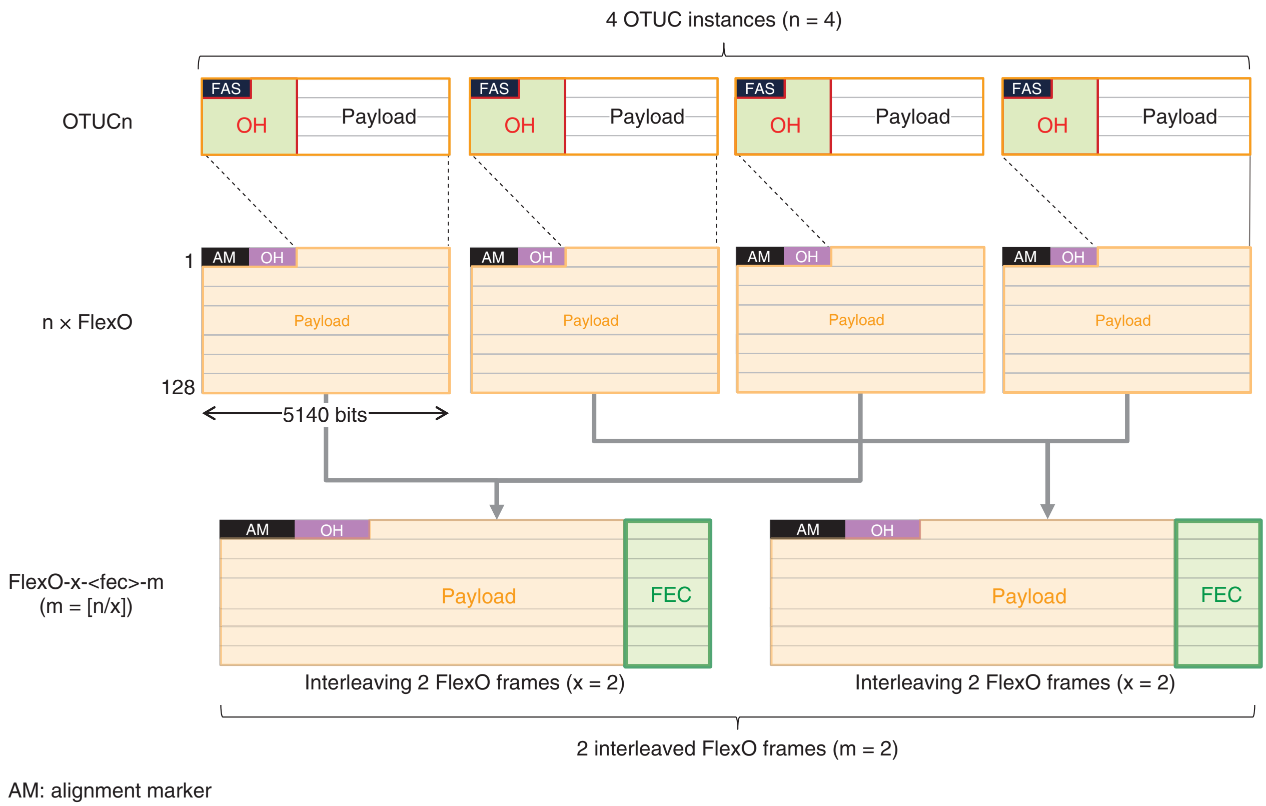


Fig. 4. FlexO frame example.

which is the same as that of 200G and 400G Ethernet, which have transmission distances of up to 10 km.

NTT has been actively involved in and contributing to FEC discussions, believing that it is necessary to specify an FEC for metro and long-distance transmission to apply OTUCn/FlexO to practical OTNs. We took a leadership role in standardizing the best FEC possible. For example, we summarized various opinions from many network service providers and vendors and submitted many contribution documents to Q11/15 and Question 6 (Q6/15). In addition to standardization activities, NTT also has conducted and demonstrated inter-operability verification trials on a candidate FEC between optical devices developed by different vendors. After more than five years of these activities, Staircase FEC was successfully standardized in 2018 as 100G FlexO FEC (Recommendation G.709.3). At the same time, Staircase FEC was also adopted as the FEC for long-reach OTU4 application in Recommendation G.709.2. NTT has also played a major role in open FEC (OFEC) standardization for 450-km 200G/400G FlexO applications (G.709.3 in 2020). OFEC is also adopted by the Open ROADM (reconfigurable optical add/drop multiplexer) Multi-Source Agreement (MSA).

4. OTN secure transmission

Due to the recent worsening of security risks, Q11/15 began discussing OTN cryptographic standardization, which covers two functions: encryption to prevent wiretapping and authentication to verify if the received frames were tampered with. Prior to the standardization study, there was a discussion about the order in which to proceed with the standardization. Since FlexO OH can carry more cryptographic control information than OxU (OPU/ODU/OTU) OH and it seemed easy to study, Q11/15 decided to start with FlexO encryption. OxU encryption will follow after FlexO.

Figure 5 illustrates FlexO frame encryption. The encryption area of FlexO consists of a payload region and a portion of the OH region. With regard to the encryption algorithm, it has been agreed to adopt a current algorithm (e.g., Advanced Encryption Standard (AES)*) rather than develop new ones. Studies

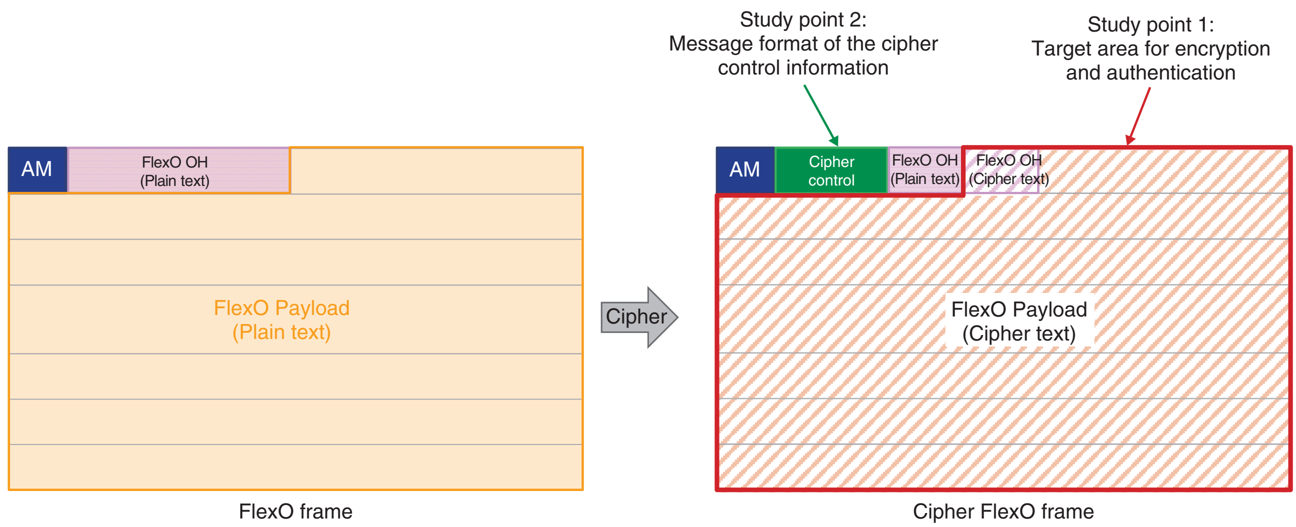


Fig. 5. Example of FlexO frame encryption.

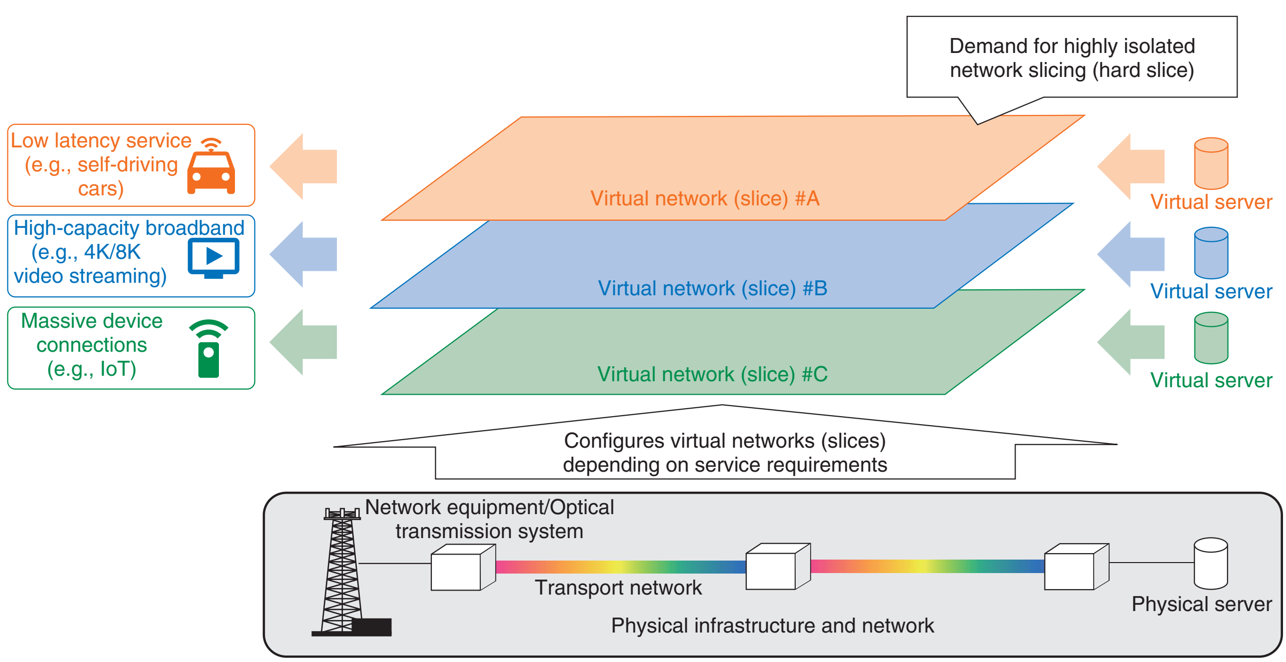


Fig. 6. Characteristics of 5G networks.

and discussion are still on-going, but agreement in the near future seems likely.

5. Slicing techniques for 5G mobile transport network

One of the characteristic requirements for 5G mobile networks is network slicing (**Fig. 6** [2]). Network slicing involves managing physical equipment (physical resources), such as servers and routers, as resources that can be partitioned virtually (virtual

* AES: An encryption algorithm standardized as FIPS197 by the National Institute of Standards and Technology, which is a standardization body of the United States.

links, virtual network functions, etc.).

In Q11/15, these network-slicing requirements are specified in G.8300; it summarizes the requirements for 5G mobile transport networks and a new multiplexing technology, tentatively called Metro Transport Network (MTN). MTN requires strongly isolated network slicing (hard slicing) to prevent interference among slices that access the common physical network.

Two candidate techniques have been proposed to achieve hard slicing; one is an OTN-based technique and the other is a FlexE-based technique. In the former, the payload area of high-speed OTN frames is divided into small areas called time slots, and low-speed frames are allocated to them so that hard slicing can be achieved by treating the OTN time slots as virtual resources. In the latter, the logical channel is used to implement hard slicing. These two techniques are specified in Supplemental G.sup.67.

6. Future activities

OTN technologies, which play important roles in the information society, will become even more important. NTT will continue contributing to standardization activities in ITU-T while staying informed about technological trends in optical transmission and advancing discussions in other standardization bodies and MSAs.

References

[1] T. Ohara and O. Ishida, "Standardization Activities for the Optical Transport Network," NTT Technical Review, Vol. 7, No. 3, 2009. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr200903gls.html

[2] S. Yasukawa, H. Sato, T. Hirota, T. Tojo, K. Endo, Y. Kasahara, and H. Suzuki, "Research toward Realizing a Future Network Architecture," NTT Technical Review, Vol. 16, No. 6, 2018. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201806fa5.html

Kengo Shintaku

Engineer, NTT Network Innovation Laboratories.

Since joining NTT Network Innovation Laboratories in 2013, he has been researching OTNs. He has been participating in ITU-T SG15 activities addressing OTN technologies since 2019.

Shoichiro Kuwahara

Senior Research Engineer, NTT Network Innovation Laboratories.

He received an M.S. in electrical engineering from Shizuoka University in 1995. After he joined NTT Optical Network System Laboratories the same year, he engaged in the research and development of adaptive chromatic dispersion equalization and standardization activities in ITU-T SG15 Q6. He is currently with NTT Device Innovation Center.

Shoukei Kobayashi

Assistant Manager, NTT Electronics Corporation.

He received a B.E. and M.E. in applied physics from Nagoya University in 1992 and 1994. In 1994, he joined NTT, where he was engaged in research on optical and wireless access network systems, high-speed Ethernet transport, terabit-class local area networks, and virtualized network architecture, and in standardization activity in IEEE 802.3. In 2009, he joined NTT EAST as a technical manager. In 2014, he again joined NTT and kicked off research in autonomous operation and maintenance systems. He is currently engaged in standardization activities in ITU-T SG15, Open ROADM MSA, and Open-ZR+ MSA.

Manual for Assisting with Troubleshooting in PHS Digital Cordless Telephone Systems

Technical Assistance and Support Center, NTT EAST

Abstract

This article introduces a manual for assisting with troubleshooting in digital cordless telephone systems. The Technical Assistance and Support Center, NTT EAST, published this manual to help service personnel in the field. This is the sixty-first article in a series on telecommunication technologies.

Keywords: digital cordless telephone, PHS, DCL, DECT

1. Introduction

Digital cordless telephones based on the personal handy-phone system (PHS) (hereafter, called the DCL) enable voice communications between a master unit called a cell station (CS) and a slave unit called a personal station (PS) via digital wireless communications. The DCL is widely used by business customers and consumers in Japan. On the other hand, the digital enhanced cordless telecommunication system (DECT), which uses the same frequency band as the DCL (i.e., 1.9-GHz band), has come into wide use as digital cordless telephones and intercoms for consumers. This indicates that these two wireless systems, the DCL and DECT, co-exist in the same frequency band, and the number of technical consultations concerning problem cases has been increasing.

The Technical Assistance and Support Center (TASC) of NTT EAST has published a manual to provide efficient and simple troubleshooting methods for digital cordless telephones for maintenance personnel on site. This manual presents simple techniques for evaluating wireless communication conditions without the need for dedicated measurement instruments, such as a spectrum analyzer or wireless protocol analyzer. It also provides a simple procedure on how to determine the cause of a problem.

2. Features of DCL and DECT

The DCL [1] and DECT [2], which both use the 1.9-GHz band, have the following features (their frequency allocations are illustrated in **Fig. 1** [3]).

- DCL: When a call channel is used, the carrier-sensing function automatically selects an unoccupied channel. It automatically switches channels when interferences occur during a call.
- DECT: Before the revision of the Radio Law in October 2017, it had five channels (from ch 1 to ch 5), and the use of three channels (from ch 2 to ch 4) was prohibited when a DCL radio wave is detected for the purpose of protecting the DCL control channels (ch 12 and ch 18). After the revision, the total number of channels was increased to six. Currently, the use of ch 1, ch 2, ch 5, and ch 6 is allowed and the use of ch 3 and ch 4 is conditionally allowed depending on their transmission power.

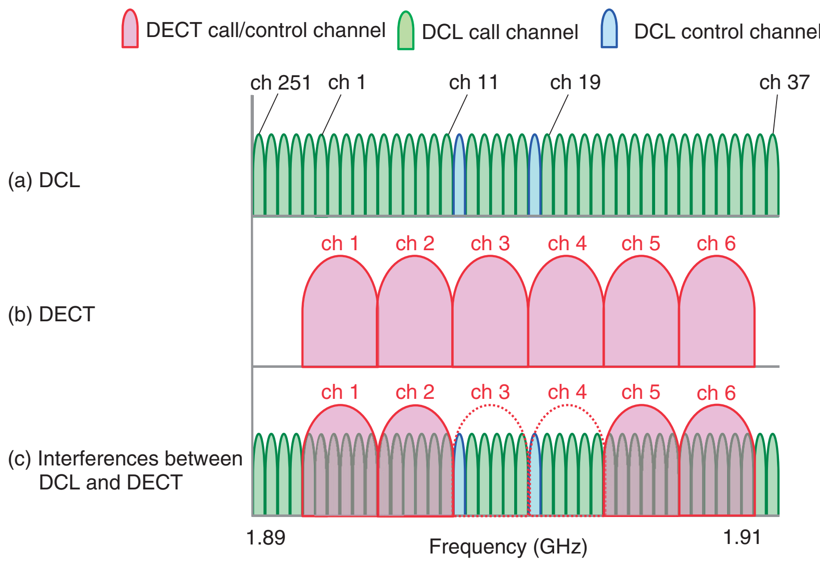


Fig. 1. Frequency allocation of the DCL and DECT.

Table 1. List of estimated causes and countermeasures obtained from DCL problems cases.

Problem	Estimated cause	Countermeasure
• Call is breaking up • No sound • Intermittent sound, etc.	Handover*	Placement redesign of CS
	Insufficient received signal strength	• Placement redesign of CS • Installation of additional CS
	Interference with the DECT	• Turning off DECT terminal • Unification of wireless system (DCL or DECT)

*Even in the case of normal handover operation, the call is interrupted for about 1 s.

3. DCL problem-isolation process flow

Table 1 summarizes the problems, estimated causes, and their countermeasures concerning the DCL categorized using the data gathered from cases we have dealt with thus far. The problems concerning the DCL claimed by customers are roughly divided into (i) "call is breaking up," (ii) "no sound," and (iii) "intermittent sound." There are three factors that lead to such problems: (a) handover, (b) insufficient received signal strength, and (c) interference with other wireless systems. The following are examples of such problems:

Let us consider a case in which multiple CSs are installed in an area and a customer holds a PS that is registered with one of those CSs. The PS monitors radio waves from the registered CS and receives radio waves from other unregistered CSs. It switches its registration to another CS when the received signal strength of the radio waves becomes weaker due to the customer moving around. This operation is called *handover* and a time lag of about one second occurs by switching CSs, which is a normal operation of wireless communication. However, some customers claim that a problem occurred and say, "the call was interrupted" or "no sound." In another case in which some CSs are closely located and their reach overlaps, a PS could attempt to keep receiving radio waves from the registered CS even when the received signal strength decreases. In this case, the handover does not work well. This may increase errors and lead to deterioration in call quality.

To solve these problems, the terminal-maintenance-mode of the PS enables easy checking of radio communication on site. This mode enables checking of the received signal strength (level), identifier (ID) of the registered CS (BS-ID), and frame error rate (FER). Various causes of a problem can be isolated by

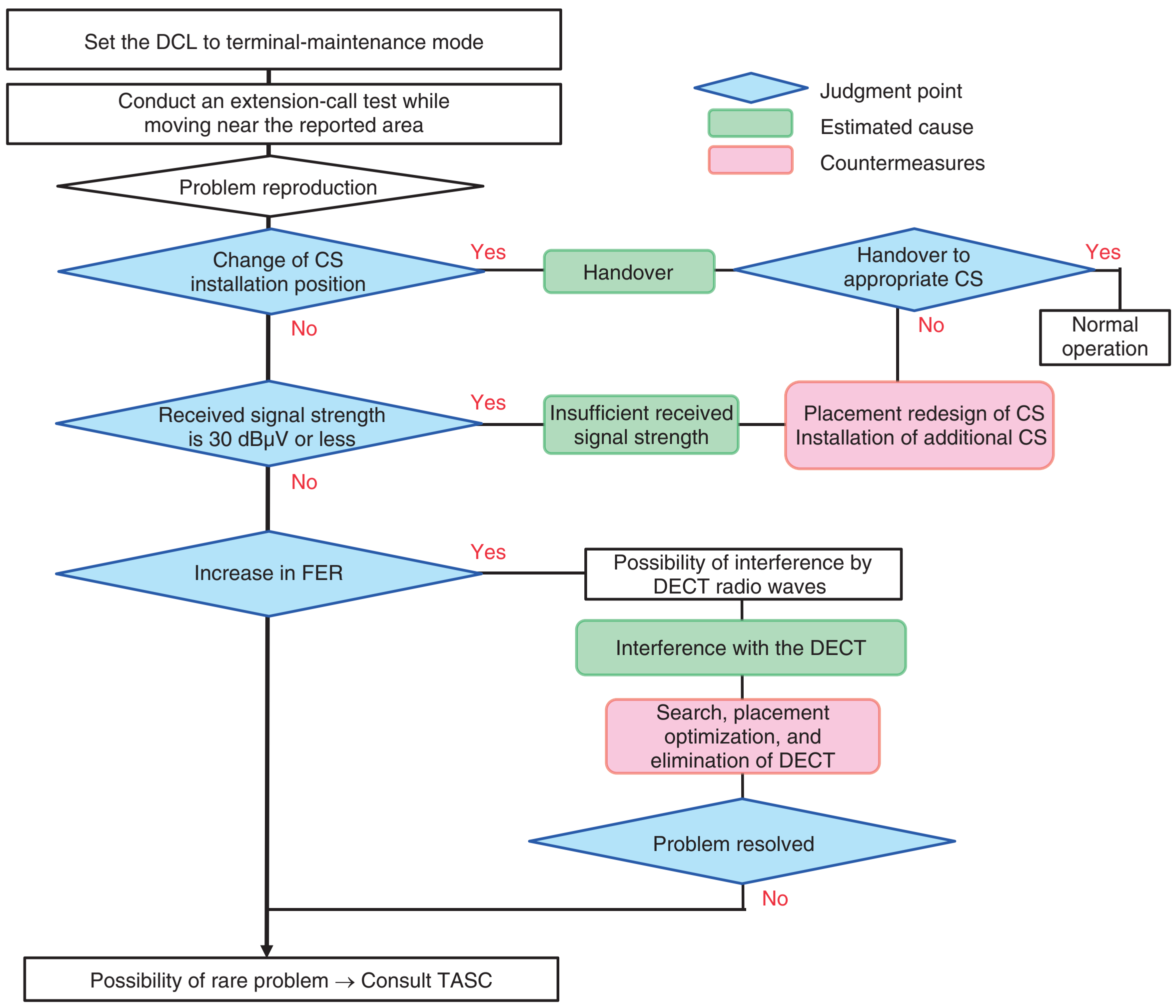


Fig. 2. Flow chart of problem isolation in the DCL.

comparing these values before and after the problem occurred.

A troubleshooting flow chart is shown in **Fig. 2**. Before troubleshooting, the area where the problem occurred is estimated through consultations with the customers, and reproduction of the claimed problem is attempted. After the problem is observed, the PS is set to the terminal-maintenance mode, and three parameters are evaluated. First, the ID of the currently registered CS (displayed on the PS) is checked, and the PS is moved to where the problem occurred. The ID of the registered CS before handover is compared with that after the handover. If the CS selection is appropriate, the received signal strength is then checked. The DSL determines if there is insufficient received signal strength and executes handover when the received signal strength is 30 dBμV or less. Around this limit of 30 dBμV, errors sometimes occur due to quality deterioration of the call. If both CS selection and received signal strength are appropriate, the FER is finally evaluated. If the FER increased (approximately 24 or higher), the DECT or other wireless systems could be interfering with the DCL.

By investigating a case according to the procedure in the manual using the terminal-maintenance mode, the cause of the reported problem could be determined and the primary countermeasure could be applied on site.

4. Example case of problem

4.1 Details of problem

A customer using a DCL of the αA1 key telephone system reported that "calls in the store are breaking

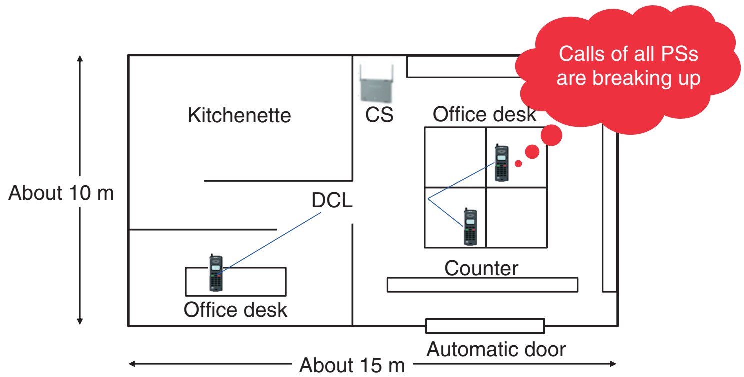


Fig. 3. Configuration of customer's equipment.

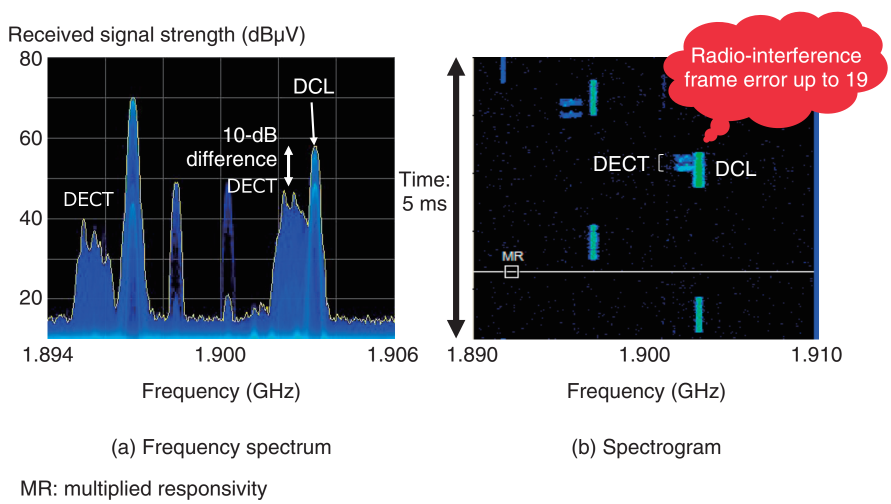


Fig. 4. Results of measuring radio environment.

up." The customer's equipment is illustrated in **Fig. 3**. A CS is installed near the center of the first floor in the store.

4.2 Investigation and results

The PSs were set to the terminal-maintenance mode on the first floor of the store, and the investigation was conducted according to the flow chart shown in Fig. 2. In this case, the handover was not executed since there was only one CS. The investigation was therefore started by checking the received signal strength. The strength in the customer's store was 40 dBμV or higher, so there was no issue with that. Then the FER was checked, revealing that the FER increased in a specific place in the store.

The results of this investigation following the manual indicate that radio waves of other wireless systems may have been interfering with the DCL radio waves. However, no equipment using other wireless systems was found in the store. To evaluate the electromagnetic environment in more detail, the radio-wave spectrum was measured using a spectrum analyzer where the problem occurred. The results of that measurement are illustrated in **Fig. 4**. The frequency spectrum and spectrogram (i.e., time change of the spectrum) when a call was interrupted are shown in Figs. 4(a) and (b), respectively. From Fig. 4(a), we can see the DCL signals exist in the frequency band of the DECT. Similarly, the spectrogram revealed that the DCL and DECT signals overlap. These results indicate that the reported problem was due to interference between the DCL and DECT radio waves.

4.3 Estimation of cause

As a result of investigating the electromagnetic environment including the area outside the customer's store, we found that DECT terminals were being used in an adjacent store and the radio waves from those terminals were interfering with other terminals in the customer's store. The DECT can usually prevent interference with DCL radio waves by detecting the usage status of radio waves with carrier sensing and switching channels depending on the results. However, in the case of this customer's environment, interference prevention would not work well because the terminals in the two stores were far apart.

4.4 Countermeasure

We confirmed that wireless signals including those from the adjacent store are transmitted by less than 12 slots in the customer's area and replaced the customer's PSs, i.e., A1-DCL-PS terminals (DCL), with A1-DECT-PSSET terminals (DECT), which can be adapted for up to 12 slots (12 terminals) in the same frequency band.

5. Current status of DCL problems

The number of technical consultations concerning the DCL using the 1.9-GHz band is gradually increasing; we had 159 cases in 2018. The main reasons of the consultations are "the call is breaking up," "cannot make a call," etc., and analysis of our on-site investigation revealed that most causes of those problems are either interference with the DECT or handover. Moreover, some customers using the DCL often report one-second interruption caused by normal handover operation as a problem. Thus, it is important to explain that handover will occur with some interruption to customers before they start using the DCL.

When a problem occurs, using the terminal-maintenance mode of PSs helps determine the problem because service personnel can determine the operating status of the PSs and radio-wave environment. Based on such information, measures that optimize the wireless environment, such as adjusting the installation position of CSs, increasing or decreasing the number of installed CSs, and adjusting their signal power can be implemented.

The "DCL Troubleshooting Manual" is available for download at CyberTasc (URL: https://www.cybertasc.com/), a site for disseminating information to telecommunication construction companies (limited access and written in Japanese).

6. Conclusion

This article introduced the content of the troubleshooting manual concerning the DCL. TASC provides this manual to improve the efficiency of such troubleshooting in the field. The Network Interface Engineering Group in TASC works on solving various problems related to the network protocol, interface, signals, and so on, toward early resolution of problems on site. TASC will continue to actively engage in technical cooperation and development and disseminate its technologies through activities such as technology seminars.

References

[1] Association of Radio Industries and Businesses, "Second-generation Cordless Telephone System (7.1 Ver.)," ARIB STD-28, 2018.

[2] Association of Radio Industries and Businesses, "Radio Equipment for Radio Stations of Time-division Multiple Access Wideband Digital Cordless Telephones (2.0 Ver.)," ARIB STD-T101, 2018.

[3] H. Ito, "Actual EMC Problems in Wireless Systems," Electro Magnetic Compatibility (EMC), Vol. 379, pp. 33–42, Nov. 2019, Kagakujyoho Shuppan Co., Ltd (in Japanese).

External Awards

ELEX Best Paper Award
Winners: Takuya Hoshi, Norihide Kashio, Yuta Shiratori, Kenji Kurishima, Minoru Ida, and Hideaki Matsuzaki, NTT Device Technology Laboratories
Date: September 15, 2020
Organization: The Institute of Electronics, Information and Communication Engineers (IEICE) Electronics Society

For "InGaP/GaAsSb/InGaAsSb Double Heterojunction Bipolar Transistors with 703-GHz f_{max} and 5.4-V Breakdown Voltage."
Published as: T. Hoshi, N. Kashio, Y. Shiratori, K. Kurishima, M. Ida, and H. Matsuzaki, "InGaP/GaAsSb/InGaAsSb Double Heterojunction Bipolar Transistors with 703-GHz f_{max} and 5.4-V Breakdown Voltage," IEICE Electronics Express, Vol. 16, No. 3, p. 20181125, 2019.

Communications Society: Distinguished Contributions Award
Winner: Toshihito Fujiwara, NTT Access Network Service Systems Laboratories
Date: September 16, 2020
Organization: IEICE

For his contribution to the operation of Technical Committees in Communications Society.

Electrical Science and Engineering Award
Winners: Hideki Maeda and Hiroki Kawahara, NTT Network Service Systems Laboratories; Asuka Matsushita, NTT Electronics
Date: September 28, 2020
Organization: The Promotion Foundation for Electrical Science and Engineering

For the development and deployment of Beyond-100G optical transport system.

Best Paper Award
Winners: Keisuke Tsunoda, Naoki Arai, and Kazuaki Obana, NTT Smart Data Science Center
Date: October 7, 2020
Organization: Information Processing Society of Japan (IPSJ)

For "Estimating Number and Dwell Time from CO_2 Concentration Based on Partial Modelling with Variable Time Window."
Published as: K. Tsunoda, N. Arai, and K. Obana, "Estimating Number and Dwell Time from CO_2 Concentration Based on Partial Modelling with Variable Time Window," Multimedia, Distributed, Cooperative, and Mobile Symposium (DICOMO) 2020, 2F-1, June 2020.

Papers Published in Technical Journals and Conference Proceedings

Relationship between Scenery Structure and Illusion Strength on the Basis of the Three-dimensional Interpretation of the Café Wall Illusion

K. Maruya and T. Ohtani

Journal of Graphic Science of Japan, Vol. 54, No. 2, pp. 13–18, September 2020.

When the shadows of a group of linear columns projected on a stepped three-dimensional (3D) plane are diagonally viewed, the retinal image shows a geometric pattern that is similar to the Café wall illusion, which is a popular geometric illusion. In this study, we investigated the relationship between the geometric structure of a scene, such as the 3D shape, viewing angle, and direction of the light source, and the strength of the Café wall illusion perceived from the retinal image of those 3D situations. The difference in the depth at which the shadows fall should be approximately 1–2% of the observation distance to result in a strong Café wall illusion. The results of our investigation indicate that the information of the shadow that falls on an object is an effective clue for estimating the shape of the object surface with relatively small depth differences in the case of large viewing distances.