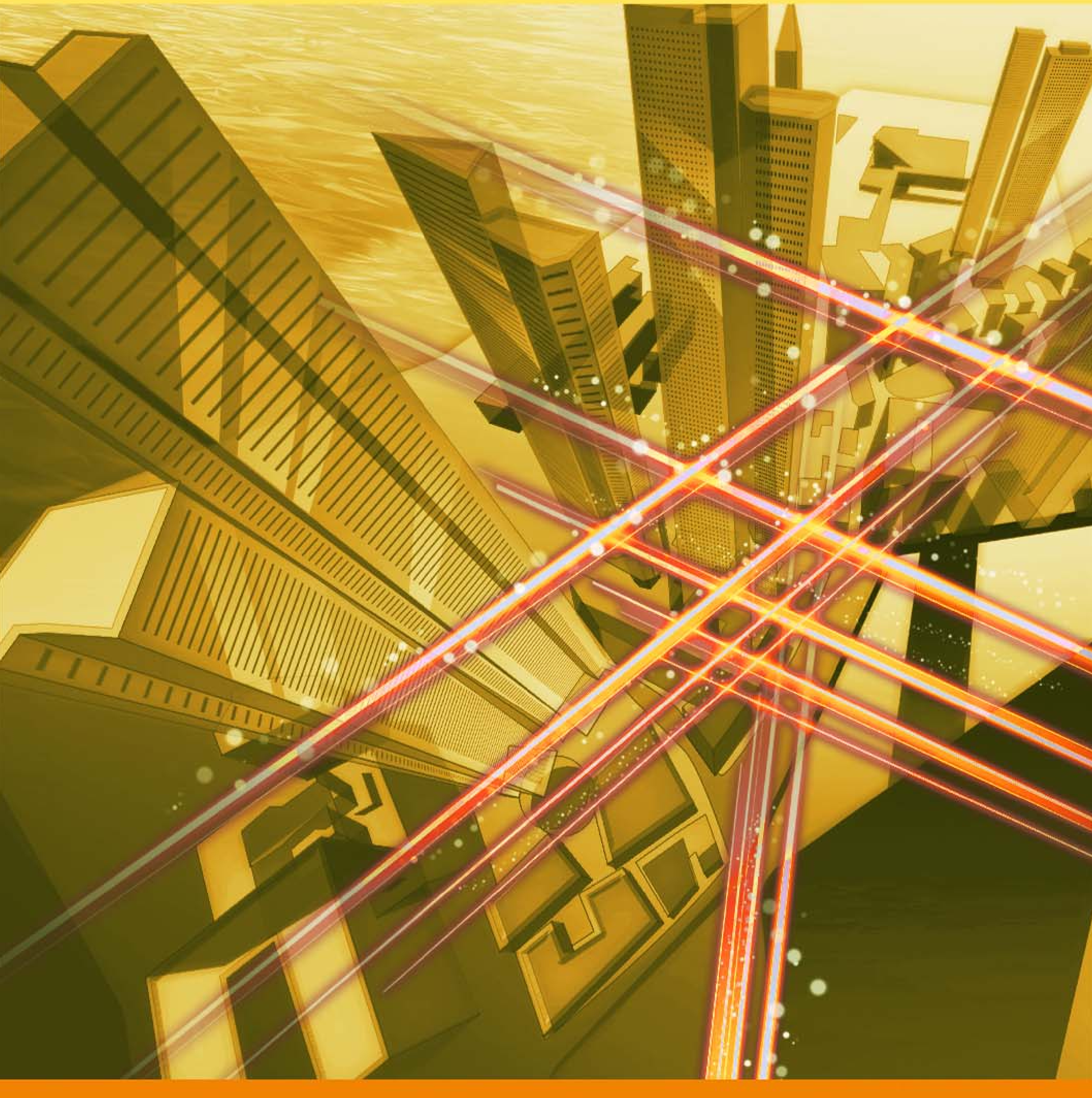
NTT Technical Review

10
2017



October 2017 Vol. 15 No. 10

NTT Technical Review

October 2017 Vol. 15 No. 10

External Awards/Papers Published in Technical Journals and Conference Proceedings

- External Awards/Papers Published in Technical Journals and Conference Proceedings

Overview of NTT DATA Technology Foresight 2017

Yuji Nomura

Abstract

NTT DATA Technology Foresight is the outlook on technology trends of the near future that is compiled by NTT DATA once a year. The objective is to find the challenges our future society will face at an early stage and serve as a compass to promote the creation of new value. The Feature Articles in this issue introduce NTT DATA Technology Foresight 2017 published in January this year; this article provides an overview as an introduction to it.

Keywords: technology, foresight, trend

1. Introduction

NTT DATA Technology Foresight is the vision of technology trends predicted to occur in the coming three to ten years. It is compiled by NTT DATA once a year. This vision notes the challenges our future society will face at an early stage and serves as a compass to promote the creation of new value.

NTT DATA aims for the betterment of society by depicting a future vision and achieving it together with various customers by foreseeing the impacts future technology will have on society and business.

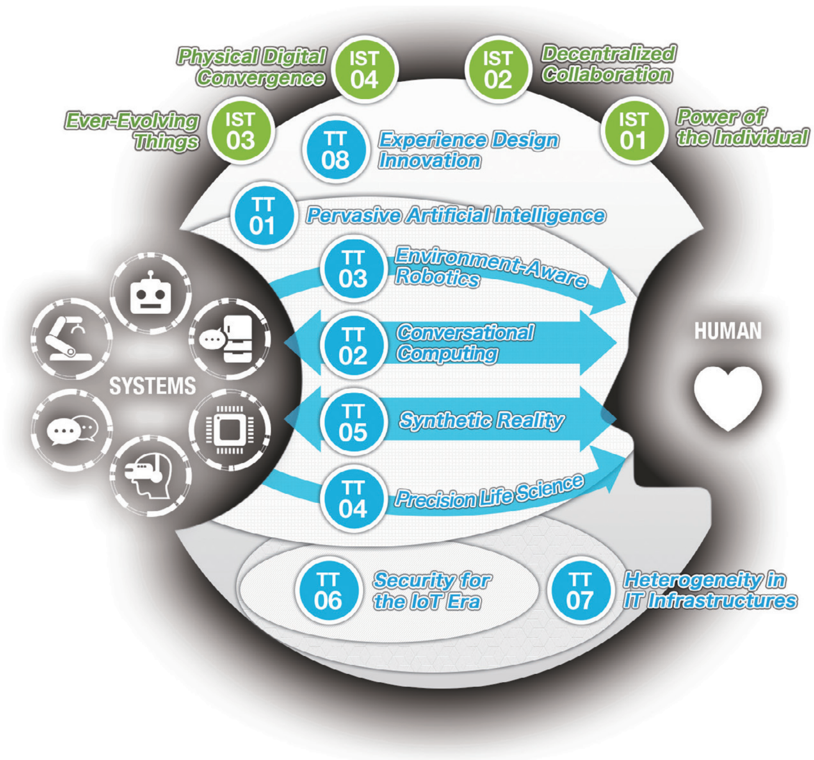
At NTT DATA, we incorporate NTT DATA Technology Foresight into our management strategy, and we are committed to technology development and service creation that anticipate changes in the business environment [1].

2. Information society trends

NTT DATA Technology Foresight 2017 has identified four information society trends (ISTs), as follows (**Fig. 1**). The digital world has become a material part of our daily lives, and we are heading toward a society in which increased value will be provided via the Internet. In addition to the growing influence of individuals (IST01) [2], open participative collaborations (IST02) [3] are revolutionizing workplaces and societies. As information technology (IT) advances, restrictions in the current society will be overcome (IST03) [4], and the structure of society will be transformed (IST04) [5].

3. Technology trends

NTT DATA Technology Foresight 2017 has identified eight technology trends (TTs). First, artificial intelligence (AI) is identified as a central technology that has become increasingly important (TT01) [6]. The next three trends are topics related to AI that are covered within the research: conversational computing (TT02) [7], robotics (TT03) [8], and precision life sciences (TT04) [9]. Additionally, synthetic reality, a technology that is transforming the interface between humans and systems, is discussed (TT05) [10]. Also, cybersecurity for the Internet of Things era (TT06) [11] and the advancement of IT infrastructures (TT07) [12] are featured as important trends. Lastly, experience design innovation has also been added as a technology trend (TT08) [13]. These trends are explained in more detail in the Feature Articles in this issue.



IST: information society trend
TT: technology trend

Fig. 1. NTT DATA Technology Foresight 2017.

References

[1] Website of NTT DATA Technology Foresight, http://www.nttdata.com/global/en/insights/foresight/index.html

[2] Y. Nomura, "Power of the Individual," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa2.html

[3] Y. Nomura, "Decentralized Collaboration," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa3.html

[4] Y. Nomura, "Ever-evolving Things," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa4.html

[5] Y. Nomura, "Physical Digital Convergence," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa5.html

[6] Y. Nomura, "Pervasive Artificial Intelligence," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa6.html

[7] Y. Nomura, "Conversational Computing," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa7.html

[8] Y. Nomura, "Environment-aware Robotics," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa8.html

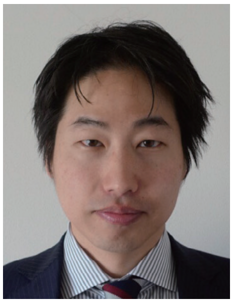
[9] Y. Nomura, "Precision Life Science," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa9.html

[10] R. Oi, "Synthetic Reality," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa10.html

[11] R. Oi, "Security for the IoT Era," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa11.html

[12] R. Oi, "Heterogeneity in IT Infrastructures," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa12.html

[13] R. Oi, "Experience Design Innovation," NTT Technical Review, Vol. 15, No. 10, 2017. https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201710fa13.html



Yuji Nomura
Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.
He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Power of the Individual

Yuji Nomura

Abstract

Our individual-centric society is leading to the transformation of current systems. Individuals are encouraging the restructuring of businesses, an increase in options, and a transformation to a more flexible society.

Keywords: individualization, on-demand economy, gig economy

1. Narrowing the information gap

The spread of the Internet has narrowed the formerly unbridgeable gap in the ability to distribute information between enterprises and consumers, employers and employees, and large businesses and small-to-medium-sized businesses. The diffusion of social media has also increased the individual's ability to spread information, which is placing considerable pressure on providers. In the manufacturing arena, lower prices made possible by mass production and uniform service can no longer meet diversifying consumer needs. This is causing a shift from mass production to high-mix, low-volume production and even to the personalization of products in accordance with consumer requests. In addition, significantly lower costs to switch providers and to switch brands of products and services have strengthened the relative power of consumers. A recent trend is consumerization, where innovation occurs in consumer markets ahead of business markets, after which the innovation is imported to front-line businesses.

Narrowing the gap between citizens' ability to access information and the government's ability to do so has resulted in significant changes in politics, the economy, and social life as a whole. Vox populi—the voice of the people—which used to be oppressed, now enjoys much greater exposure, and it is common to hear sharp individual opinions, with demands regularly made to incorporate consumer views into policies. Since a reduction in the national sense of belonging and a rise of populism are also being seen, in some cases, unexpected results are occurring in elections and referendums. Finding an equilibrium point is expected to take a little more time.

2. Development of individualization

Personalization is shifting toward individualization. Stronger advocacy of individual demands and the recent ability to acquire accurate information on individuals and their behaviors have made it clear that services that segment customers simply by their attributes, for example, age and gender, and their purchase histories, do not attain maximum results. For example, analysis of information distribution via social media, information on families and assets, and real-time analysis of location and individual situations and behaviors, are now being combined to increase closing rates and likability rates. One case reported that individualized rewards programs based on personal tastes and interests would result in more than ten thousand times as many variations in rewards programs, doubling the rate of utilization.

To achieve customer satisfaction, it is imperative that providers of products and services restructure their business models. In the medical field, treatment methods that incorporate genetic information, patient personality, and living environment in addition to the state of illness and comorbidities, will likely become more commonplace. Although concerns about privacy may make the application of these methods optional, evidence of the significant impact on users will likely promote continued adoption of these techniques.

Three-dimensional (3D) printing is a system that supports individual needs more accurately and rapidly than conventional molding methods and is therefore expected to expand rapidly in the future for prototypes, finished products, and parts. The spread of 3D printing will lead to the shift of manufacturing locations closer to the final consumption points, enabling individualization even of manufacturing processes. This may result in the restructuring of supply chains as a whole. A recent IDC study forecasts that the 3D printing market will grow at the compound annual growth rate of 22.3%, with a market size of approximately 29 billion dollars in 2020, almost 2.2 times the size of the market in 2016 [1]. Expansion of 3D printing is expected particularly in the dental and medical implant treatment fields.

3. Rise of new businesses

Digitization has caused rapid changes in the business models of existing companies. For example, the way to purchase music has evolved from record albums and compact discs to individual pieces of music bought online. In the publishing industry, consumers can now subscribe to newspapers on an article basis, radically shifting distribution systems. The service industry has seen the appearance of FinTech (financial technology) companies, which specialize in their areas of strength to provide functions to existing financial institutions. This indicates a change from the provision of uniform service to the provision of service restricted to certain functions. Some financial institutions offer only digital services without providing any branch offices, while other banks have expanded their functions to meet customer demands. In place of large companies offering uniform services, the distribution arena is experiencing unbundling, where a selective combination of multiple startups can provide similar services. An increasing number of construction and healthcare companies are also providing highly specialized services by narrowing down their offerings to certain functions. Startups have the advantage of greater proximity to customers.

Digitization has led to standardization and lower transaction costs, reducing the advantage of one company providing all functions. Large organizations are also hindered by bureaucratic structures and slow decision making, resulting in the need for a system to create and adopt innovations. For this reason, collaboration between large companies and startups has become more common. In addition, acceleration of the information technology (IT) revolution is blurring the boundaries between industries. For example, some IT-related startups are entering the automobile and space industries.

4. Expansion of on-demand economy

The on-demand economy, in which products and services are provided based on demand, continues to grow. In the United States, over 40% of adults have used this type of service. The spread of smartphones and social media has enabled more effective matching of supply and demand, leading to more efficient use of idle assets.

Because consumers can now become providers, an increasing number of markets are in flux. While services based on sharing of personal cars and lodging facilities are generating resentment from the taxi and hotel industries, users who are unsatisfied with the inflexible service of the existing industries support new entrants not subject to regulations, causing disruption to the existing business model. This trend has even had an impact on the automobile industry, which has entered the business of sharing vehicles, despite the potential negative effect on sales volumes. In tandem with the expansion in the types of services offered, for example, housekeeping services and home delivery of food, the number of users is also increasing. The majority of users are currently young adults who are trend-conscious and knowledgeable in IT, but the market for this type of business is expected to expand in the future.

5. Penetration of the gig economy

An increasing number of workers, especially in advanced countries, are freelancing instead of working for a specific company or organization, thus propelling the diversification of individual work styles. This style of work with freelancers and short-term contracts is referred to as the gig economy. In addition to digitization and mobilization, which has separated work from the workplace, expansion of the on-demand economy and crowdsourcing has paved the way for utilization of the abilities and resources of individuals. This is one of the key factors causing the increase in the number of full-time and part-time freelancers. In particular, more individuals with highly specialized skills have access to global markets and select a job that takes advantage of their abilities regardless of the location.

The United States has approximately 53 million freelancers, approximately one-third of the workforce. This number is expected to increase to more than 50% of the workforce by 2020 [2]. Similar trends in the number of freelancers can be seen in Europe and Japan, with large companies increasingly hiring freelancers. Employers that leverage freelancers to strategically use their specialized expertise instead of to simply reduce costs or adjust the number of employees experience enhanced competitiveness. If companies and organizations continue to utilize external skilled workers for important tasks, organizations will become more open and flexible.

6. Experiencing the power of the individual

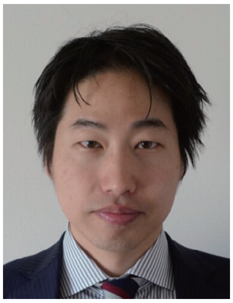
The power of the individual is already fundamental in the world, bringing changes to the relationships among entities in society. Although a negative aspect exists where fake online news is affecting even politics, the impact of the individual will likely continue to expand, encouraging modifications in industries, reexamination of regulations, and construction of new systems. The power relationship (power balance) in society will also likely continue to change. As a result, companies need to look beyond these changes and adequately prepare for their next steps in the future.

References

[1] IDC Press Release, "Worldwide Spending on 3D Printing Forecast to Grow at a Compound Annual Rate of 22.3% to Nearly $29 Billion in 2020, According to IDC's 3D Printing Spending Guide," Jan. 2017. https://www.idc.com/getdoc.jsp?containerId=prUS42211417

[2] A. J. Nielsen, "The Freelancing Business Is Booming in Europe and America," Freelancer Worldwide A/S, Mar. 2016. https://freelancerworldwide.com/the-freelancing-business-is-booming-in-europe-and-america/?lang=en

Trademark notes



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Decentralized Collaboration

Yuji Nomura

Abstract

Many people and all kinds of things will be linked to the Internet in the near future, resulting in innovation. Each component will act autonomously, and a new ecosystem will be built where relationships will change dynamically.

Keywords: Internet of Things, blockchain, API economy

1. Peer-to-peer society

The Internet, for which no centralized management mechanism exists, has changed the information distribution system via peer-to-peer networks, where users are interconnected on an equal basis. Organizational structures are being flattened and decentralized, and open systems are spreading throughout society at large.

2. Building a collaborative economy

All types of information are circulating on the Internet, which is used by more than 3.4 billion people in the world. Large amounts of knowledge and know-how are contained within online encyclopedias, which let anyone edit topics at will. In this collective intelligence approach, individual authors contribute knowledge from their specialty areas. Because inaccurate information is corrected over time, encyclopedias created in this way are comparable in accuracy to those edited by experts. Some believe that these encyclopedias offer detailed information that ordinary encyclopedias do not address.

Other examples for similar uses of collective intelligence include investigating the causes of illnesses and developing therapeutic methods based on various kinds of information provided by the patients, and participatory sensing where individuals' smartphones are used for monitoring the environment and acquiring data on traffic conditions. While ordinary sensors are more precise than participatory sensors, they are expensive, and thus, measurement locations are limited. Although sensors such as those in smartphones may be lower in precision, their ability to collect huge amounts of data over wide areas results in a low rate of overall errors. Participatory sensing is also used for purposes such as map information updates and urban planning.

To encourage open innovation, an increasing number of companies are utilizing crowdsourcing, which is the practice of soliciting the knowledge, know-how, services, and work from the general public. Its uses have expanded to include the development of new products and services, generation of problem-solving ideas, and the application of specialized expertise and skills. Professionals also use crowdsourcing to leverage their skills and expertise by identifying opportunities.

Crowdfunding is the practice of soliciting funds from the general public, enabling innovation by opening up new opportunities of funding startups. It diversifies financial risks and introduces a new decision-making mechanism for financing that enables even companies without collateral to raise funds if they win support for innovative ideas. A growing number of large companies as well as small businesses use crowdfunding for development of new products because they can derive benefits such as the improvement of ideas during the development stage and prediction of reactions via communication with the market.

3. Growth of the Internet of Things

The Internet of Things (IoT) has always existed as a concept, but not until recently has it started to become a reality. The addition of a sensor and a communication feature makes it possible to convert various things to digital devices, including smartphones, information devices, and household appliances. This enables the devices to send data on any conditions or changes in the devices themselves or in the environments and for people to remotely monitor, manage, and control the devices. It has been predicted that by 2020, as many as 20 to 30 billion devices will have access to the Internet [1–3] and will be able to function in coordination with one another.

Autonomous problem avoidance and condition optimization delivered by artificial intelligence technology and big data analysis are also anticipated to become commonplace. Some examples include control of the demand for electricity by balancing and prioritizing the usage of home appliances, and collision avoidance of self-driving cars and drones by mutual communication.

In manufacturing, IoT brings an industry revolution called Industrie 4.0, or the Industrial Internet. Industrial Internet sensors are installed on parts and manufacturing machinery and linked to people (laborers), production plans, and processes to optimize the entire lifecycle of a product, including procurement, production, and post-shipment. It is now possible to change a production plan on a real-time basis in accordance with changes in the market, to replace a production process in response to a mechanical failure, and to improve operational efficiency by monitoring product conditions. This is giving rise to new business models aimed at increased profitability by optimizing machine operation. The entire supply chain including distribution can be coordinated, and plans are underway to detect a delay in the arrival of specific parts ahead of time in order to optimize the production plan that covers multiple factories.

4. Emergence of blockchain technology*1

Virtual currencies do not have an issuer such as a national government or central bank. Instead, reliability is ensured through the use of decentralized, distributed ledgers, and an ‘Internet of money’ is formed that circulates the values of such currencies. Not only can virtual currencies be used as a means of settlement in lieu of real currencies, they can also be used for managing and transferring rights and contracts in coordination with tangible and intangible assets. Because blockchain technology is highly transparent and immutable, it is gradually becoming used in the registration of real estate, distribution of copyrights, and management of medical information. Decentralized e-commerce and market forecasting, where there is no central management entity, are also starting to emerge. Usage in voting and notary services is also being sought.

A decentralized autonomous organization (DAO), where no centralized governing system exists, is operated autonomously according to predetermined rules. One characteristic of transactions in a blockchain is that they cannot be altered or canceled once they are recorded, and this may become a problematic issue. Nevertheless, DAOs can potentially be used for shared services, investment funds, and asset management. For example, Estonia has implemented a notary service for marriages, births, and contracts using e-Residency, which is used by foreign nationals. Identity document issuance, notary services, and other services offered by a virtual nation also have the potential for widespread use in the future as a means of verification without depending on a national government.

5. Rise of API economy*2

Services are increasingly integrated by virtue of inter-company coordination. Sharing information and systems leads to coordination that involves a wide

*1 Blockchain technology: The core technology for virtual currencies such as Bitcoin.

*2 API economy: A trend for a company to disclose (or create a platform for) its API (application programming interface)—a mechanism to call a software function—thereby making the coordination of information systems with other companies easier and creating new value and business.

variety of concerned parties, including those in different industries. This in turn leads to the creation of highly convenient new services and innovations. Examples include the coordination of flight information with airport transportation services and hotel reception, utilization of information obtained from automobiles for proposing insurance and maintenance services, and the process of searching for a house to purchase on a smartphone and applying for a home loan. Because services are automatically proposed based on individual situations, users do not have to spontaneously find services by themselves. This will lead to easy decision making for purchases. Contextual commerce integrates social media and payment services, enabling an on-the-spot purchase of a product a user wants. This type of commerce is changing consumers' purchasing behaviors.

6. Migration to a decentralized society

Although only some progressive organizations are currently practicing a decentralized approach, it is anticipated to spread widely as a system of value exchange that does not require an intermediary. Some systems and tasks that currently have centralized management are considered to be the best initial candidates for decentralized systems. These include those that do not have central managers, for example, systems for international payments and remittances, and the functions and tasks of government agencies suitable for outsourcing to the private sector. However, many tasks are not suitable for a decentralized system, including transactions of listed stocks, where situations may change instantaneously, and areas that require substantial decision making.

Government shutdowns of the Internet, including partial shutdowns, occurred as many as 56 times in 18 nations in 2016 [4]. In a truly open, decentralized society, even the government would not be able to shut down such services. An open, decentralized society is one with high transparency and no conflicts. At the same time, it is a society with no leader. The speed and spread of the migration to a decentralized society will likely depend on future system planning and the degree of societal acceptance.

References

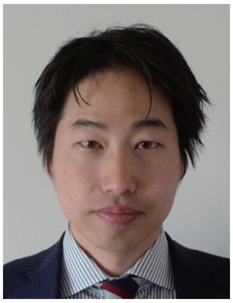
[1] IDC, "Worldwide and Regional Internet of Things (IoT) 2014–2020 Forecast: A Virtuous Circle of Proven Value and Demand," May 2014.

[2] Ericsson, "Ericsson Mobility Report," Nov. 2016.

[3] IHS Technology, "IoT Platforms: Enabling the Internet of Things," Mar. 2016.

[4] Website of accessnow, https://www.accessnow.org/keepiton/

Trademark notes



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Ever-evolving Things

Yuji Nomura

Abstract

Big data analytics will fuel innovation. Even after products are shipped, they will become ever-evolving things through functionality and performance enhancements. This in turn will boost value for customers and promote business model transformation.

Keywords: smart machines, artificial intelligence, cybersecurity

1. From goods to services

With the development of globalization and the expansion of digitization, more products are becoming commodified. In the meantime, the transition from an industrial to an intellectual society has changed the sources of value from tangible things and assets to the use of intangible information, design, and functions. The same product may have different value depending on the user's sense of value, situation, and usage. In some cases, products have emerged from price-cutting wars with added value, ending up with an increased price point. The law of *one* price, which indicates that value is attributed to products, is being shifted to *multiple* prices, which means value is derived from products by using them.

In the manufacturing industry, *servitization* that provides products as a solution is becoming more common. This accompanies the paradigm shift from the idea that value exists in products themselves to the idea that value is generated by using products. For example, customers who wish to shop efficiently and those who wish to enjoy their shopping experience look for different kinds of value such as the atmosphere of the store and the service of the sales staff. Redefinition of the value offered to customers is also happening in the financial, medical, and social welfare fields.

2. Analysis of big data

The development of information technology (IT) has enabled all kinds of information produced by people and things to be accumulated. Real-time analysis of varied and large amounts of information generated in real time helps visualize the conditions of customers, markets, society, and the environment. Signs of change and correlations that have been elusive can now be linked to improvements in customer satisfaction, development of new products, diagnosis of illnesses, and development of medicines. While privacy and other issues must be considered, the results of big data analysis may prevent problems based on the correlations of behavior patterns and environmental conditions with success and failure, illness and accident, and criminal action.

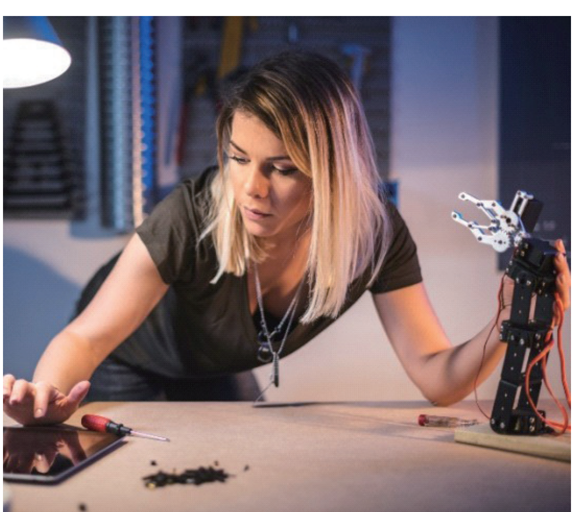
Analytical algorithms are the key to deriving value from big data. For example, analysis of customer attributes may generate different results depending on the data and analytical algorithms used. In the consumer finance industry, traditional financial information has included revenues, debt balances, and credit histories. In addition to these, an approach is emerging that determines the credit risks of individuals based on their overall information, including financial behaviors that were traditionally overlooked and information seemingly unrelated to financial behaviors such as the way they sign documents, their major in college, and their postings on social media. While demographic groups with no credit histories can now receive loans, which leads to the correction of some disparities, loans to applicants with high default risks have decreased, allowing some lenders to reduce losses from defaults by more than 20%.

3. Ever-evolving things

Because automobiles, machines, consumer

electronics, and other things are connected to the Internet, processes common in the world of IT are expected to be used widely in the physical world. For example, when software updates of computers and digital devices are applied to a wider range of things, users may be able to enjoy the advantages of added features and improved performance without trading in their things. While in the past, goods were replaced whenever a new product with enhanced features appeared, it may become common to update only the software while continuing to use the same device (hardware). If that happens, design concepts will change so that software will provide all kinds of features. Modularization, which enables only the components of a device to be replaced depending on the software features, and agile development methods, which are common in software development, might be introduced.

Cars that can be converted to self-driving vehicles by installing software, and robots whose functions expand by adding applications have already been introduced. Robots that learn tasks and those equipped with artificial intelligence (AI) with self-learning abilities also exist. It used to be that the different tangibility of goods and services produced other differences. For example, there was a time difference between the production and consumption of goods, while those of services were simultaneous. The value of goods was determined at the completion of production and deducted through consumption, while that of services was co-created by producers and consumers. Things whose value increases as they are used have made it meaningless to make a distinction between goods and services. This could create an impact on industrial classification and how depreciation is treated in accounting.



4. Emergence of smart machines

Recent years have seen the emergence of self-driving cars, drones, robots, and other machines equipped with AI (collectively called smart machines) that have self-learning functions and move autonomously. The spread of smart machines is expected to change the roles and functions of both people and machines in society. For example, more factories and other facilities may become unmanned, with robots monitoring machines. Self-driving cars and drones will be handling logistics, and virtual assistants will be serving customers at banks and stores. Robots can already communicate with other robots to share what each has learned on its own. In the future, it is believed that smart machines will also be sharing perceptions such as those concerning collision avoidance, acting in collaboration with one another. While society is human-centered, smart machines are expected to assume a central role in the social infrastructure and social control, giving rise to restructured social systems and processes.

5. Human–AI collaboration

Although AI now has the capability to beat top-class professionals in Japanese chess (*shogi*) and *go* through self-learning, this type of AI is specialized in specific skills. There is still no AI that possesses the overall high intelligence of humans. In light of the increasing dependence on decisions made by AI in the future, it may make sense to provide AI with a basic education that lets it make correct decisions similar to human children, allowing for versatility and high intelligence. Unfortunately, AI could also be taught impropriety by malicious developers, enabling it to make incorrect decisions or misjudge between right and wrong. With its increased use, AI will be expected to make decisions outside its expertise such as those related to ethical problems. In short, it will be essential for AI to acquire common sense.

Some non-profit organizations have already launched educational institutions that focus on AI. In the future, the education of AI may become critical. Meanwhile, if a smart machine (AI) causes an accident due to its self-learned ability, an issue will occur as to the extent of liability that may fall on the owner, user, manufacturer, or software developer. Some have started to consider giving smart machines the legal status of a person, subjecting it to liability.

One study reported that the employment of 40% to 50% of a nation's labor force could be replaced by AI and robots [1]. In school education and worker training, it may be necessary to have humans learn a high level of cooperativeness and creativity, which are difficult to substitute with AI. However, the employment opportunities replaced by AI could exceed those newly created for humans. For this reason, it may be necessary to first examine the social system, including the social security concept, the possibility of shorter workdays to increase the number of employees, and the implementation of basic incomes.

6. Strengthening of cybersecurity

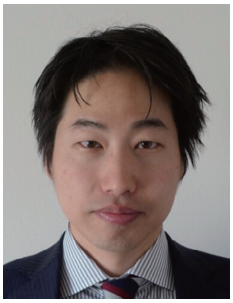
It is believed that real-time situational judgment and continuous communication among smart machines will help avoid accidents in the future. This could make the quality and reliability of communication functions even more critical. For example, if exploitation of security vulnerabilities causes a large-scale blackout or traffic jam, the entire society might become dysfunctional. Hacked smart machines might even start assaulting humans.

One way to avoid hacking is to refrain from having them be constantly connected to the Internet. However, this would make it impossible to take full advantage of existing information. Although strengthening countermeasures against cyberattacks is necessary to ensure security, it will also be necessary to build fail-safe systems to minimize the impact of potential failures and to prevent unexpected actions of a smart machine.

Reference

[1] Nomura Research Institute, "Coexistence with Artificial Intelligence—Interview with Michael A. Osborne by Masaki Fukui—," lakyara, Vol. 234, Mar. 2016.
https://www.nri.com/~/media/PDF/global/opinion/lakyara/2016/lkr2016234.pdf

Trademark notes



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Physical Digital Convergence

Yuji Nomura

Abstract

The physical-digital convergence will broaden in scope. People will freely go back and forth between physical (real) and digital environments, and between offline and online without being aware of the borderlines. Restrictions in time, space, and ability will be relaxed; therefore, new value will be created.

Keywords: smart cities, self-driving cars, digital business

1. Transition to smart cities

The United Nations estimates that the world population will be 11.2 billion in 2100 (medium variant projection), approximately 1.5 times the population of 7.3 billion in 2015 [1]. Although some advanced countries can expect an increased population via the influx of immigrants, a majority will face a dwindling population. To take efficient measures for expansion of social infrastructure and the growing demand for energy in regions where population will increase, and to avoid the risk of excessive infrastructure in regions where population will decrease, the focus on future urban planning should be shifted from physical and technological measures, which support social infrastructure through construction, to digital measures through ideas and design, which complement the infrastructure through improved utilization density and rate.

For example, traffic jams can be eliminated by using advanced traffic signal control, ridesharing, and the implementation of self-driving cars. Power shortages can be addressed by the implementation of smart grids and demand response systems. In addition, the combination of remote monitoring and preventive maintenance will likely lead to a reduction in maintenance management costs and increased longevity of aging tunnels and bridges. In smart cities, physical infrastructure and digital information will be used in combination to optimize infrastructure.

Automobile manufacturers and information technology companies are currently testing self-driving cars on the road, or will soon reach this milestone. Technological development has preceded rule formulation until now. However, the U.S. Department of Transportation has already published its own guidelines for self-driving cars, with the United Nations following to establish international standards. Although it will still take time to establish effective international rules, self-driving cars could be in practical use in the near future. Commercialization of autonomous, self-driving car technology will likely reduce individual ownership of automobiles. Buses will no longer need to run on specific routes according to a schedule, and there will no longer be a distinction between buses, taxis, shared cars, or rental cars. The transportation system itself is expected to transition to a more flexible model where vehicles are run on demand. Urban planning will also shift toward self-driving cars and efficient vehicle dispatches.

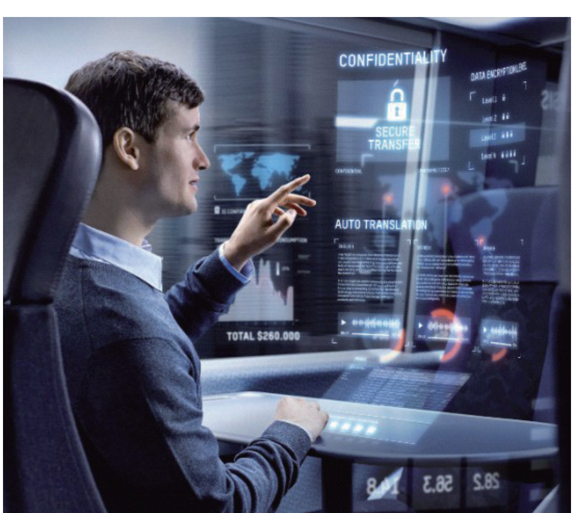
In the area of agriculture, the expansion of agricultural production is essential to support needed increases in food demand associated with increasing populations. It is anticipated that digitization will lead to a stabilized food supply. For example, sensors are already measuring temperature, humidity, and soil conditions. Similarly, drones are spraying fertilizers from the air and monitoring crop growth. In addition, unmanned farm machines are plowing land and harvesting crops. Because the leading cause of poor crop growth is weather, farms are also using mesh weather forecasting data for optimal soil management. Improvement in the accuracy of harvest prediction models also enables the pre-arrangement of storage, transport, sale, and processing of crops, with an expected benefit of reduced wastage. Digitization efforts are also beginning to take place in the fishery and livestock industries.

2. Changes in the relationship between machines and humans

The relationships between humans and machines are changing. In the past, users interacted with machines according to the requirements of the machines. Nowadays, intuitive interfaces are being introduced in machines that make interactions with them similar to interactions with other people. Humans will increasingly provide instructions to machines using postures and facial expressions (image recognition) as well as words (voice recognition). Technology making it possible to identify an individual without examining the face has also recently been developed, and the recognition rate of major voice recognition software now exceeds 90%.

Technologies have advanced to a level close to human capacity. Some predict that 50% of web searches will use images and voice by 2020. It is also expected that control of household appliances, reservations for restaurants, and orders to e-commerce stores will be achieved by talking naturally to a smart home device. (Even switches and panels will disappear from household appliances). A household communication robot will not only be able to engage in conversations with a user but also to detect emotional changes or symptoms of an illness based on voice patterns, enabling it to mitigate discomfort. A robot that can detect changes in the user's body conditions immediately and notify healthcare staff will significantly help senior citizens to lead independent lives.

Companies and their customers are starting to use a service called chatbot, which uses text (written characters) and voice to communicate naturally. Applications of chatbot are expanding from travel reservations and shopping to taxi dispatches. Currently, users need to call different chatbots that meet their needs, and communication is mostly text-based. However, in the future, it is predicted that virtual assistants will converse with users using natural language, while controlling separate chatbots.



Individual authentication will no longer require cards or passwords. Instead, touch and conversation will allow for the detection of finger and voice prints, biometrically authenticating an individual. The era of humans adjusting to machines is being transformed into an era where machines adapt to humans. Autonomous self-driving cars can be considered an example of machines adjusting to humans. With a natural interface, the digital divide, which stems from not knowing how to use machines, will probably also disappear in the future.

3. Relief from restrictions

The spread of remote healthcare and telecommuting can be regarded as evidence of the convergence of the physical and digital worlds, which has enabled services to overcome the restrictions of time and space to come to humans. Whereas simple images and audio were used to overcome distances in the past, environments artificially created by virtual reality (VR) technology and telepresence will make the experience more realistic. Augmented reality (AR) technology, where the real world and digital information are superimposed in real time, and mixed reality (MR) technology, where virtual objects are combined with the real world, are also gradually coming into use as well. Although the need for a head-mounted display was a bottleneck to the spread of VR, smartphones can now easily provide the experience. The main uses of VR are currently entertainment-related such as games. However, VR is starting to be used in the fields of education, sports training, simulations, telework, and remote support. AR is also more common in game applications and is now starting to be more widely used in areas including tourism, events, promotions, construction, and interior design. Applications focused on agriculture and fishery are also expected.

Wearable terminals, which measure and record biometric information and track physical activity, can, if worn at all times, capture health conditions. This technology may lead to the early detection of factors leading to illness and/or changes in symptoms, enhancing preventive medicine. The relationship between medical professionals and patients may change. The attitude toward healthcare may also change from people seeing a physician when they feel sick to having ongoing monitoring and communication.

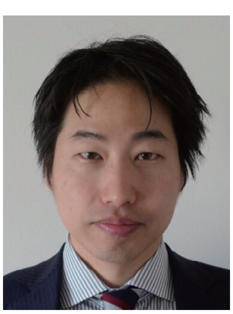
4. Transformation to digital business

The convergence of the physical and digital worlds is accelerating the transformation of companies to a digital form of business. In the past, humans made all business decisions. However, in digital business, machines, things, and other business components will also make appropriate decisions. For example, if the delivery of parts is delayed in a factory that has advanced Internet of Things technology, the machines and things in the factory may decide to switch production lines to manufacture different products. Humans would react based on this decision. There is a high probability that existing organizations and systems cannot optimize the benefits of digital business. As a result, not only the company, but the industrial structure itself will probably change.

These changes will also transform the overall social system. As gaps exist in the ability to adapt to rapid changes, new social disparities may arise. A gap between global and in-country rates of change may lead to economic imbalances. Accordingly, it may be necessary to start building a culture through education that tolerates more rapid changes.

Reference

[1] United Nations, Department of Economics and Social Affairs, Population Division, "World Population Prospects: The 2015 Revision, Key Findings and Advance Tables," 2015.



Yuji Nomura
Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.
He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Pervasive Artificial Intelligence

Yuji Nomura

Abstract

Widely accessible advanced machine learning will result in the expanded use of artificial intelligence (AI). AI will increase convenience, resolve intellectual labor shortages, and drastically advance science. Consequently, mastering AI will become a critical component of competitiveness.

Keywords: artificial intelligence, deep learning, machine learning

1. Pervasiveness of deep learning

In 2016, the big news on artificial intelligence (AI) was the major developments in the technology known as deep learning, which mimics the brain cell activity of animals. Deep learning is expected to have higher accuracy than past AI technology and feature extraction, which can also be automated, making it available to a wider range of users. In addition, open environments and open source software are resulting in accelerated advancements in deep learning, and TensorFlow, DSSTNE (Deep Scalable Sparse Tensor Network Engine), CNTK (Microsoft Cognitive Toolkit), and other frameworks and cloud environments for deep learning have been created, further enabling its use. For specific applications such as image recognition, there are now web services that let users perform deep learning and even prediction without programming.

The use of deep learning is rapidly becoming widespread. For example, farmers are using AI to recognize and learn images of cucumbers and to distinguish ratings of crops, and dermatologists are using AI to recognize and learn images of symptoms and diagnose skin cancer. In the future, the use of deep learning is expected to become even easier, and commoditization will likely follow.

2. Expanding AI applications

AI applications are expanding in various fields by leveraging deep learning. AI is already used in healthcare for diagnosis and drug discovery, in finance for stock trading and credit decisions, and in retail for marketing and customer service. In addition, AI technology supports our daily life in things we take for granted such as web search, path search, and translation activities. Thus, AI will most likely spread to all fields in a natural way.

Robotic process automation (RPA), a system that enables digital robots operating with rule engines and AI to automate white-collar tasks, is also becoming popular. RPA automates tasks without the need for programming, and it is currently used to perform rule-based automation of routine tasks. However, it is believed that advanced AI will be used to automate more sophisticated, non-routine tasks such as analysis and decision making in the future. Some predict that more than 100 million global, full-time intellectual workers will be replaced by AI by 2025 [1]. This will likely contribute to the resolution of labor shortages.

AI is also being tried in research applications. For example, AI was able to reproduce the condition of a gas that requires complex control and that is difficult to reproduce. AI used a method that would never occur to humans and completed the task in an hour [2]. Soon, AI may advance rapidly in research-based fields such as science and physics, enabling dramatic progress.

3. Further evolution of AI technology

AI is rapidly advancing in image and voice recognition, and it is still evolving in other areas. One of these areas concerns the understanding of meaning. In addition to using text to explain the contents of an image, some researchers have recently been actively trying to generate images based on text, thereby generating images that are close to the *meaning* of the text. Inter-exchangeability between language and images may mean that AI is getting close to understanding meaning.

Another area of evolution has to do with the issue that large amounts of teaching data are necessary for AI to learn. Humans have a natural ability to learn inductively based on events that occur and to identify an object after it has been taught only once. Technology called *deep reinforcement learning* is attracting attention as a potential way to mimic this ability. Reinforcement learning is the autonomous learning of subsequent actions to take or the condition that needs to exist based on experience. Deep reinforcement learning combines reinforcement learning and deep learning. With this technology, there is no need to provide in advance the teaching data necessary for machine learning. All that is required is to set up the desired actions and conditions. AI will then use repeated trial and error to learn the task. This technology is already used in AlphaGo, an AI *go* playing machine that surprised the world by beating a professional *go* player in 2016. It is also being used in self-driving cars and in robots used in factories.

Many studies are underway to make machine learning possible based on small amounts of teaching data for tasks to which reinforcement learning cannot be applied. The fundamental concept of this is called *transfer learning*, where the knowledge acquired when learning one task is applied to the learning of another, streamlining the learning process. For example, in an image categorization task with limited amounts of data, teaching just one image from a category that AI had not yet learned enabled it to achieve a level of accuracy that was almost equivalent to a case where large amounts of data were taught. If this technology is implemented, AI is expected to learn things and events at a dramatic speed, further expanding its use.

4. Importance of knowledge and technology of AI utilization

While the advancement of deep learning is remarkable, it is not versatile. Therefore, it is critical to understand the advantages and disadvantages of deep learning before identifying its application areas. Under some circumstances, rule bases, probability/search models, and other traditional algorithms may have to be selected. For example, in standard chess, humans cannot beat AI. However, in advanced chess, collaboration between AI and a human achieves better results. Therefore, it is critical to identify the areas and the best method for such collaboration.

Since technologies such as transfer learning, where small amounts of data enable learning, are under development, large amounts of learning data are still necessary to achieve a practical level of accuracy for AI in a new area. For this reason, it is necessary to determine how much learning data can be acquired in advance, and whether or not a system can be built that can perpetually accumulate data during operation based on feedback. In addition, preprocessing such as data cleansing, and the know-how on parameter tuning are important to achieve high accuracy. Machine learning has reached a point where a certain level of accuracy can be achieved. However, the knowledge and technology for utilizing AI will be a critical differentiator for companies in the future.

5. Toward AI that further contributes to humans

Deep learning is based on complex network structures similar to actual brain neurons. The high accuracy achieved with deep learning has become possible thanks to improvements in computer processing. At the same time, the process to obtain the answer has also become more complex, making it difficult for humans to understand the judgments and reason for the end result. This is one of the challenges of deep learning.

For instance, when humans look at an image of an elephant, they can recognize it as such, although they are probably not making a judgment based on logical reasoning. However, if they are asked for their reasons for deciding the image is an elephant, they site details such as its long trunk and huge ears. Thus, just as humans can explain the reason after the fact, AI may be able to provide accurate explanations for the process it uses to reach its decision.

A project is underway that aims to enable the explanation of the basis of output results [3]. If AI can acquire the ability to explain its judgment and reasoning in the future, it will be easier to improve accuracy and apply it to cases where human lives are involved, such as in self-driving cars.

Discussions are also underway on the singularity of AI's abilities to exceed those of humans. Many non-profit organizations have been established with the purpose of keeping AI from becoming private and abused, developing AI to contribute to society, evaluating AI's impact, and establishing development principles [4]. Social, ethical, and legal issues surrounding AI will need to be resolved. These efforts will increase in importance in the future. Although AI is not yet fully developed, it is important to discuss its development and effective use as well as its risks now. Doing so will contribute to the benefits of using AI for humans, such as the resolution of labor shortages and energy issues.

References

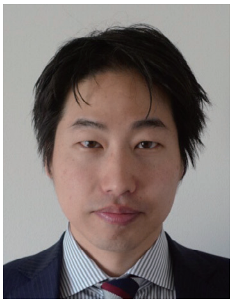
[1] McKinsey Global Institute, "Disruptive Technologies: Advances that Will Transform Life, Business, and the Global Economy," May 2013. http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/disruptive-technologies

[2] P. B. Wigley, P. J. Everitt, A. van den Hengel, J. W. Bastian, M. A. Sooriyabandara, G. D. McDonald, K. S. Hardman, C. D. Quinlivan, P. Manju, C. C. N. Kuhn, I. R. Petersen, A. N. Luiten, J. J. Hope, N. P. Robins, and M. R. Hush, "Fast Machine-learning Online Optimization of Ultra-cold-atom Experiments," Scientific Reports 6, 25890, 2016.

[3] D. Gunning, "Explainable Artificial Intelligence (XAI)," http://www.darpa.mil/program/explainable-artificial-intelligence

[4] Open AI, http://www.openai.com/
Partnership on AI, http://www.partnershiponai.org/
Ministry of Internal Affairs and Communications, AI Network Social Promotion Conference (in Japanese), http://www.soumu.go.jp/main_sosiki/kenkyu/ai_network

Trademark notes



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Conversational Computing

Yuji Nomura

Abstract

The advancement of voice recognition technology with enhanced context/emotion interpretation will make natural and seamless people-to-technology interactions possible. Such intelligently interactive systems will change human behavior, societal interactions, and decision making.

Keywords: voice recognition technology, chatbot, emotional recognition

1. Expansion of interactive interfaces

It is now customary to see people talking to their smartphones. In fact, a personal voice assistant installed in smartphones and smart devices is being used in a variety of ways from checking the day's schedule and weather, to searching for nearby restaurants. It has even become commonplace to use text messages to inquire about, order, pay, or send money via wire transfer for products. A recent survey indicates that approximately 89% of users wish to use a messaging tool for business communication [1].

In addition to smartphones, voice assistant terminals installed in residences and offices are also becoming popular in both Europe and the USA. These terminals are always ready, so when a user gives them instructions simply by talking to them, they will start music, adjust air-conditioner temperatures and room lights, or order products.

Behind the popularization of interactive interfaces is the fact that as smartphones have become popular, voice conversations with computers and the use of text chatting have taken root in our culture. The number of active global users who use messaging applications (apps) at least once a month is now over 3 billion. One of the factors for this may be that users have simply gotten tired of using ordinary apps. Users feel it is bothersome to install and learn to use each individual app. In fact, one in four users will quit using a new app after using it only once. In contrast, messaging tools use natural languages, which are familiar to humans, and users can immediately start using these tools irrespective of their information technology literacy.

The advanced processing capability made possible through interaction technology, which was enabled by the evolution of artificial intelligence (AI) in recent years, is also boosting the popularity of messaging tools. Chatbots, programs that use AI to automate communication, are becoming particularly popular. Chatbots can provide a wide range of services from flight reservations to real estate suggestions. In addition, plans are in place to use chatbots in banks to provide account balances and wire transfers and even financial plans based on usage patterns. Interactive services in coordination with AI will likely continue to increase in the future.

2. Development of conversation support technology

The advances in deep learning technology in recent years have significantly increased the accuracy of recognizing images and other patterns. In October 2016, the error rate of voice recognition technology reached 5.9%, equivalent to that of speech-to-text experts. In addition to voice recognition, lip reading with AI has achieved an accuracy of 46.8%, approximately four times as accurate as professional lip readers. This may enable users to interact with computers in situations where they are not able to speak or are in a distant location. Computers cannot only read but are also capable of varying tone and pausing speech depending on situations, producing voices close to human voices.

The use of technology to recognize emotions from voice, expressions, and text has also been spreading. For example, analysis of emotional changes in viewers enables the identification of effective scenes in commercials and differences in responses depending on countries and cultures. This technology is used widely in marketing. Call centers also use similar technology to support their customers with consideration for their emotions. Future uses of computers will expand from human assistance purposes to reading human emotions for direct interaction. For example, a computer that perceives anger in a user driving a car might talk to the user to calm him or her down.

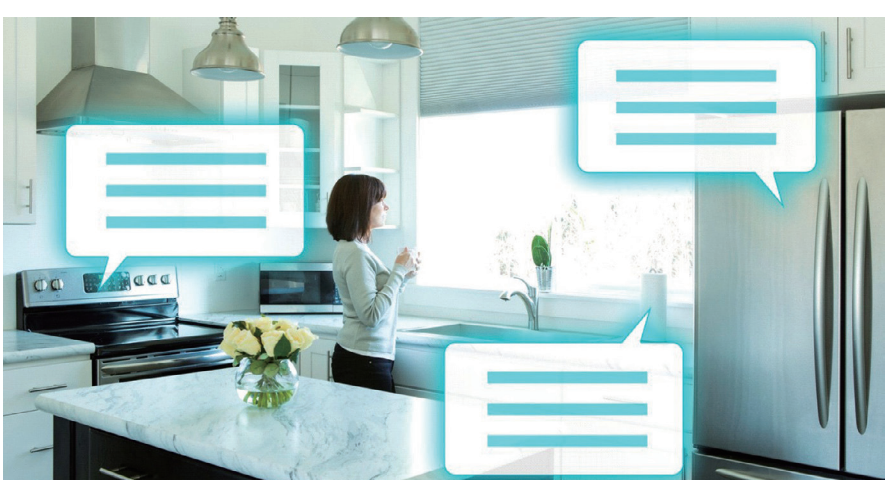
Unlike humans, computers still cannot understand contextual meanings that hide behind words. For this reason, humans may feel stressed and disappointed about the interaction, and there may come a period of disillusionment with these technologies. In the near future, however, a context-understanding technology may be launched just as technologies such as voice and emotional recognition, and voice synthesis have been developed. These technologies, which support human interaction with computers, will enable more natural interaction, as they more accurately understand the nuances of a conversation, including a user's intent and emotion. It is anticipated that this will further expand the use of interactive computing.

3. Interactive computing that innovates the world

Interactive computing in coordination with AI will result in ultimate personalization. Traditional personalization methods use past behaviors such as viewing and purchase histories. This method is often used to present to the user products already purchased and related products in which the user is not even interested. However, an interactive system can use conversations to interpret a user's intention and respond accurately even to complex requests. This system, when combined with information from sensors, can respond to a user's individual situation immediately. Personalized information will be available that defies comparison with what is used today, changing customer service, marketing, and advertising.

Conversation is the most natural communication tool used by humans. With a conversation-based interface, many future actions will be completed using interactive apps. Instead of opening a different app for each purpose, a system will likely be developed in which a personal assistant app listens to a user's requests and then distributes the tasks to other apps and chatbots based on the content. As automatic responses from AI become possible, immediate and appropriate responses will be available 24/7, which will have an enormous impact on the relationship of individuals to society. The traditional individual-to-society connection mainly consisted of one-way information by companies, such as emails and online advertisements, and a temporary connection by telephone or direct visits to brick-and-mortar stores. With interactive computing, however, the company-to-individual, and individual-to-individual connections will become bidirectional and continual.

Interactive computing even has the potential to change the means of decision making. Currently, users need to select relevant data from an overwhelming amount of information and make decisions such as product purchases and travel arrangements. In interactive computing with AI, the information required is narrowed down step by step through interaction, leading to more natural decision making. Thus, not only will the process of decision making change, but satisfying and prompt decision making will also become possible.



4. The future of interactive computing

Interactive computing goes beyond a mere convenient interface based on speech. Accumulation of interaction histories and sensor information will increase the accuracy of behavior predictions based on the situation and preferences of the user. By forecasting users' intentions ahead of time and providing the necessary information from the system, interactive computing may well become a virtual enabler of human behaviors prior to individuals becoming aware of the desire for the behavior. For example, the system may be able to provide users with the information and data they need before they realize it.

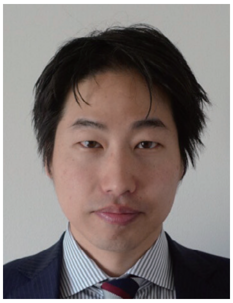
The ultimate communication may not be via language, but the immediate and accurate transfer of intentions. Studies on a brain-computer interface, which reads human intentions using brain waves to control devices, are being actively pursued. As a result, it may become possible to control computers and vehicles by merely thinking about something. The ability to read complex thoughts may eliminate the problem of not being able to express one's thoughts in words, giving rise to interactions with no gap in communication.

Interactive computing is not merely an app on a smartphone or a personal computer. It will be installed as a standard feature in all kinds of devices, while making its way to becoming a new computing infrastructure. It will not be long before we can simply talk to a device to get any information or product, at any time and from anywhere such as at home, at a store, or in a car.

Reference

[1] Twilio, "Global Mobile Messaging Consumer Report 2016." https://www.twilio.com/white-papers/global-mobile-messaging-consumer-report-2016

Trademark notes



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Environment-aware Robotics

Yuji Nomura

Abstract

Advancements in perception technology for images and voice are enabling robots to acquire enhanced environmental awareness and providing opportunities to exploit its use within products such as self-driving cars and drones. These higher-level operational capabilities will transform the industrial structure.

Keywords: environment recognition technology, self-driving car, drone

1. Development of technologies for recognizing the external world

It is safe to say that robots have better eyesight than humans do. One reason for this is that deep learning technology has increased the accuracy of image recognition. This was proven at the 2016 ImageNet Large Scale Visual Recognition Challenge (ILSVC), in which a robot identified the name of an object in an image with 97.0% accuracy, as compared to 94.9% for humans. In addition, SLAM*, a technology that simultaneously estimates a robot's own location and creates a map of its surroundings based on information from a camera and a sensor, has enabled highly accurate capture of a three-dimensional (3D) space. Yet another new technology allows the capture of a space with only a smartphone using a monocular camera, which will enable the effortless creation of indoor 3D maps. This technology will likely become widespread in many places such as commercial facilities, warehouses, and factories.

Voice recognition has also become more accurate, achieving human level and giving robots effective ears as well as eyes. Furthermore, robots can recognize things that humans cannot such as ultrasonic waves, infrared light, and magnetism. This capability is inherent to machines. It is anticipated that robots that possess the perceptual aptitude of humans in addition to other capabilities will rapidly expand their range of applications.

As the ability of spatial recognition improves, robotic contests, whose purpose is to enhance the functionality and performance of robots, are on the rise. For example, at the Amazon Picking Challenge, robots compete in their ability to place products on shelves and remove them, while at RoboCup, robots play soccer on a soccer field or compete in rescuing humans at a disaster site. In addition, at the inaugural event of a robotic car race called the DARPA Grand Challenge in 2004, none of the participants were able to reach the finish line. However, five cars completed the race in 2005, building the foundation for the self-driving technology of today.

2. Lower costs and evolution of hardware that supports robots

LIDAR (light detection and ranging, or laser imaging detection and ranging), a sensor system smaller than radio radar systems that uses light and recognizes 3D spaces, can measure extremely small particle sizes. Because it can even recognize the shape and moving speed of objects, it is receiving special attention as the potential *eyes* of self-driving cars. Although LIDAR is costly today, manufacturers are targeting a cost of 100 U.S. dollars or less in five years. Efforts are also underway to install a LIDAR sensor in a single microchip with a potential price of only 10 dollars. This type of LIDAR could be installed in many devices, and its use would quickly spread to robots and home electronics.

Concurrent with this, an effort to achieve a self-driving function without sensors such as LIDAR is underway using improved camera performance and

* SLAM: simultaneous localization and mapping

artificial intelligence (AI) to recognize objects and measure distances. Either of these evolving technologies may become the optimal choice for robots' eyes.

The growing field of biomimetics models the superior functions and structures that living organisms have attained and applies these results to technological development. Biomimetics is helping to further develop robotics. For example, robots equipped with a tactile sensor that models human pain are now able to feel discomfort upon impact and to act to avoid the impact. This will likely enable the reliable use of robots in situations where they need to work closely with humans.

3. Popularization of self-driving vehicles, drones, and task-oriented robots

A self-driving car may be considered a robot that operates autonomously while recognizing its surroundings. It is also the kind of robot that is garnering the most attention nowadays. Information technology companies have entered the self-driving car race, in addition to automobile manufacturers, accelerating the trend toward mergers and acquisitions. The year 2016 also saw proof-of-concepts of self-driving buses and experimental services by self-driving taxis on public streets. In addition, an autonomous trial run of the world's first self-driving delivery truck was conducted on a 190-km stretch of expressway. Importantly, the arrival of deliveries by self-driving trucks is expected to significantly reduce the current truck driver shortage, which continues to increase due to the rapid expansion of e-commerce. It will take time before a completely autonomous car—which does not need human intervention under any conditions—is introduced. However, autonomous driving is already available under certain circumstances.

Additionally, drones are being utilized for a wide variety of business purposes including surveying, 3D map creation, inspections, security, search and rescue, investigations, deliveries, and for entertainment purposes. Drones can fly over difficult terrain and hard-to-reach locations at low cost and identify accurate spatial information of a location. As a result, they hold the potential to increase efficiency and provide new services in an incomparable way.

Robots are also spreading their working arena to commercial facilities, households, and public spaces. For example, there are now robots that use a camera and sensor to patrol the product display shelves to find out-of-stock products, wrong product placements, and messy displays, raising the potential for significantly reducing labor. To assist in everyday life, there are now self-driving vacuum cleaners and communication robots, as well as those that suggest recipes using ingredients stored in the refrigerator, and even those that can cook.

4. Advanced tasks and mass customization

In addition to the automation of simple tasks that humans have performed in the past, even advanced tasks that only experts were able to perform are now being automated. For example, agricultural applications include a drone equipped with a camera and sensor that sprays pesticide only over areas inhabited by pests, or that adjusts the amount of fertilizer depending on the crop growth conditions in a particular area. This drone can perform tasks with a precision far higher than that of humans, and such automation can result in significant savings on pesticides and fertilizers.

Individual customization of products has been difficult to bring to fruition due to cost-related problems. However, in the future, autonomous factories may emerge where robots acquire data from manufacturing machines and sensors, as well as from sales and material procurement departments, and use that data to autonomously determine the necessary materials, most efficient manufacturing process, and methods for coordinating with other machines, thus automatically changing production lines. As a result, mass customization may become a reality.

5. Economic impact of robot popularization

The global robot-related market is predicted to more than double from 91.5 million U.S. dollars in 2016 to 188 million dollars in 2020 [1], with the competition of functions and pricing of robots increasing in the future.

In particular, the automobile industry will very likely reach a significant tipping point. During the development phase of self-driving technology, the car's driving performance has been the major focus. However, once a fully automatic self-driving car is introduced, proficient driving performance will be assumed, and the transportation experience itself will become the determining factor. This means that the car industry will likely shift from the traditional form of selling *things* to that of selling experiences and services. The definition of customers will also change from people who wish to own cars to all people who have transportation needs.

6. Discussions on required reforms of the legal system

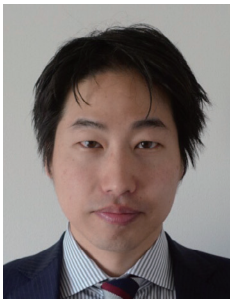
High-performance robots will also bring higher risks of injuring humans and infringing on privacy. Accordingly, the future development of robots will require the resolution of these problems, including new legislation.

Various changes are inevitable in the mid- to-long-term future. Discussions are currently underway to impose a robot tax on owners under the assumption that robots are electronic humans. In addition, the introduction of basic income grants to all citizens in order to maintain a minimum standard of living has been discussed a great deal, with experiments starting in Finland and San Francisco. These discussions are based on the assumption that social structures will significantly change with robots and AI replacing people in the workforce. However, as with computers, new professions will emerge, but different skills will be required. In addition to systemic adjustments in taxes and life security, the education programs needed to fill this skill gap will become vital in the future.

Reference

[1] IDC, "Worldwide Semiannual Commercial Robotics Spending Guide," 2017.

Trademark notes



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Precision Life Science

Yuji Nomura

Abstract

Genetic analysis, biosensors, and electronic health records continuously generate data related to individuals, accelerating the field of data-driven life sciences and enabling root cause analysis of genetics, personal habits, and environmental factors to aid in the treatment and preventive care of individuals.

Keywords: DNA analysis, preventive care, artificial intelligence

1. Worldwide analysis of genomes

How much would the average human life expectancy increase if cancer disappeared from the world? This could become a reality if ongoing progress continues to be made in genome research. In February 2016, the United States launched the National Cancer Moonshot, a project aimed at eradicating cancer. As much as 1 billion U.S. dollars was budgeted for this national project, and one of its major goals is the collection and use of genome data, which is the total information on one person's genes. The objective of this project is the development of precision medicine, which provides prevention and optimal treatment for an individual based on their genome, living environment, and electronic health record.

The United States is not the only country that is developing precision medicine. In 2012, Great Britain contributed approximately 3 billion pounds to the launch of its 100,000 Genomes Project, a national effort to collect and analyze genome data for 100,000 people by the end of 2017. China also announced that it would invest 9.2 billion dollars in a similar 15-year research project.

2. Development of genetic analysis and biosensors

Fueling global analysis of genomes is the enhanced functionality and rapidly declining price of next-generation sequencers. For example, the genome of a person can now be read for 1000 dollars or less. In fact, a do-it-yourself kit for individuals is now available in the market.

The genes of microbiomes, which are microbial populations present within or on the surface of the human body, are also being researched. Microbiomes influence the human immune function and ease of nutritional intake, and are considered an important element for understanding health conditions. For example, medicines are being developed to treat intestinal diseases based on the microbiomes present in human intestines. The microbiome-related market is predicted to reach about 899 million dollars by 2025 [1].

Genetic analysis has an influence not only on medical care, but also on agriculture, food, and other industries. For instance, the genes of hops, the main ingredient of beer, as well as coffee are now being analyzed to produce varieties that are easier to grow, resistant to pathogens, or more bitter. In addition, microorganism soil analysis can determine which soil is most appropriate for specific crops.

The development of biosensors has also enabled the collection and analysis of important health data such as the number of steps a person takes (e.g., in a day), heart rate, blood pressure, skin moisture, and breathing. Emerging measurement devices include a sensor that prevents heatstroke by inferring thirst based on the level of moisture of the skin, and a wearable device that estimates the degree of stress based on the rate of breathing. Smartphones are also being used for blood pressure and pulse rate measurement, eyesight and hearing examination, and ultrasonography. Consequently, it will not be long before users can more easily manage their own health without going to the hospital.

3. Acceleration of life science through artificial intelligence

Artificial intelligence (AI)-assisted drug development is similarly attracting great attention. Generally, a period of more than 10 years, and research and development costs in excess of 1 billion dollars are required to develop a new drug. More than 100,000 proteins exist that can be candidates for the causes of diseases. The number of combinations of chemical compounds working in these proteins runs upwards of the 60th power of 10. As a result, it takes two to three years just to discover a new drug candidate. In addition, only 1% of new drug candidates will be proven effective, leading to even more time needed for non-clinical and clinical tests. For this reason, drug prices remain high.

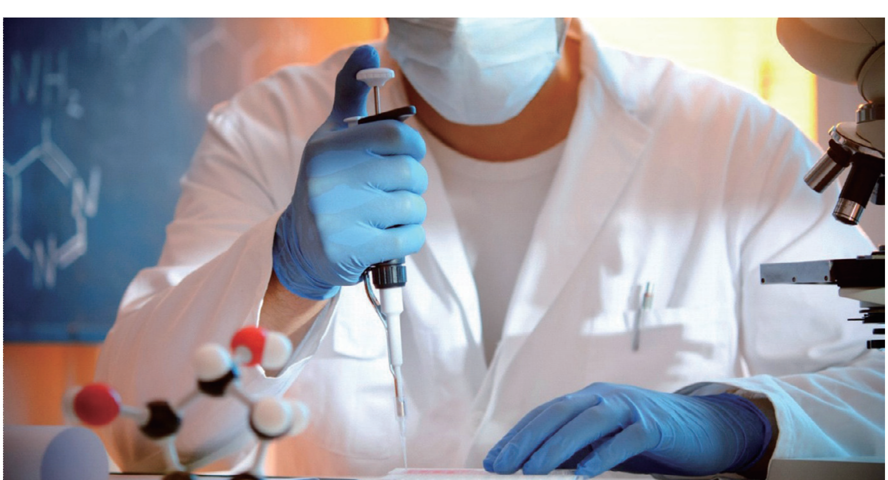
However, if AI learns the data on the molecular structures of existing drugs and their actions, it can then discover the new chemical compound candidates that can act on the proteins that lead to diseases. In one dramatic example, a new drug candidate was discovered in only one day by utilizing AI [2]. If drug candidates that cause harmful side effects are eliminated and new effective drugs can be discovered at a higher probability rate, consumers can expect significantly shorter drug development cycles and lower costs.

The use of AI is also spreading to medical practice. AI can now use treatment data accumulated in the past to assist in differential diagnosis, where symptoms are identified based on an interview. The names of diseases are enumerated in order of high to low probability with required treatment details. Oversights can be avoided in this way regardless of the experience of the physician. In February 2016, DeepMind, a subsidiary of Google known for developing an AI *go* application that beats top professional *go* players, launched DeepMind Health with the purpose of developing AI that specializes in the medical field. In coordination with hospitals, the company has already developed applications for physicians and nurses. After learning past cases, the AI analyzes a patient's condition to determine acute kidney damage that requires an urgent response. In addition, efforts are ongoing to use AI to detect signs of eye disease by reviewing past diagnostic images, symptoms, and treatment details.

4. Expansion of individualized medical treatments and healthcare

Improvement of genetic analysis and AI technology is steadily enabling more individualized medical treatments. For example, it is now possible in the treatment of breast and lung cancer to select the therapeutic drugs that are most appropriate for the genetic structure of the patient. Genetic analysis can also detect an individual's future physical tendencies such as a higher possibility of cancer, thereby enabling preventive measures. This popular trend is significantly contributing to medical care.

However, one important issue is the disparity in the amounts of analyzed genome data depending on nations and races. A majority of genome data comes from the United States, Europe, and Japan. The data on East Indians, who account for approximately 20% of the world population, are estimated to account for only 0.2%. This means that drugs and treatment methods based on the current genetic data may not work for many people. It will be necessary to conduct comparison and analysis using groups of many different backgrounds. To do this, the cost of genetic analysis will need to drop further, and governments must lead efforts to build systems for genome data accumulation.



5. From treatment to prevention

We are now entering a prevention-centered world. Cookie-cutter treatments are no longer adequate. Disease prevention must become more precise and appropriate for each individual. Compound analysis of an individual's genetic information, vital sign sensor data, and lifestyle will enable the prediction of risk factors that increase the probability of falling ill, as well as appropriate preventive measures prior to the appearance of symptoms. In this way, the focus of medical care is transitioning from the treatment of symptoms to the prevention of causes.

A number of such services have already emerged. A prime example of this is the delivery of food that is optimal for an individual based on the results of his or her genetic and blood tests. An aging world population is expected to result in skyrocketing medical costs. However, if an individual's health can be enhanced through preventive measures, a significant reduction in medical costs can be achieved.

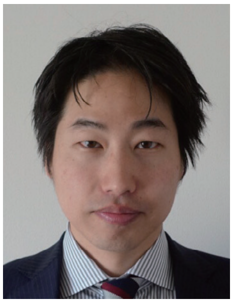
What kind of future will await us if individualized prevention and treatment become a reality? The average life expectancy will increase, and the day may come when many 90-year-old people remain working and in good health. Although the average life expectancy of humans has increased, the individual longevity record has not changed since 1997 when a female died at age 122. Further technological innovation may be required to enable humans to reach a more advanced age.

References

[1] Markets and Markets Press Release, "Human Microbiome Market Worth 899.1 Million USD by 2025," Mar. 2017.
http://www.marketsandmarkets.com/PressReleases/human-microbiome.asp

[2] Atomwise Press Release, "Atomwise Finds First Evidence towards New Ebola Treatments," Mar. 2015.
http://www.atomwise.com/atomwise-finds-first-evidence-towards-new-ebola-treatments/

Trademark notes



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.

Synthetic Reality

Rena Oi

Abstract

With the rapid evolution and diffusion of virtual reality and augmented reality devices, the digital and real worlds are being further integrated and expanded. Human perception will become synthesized within new three-dimensional spaces, enabling the sharing of knowledge and distributed experiences.

Keywords: virtual reality, augmented reality, immersion experience

1. Rapid evolution and diffusion of virtual reality devices

Head-mounted displays (HMDs), which provide immersion experiences in virtual space, finally reached full-scale adoption in 2016. The technology began when different companies sought to replace human visual perception with that of virtual reality (VR) and introduced VR-HMD devices, which achieved widespread popularity in approximately 2010. These efforts coincided with the time when VR-HMD technologies became available at affordable prices. Typical units consist of a computer that generates virtual three-dimensional (3D)-space images in real time, a position-tracking technology that tracks human movements, and a small display that enables high-definition images. With prices one-tenth of those in the past, these products are rapidly becoming popular in many households.

2. Immersion experience generated by VR devices

The experience of replacing human eyesight with a virtual space is a sensation that humans have never before experienced. It is completely different from that of 3D television, where the audience passively watches images that seem to pop out of the screen. With a VR device, the user's eyesight is completely covered by a virtual space. A movement of the user's neck triggers the rendering of additional virtual space, which rapidly heightens the sense of immersion. In addition, moving the face closer makes an image look larger.

Successful VR-HMDs have inspired great anticipation for the development of a device that replaces tactile and other human senses with VR. However, a method to recreate the sensation of touching an object, lifting it to feel its weight and feeling its movement or temperature as well as other sensations is still in development. An enormous technological barrier needs to be overcome before minute finger movements can be sensed together with appropriate tactile sensations provided as a response. To accomplish this, the development of an intuitive controller that minimizes the user's recognition of isolation from the real world, and a combination of a visual effects device and a suit that enhances sensory experiences will be required.

3. Pioneering practical applications of augmented reality

The development of technologies called augmented reality (AR) and mixed reality (MR) also made significant progress during 2016. While VR means a complete replacement of the real-world view with virtual space, in AR the real world is blended together with virtual space. Unlike VR, the goggles used in an AR-HMD are transparent, letting the user see the real world. Virtual space is pasted on real-world space to create an intermingled world. Expected applications for this technology include learning, training, and design. For example, the user can make a gigantic aircraft engine appear virtually in the room, observing its details at will by getting close to and into the object.

Many developers and forward-thinking companies have been focusing their investments on AR to speed its evolution. However, AR-HMDs currently have several technological obstacles, and devices on the market are only for developers. One significant limitation is that the area viewable in VR is still narrow, enabling VR to blend with the real world only in a specific area. In spite of its present limitation, the superimposition of VR to a human's natural eyesight may someday lead to a revolutionary user interface (UI) to connect humans to information systems.

AR is also expected to generate new applications designed for the flat screen of a smartphone. An example of the many commercial uses of such applications include a situation where the user takes photos of his/her room and superimposes the image with that of a new piece of furniture prior to purchase. Although VR is used for product introductions and advertisements, more widespread adoption has been hindered by psychological and physical resistance to wearing goggles. Smartphone AR can overcome this obstacle. Furthermore, AR may well be effective even under the current circumstances, where the development of a device that recreates human senses, except for eyesight, is still in its infant stage.

For example, delicate assembly work and training in medical practice require the use of actual tools or a full-sized model to repeat a task or operation while feeling a tactile sensation. AR can paste a patient body or a material that changes while being processed into the real space using VR. This enables effective and repeatable training that blends VR with a reality where an object such as a model and a tool can be physically touched.

4. Improvement of devices

An HMD interface that connects the user to a virtual space still has significant room for improvement. For example, an HMD needs to be connected by a thick cable to the base unit of a high-speed computer. For this reason, the user with an HMD cannot move about at will. An HMD should optimally be wireless. The next best thing would be to have an HMD without a base unit, which is being proposed. This HMD copies an idea from a smartphone processor that has an advanced computation ability and also conserves power use. As the built-in lenses become thinner, the current situation, where a large, heavy, black box is worn in front of the face, will gradually improve.

This technology will also be combined with the reproduction of images with a definition of over 8K, which exceeds the level recognizable by the retina, and with a response delay of 1 ms or less. In addition, with the release of more accurate, innovative, and outdoor-ready position-tracking technology, the so-called *drunken* issue, caused by the gap between the position information detected by humans and the virtual space, the difficulty in outdoor use, and other limitations, will most likely be mitigated. There is also a proposal to develop a system that enables collaboration without HMDs by pasting a virtual space in the physical space, with the ability to interact with it using projection mapping technology.

5. Connection with 3D space

As HMDs advance and improved interfaces are provided, it is expected that people will use VR on a daily basis. An interface called VROS (virtual reality operating system), which provides an integrated UI by combining different virtual spaces created by the computer, will emerge as well. This type of UI will probably blend the human-to-computer interactions with the real world, making us forget that it is a system UI.

However, the essence of the current fast-moving technical evolution may be a new exploration into the unknown world of

human senses. Thanks to the diffusion of HMDs and the accompanying increase in the number of developers, as well as huge amounts of user feedback, in-depth knowledge about different human senses is being accumulated. For example, the size of an object identified by human eyesight has been assumed to change depending on the person's attentiveness. This is a challenge for 3D content producers as they struggle to determine the display size of an object that humans feel is appropriate. Some have reported that the actual-world action of catching a ball flying toward you is done with different movements or speeds if the same action is presented through VR. Knowledge such as this will not only be used in the recreation of human senses as a form of entertainment, but also in the creation of more natural interfaces, contributing to the improvement of the system-human relationship.

Rena Oi
Assistant Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.
She received a B.A. in sociology from Keio University, Tokyo, in 2008. After joining NTT DATA in 2008, she worked on the development of a customer contact system. She joined the NTT DATA Technology Foresight team in 2017.

Security for the IoT Era

Rena Oi

Abstract

Internet of Things devices have enhanced the value of data collection by collecting more detailed and broader information. However, they have also increased the risk of data breaches and large-scale cyberattacks. While value associated with utilization is being realized, it is now necessary to change the way we treat and protect data.

Keywords: cyberattacks, stray IoT, personal data

1. Escalation of cyberattacks

The escalation of cyberattacks and the damage they cause know no boundaries. Today, the process that starts with these attacks and ends with the extortion of funds has been established as a business model. As criminal programs proliferate, more copycat offenders emerge, which requires the techniques and tools for defending systems against cyberattacks to be continuously strengthened. This vicious cycle is accelerating.

Data breaches are steadily increasing, and billions of individual user accounts are being leaked. This information includes account data of online services, email addresses, associated passwords, and sometimes secret questions and answers. This personal information is traded on the dark web, an Internet space accessible only by a special means. The damage is usually exposed two or three years after the fact. As a result, it is imperative to foster human resources that possess the skills to detect leak-seeking cyberattacks, deal with the ever-changing modus operandi, and continue to protect information. Because it is virtually impossible to completely prevent attacks, the promptness of detection and recovery is vital. Efforts are underway to use technology to disconnect the affected section from the network as soon as possible as well as to implement artificial intelligence (AI)-assisted automatic protection.

Ransomware damage also continues to spread. Ransomware, also called ransom-demanding malware, encrypts infected data in a computer to make it unusable and demands a ransom for decryption. It became common around 2013 and remains an effective means of attack. Cyberattackers purchase low-cost ransomware tools on the market, distribute them, and wait to obtain ransoms, repeating this *business*. It is difficult to control this crime except by improving the computer literacy of the public. The damage is expected to continue.

2. Increase in cyberthreats due to Internet of Things

In 2016, the Internet of Things (IoT) was much discussed as an element that makes cyberattacks more serious. IoT devices, which include surveillance cameras, network devices, video recorders, and electric meters, are powerless, stand-alone computers that are always connected to the Internet and therefore exposed to the risk of information breaches and takeovers. These devices tend to run for a long time with a widely known initial password. Cyberattackers look for these *stray* IoT devices on a daily basis, and upon finding one, exploit its vulnerability.

By the end of 2016, the number of stray IoTs in the world rose to hundreds of thousands. Cyberattackers organized them as a botnet, a group of bots working for cyberattackers, and used them in the largest DDoS (distributed denial of service) attacks in history. Huge numbers of packets were sent to specific websites in classic attacks intended to disable access to those sites and resulting in enormous damage where multiple online services were suspended. A variety of similar attacks ensued. As with other cybercrimes, the IoT botnet is probably already established as a business and its use is on the rise.

The search for specific security measures to respond to IoT security threats is gathering speed. With the computer security knowledge already established for IoT, it will probably be possible to eliminate the vulnerability of IoT. However, the difficulty of dealing with the costs for such measures, combined with the lack of incentives for action, may present obstacles. The imposition of liability to IoT device manufacturers, and authenticating and labeling these devices to prevent the emergence of stray IoTs, is currently under discussion. In Japan, an IoT security guideline was published, and detailed measures are under review.

3. Rapid rise in quality and quantity of personal data due to IoT

The value of secured data, especially personal data that should be protected and kept intact, will be changing. Traditional personal data include data that users registered themselves and data obtained when users' behaviors on the Internet were recorded. In the future, personal data will include behavioral and biological information of individuals captured from the physical world by IoT devices. As the amount of information continues to increase, it is assumed that a dramatic improvement in analytical ability will determine more detailed attributes of individuals. Even today, personal data on the Internet such as clicking actions on a browser, postings on a social networking service, and products selected, are thoroughly collected. For example, the reason for the extremely accurate recommendations made for on-demand videos is the presence of a system that records and analyzes the user's viewing history, time, search history, pausing, fast-forwarding, and even scrolling actions when selecting a video.

In the future, more information will be added to the data on individual behaviors that are observed by IoTs in the physical world. For example, wearable devices will obtain heart rate, blood pressure, running distance, speed, and position information. Navigation systems will collect the driving area, speed, and vehicle condition. Security cameras will catch images of faces, which are automatically identified and analyzed. A system will record the rail stations where people get on and off a train, the duration of the ride, and the time it takes to change trains. As long as the user agrees, all this information will be attached to personal data, making personalization more robust.

However, additional data created through estimation by AI based on huge amounts of combined personal information could bring negative results and disadvantages. For example, if estimated data on personal details such as health condition, healthy life expectancy, and the possibility of illness in the future are created, all kinds of personalization could happen, from the presentation of certain drug advertisements to the control of advertisements soliciting for insurance policies or loans. This negative type of estimated data is produced before the individual realizes it, so it cannot be changed even if the information is wrong. This will cause disagreements if individuals have no way to alter the wrong information.

4. Status of security and information

As we face problems about the treatment of widely collected personal and estimated data, we can no longer avoid the key question of who owns the data. Currently, information about private citizens belongs to companies that collect and analyze that information. As a result, many countries require the agreement of private parties on the collection of their information and/or have legislation that governs the use of such information by third parties.

Trials are underway to build a data trading market that promotes the fee-based distribution of valuable information. Meanwhile, there are services that use anonymization technology to process the collected personal data into forms that do not enable individuals to be identified so that the data can be used in marketing without approval. The GDPR (EU General Data Protection Regulation), which aims to protect particularly critical personal data, is expected to go into effect in May 2018. The time has come for companies to build specific plans to adhere to this framework.

Large companies, who hold a dominant position on the Internet, particularly in the area of information distribution by smartphones, have already established a system where individuals who provide personal data can control their own information. In addition, in an attempt to acquire the information on IoT devices, large companies are expanding their service menus. They are also providing a console where users themselves can configure the opt-ins broken down by types of information, the viewing of collected information, and the customization of advertisements. In other words, large companies are building a system to store information, while meeting the changes in the status of information.

It is possible to create an organization to counteract this oligopoly and to actively distribute information. However, it will be a long-term challenge to establish a business model where the new organization will acquire the trust needed to take care of large amounts of information and protect it. One of the first ideas may be data exchange markets for open data collected by IoT devices, which a public organization owns. Accumulating such experiences will lead to the next stage of dealing with valuable information, such as personal information, which generates more profit.

Rena Oi

Assistant Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

She received a B.A. in sociology from Keio University, Tokyo, in 2008. After joining NTT DATA in 2008, she worked on the development of a customer contact system. She joined the NTT DATA Technology Foresight team in 2017.

Heterogeneity in IT Infrastructures

Rena Oi

Abstract

To supply the massive computing performance required for artificial intelligence and the Internet of Things, new infrastructure is needed for both general use and specific purposes. Cloud services will rapidly enable such a flexible future infrastructure.

Keywords: graphics processing unit, field-programmable gate array, edge computing

1. Change in the leading processor

The evolution of technology is insatiable, demanding more and more calculation ability and processor power. In the field of artificial intelligence (AI), it is robust processor power that led to the breakthrough and commercialization of deep learning. In self-driving cars, it is processor power that enables real-time identification of objects surrounding the car from four directions to determine the next operation of the steering wheel. It is processor power that, based on the movement of the user, allows real-time generation of high-definition images on a virtual reality display. Finally, it is processor power that makes possible the more in-depth analysis of Internet of Things (IoT)-generated data for marketing purposes.

Processor manufacturers continue to compete in the development to meet the insatiable demand for advanced calculation ability. In the past, the calculation ability of the central processing unit (CPU), a multi-purpose processor, was continuously being improved as the standard. Now processors that excel in parallel execution of many homogeneous processors have become the leader in calculation. Because emerging technologies of recent years required large-scale parallel processing, the adoption of the graphics processing unit (GPU) has increased.

In the AI area, the GPU achieved an efficiency rate ten times higher than that of the CPU, making the GPU a de facto standard. In addition, the emergence of the GPU has coincided with the enhancement of software libraries, which utilize the GPU's characteristic features. As a result, the use of the GPU has been increasing from image processing, which was its original purpose, to advanced scientific computation. GPUs and processors with many cores that compete are used in devices ranked high in TOP500, which ranks the world's supercomputers. These parallel processors are also continuing to evolve, maintaining their leadership in speed for the foreseeable future.

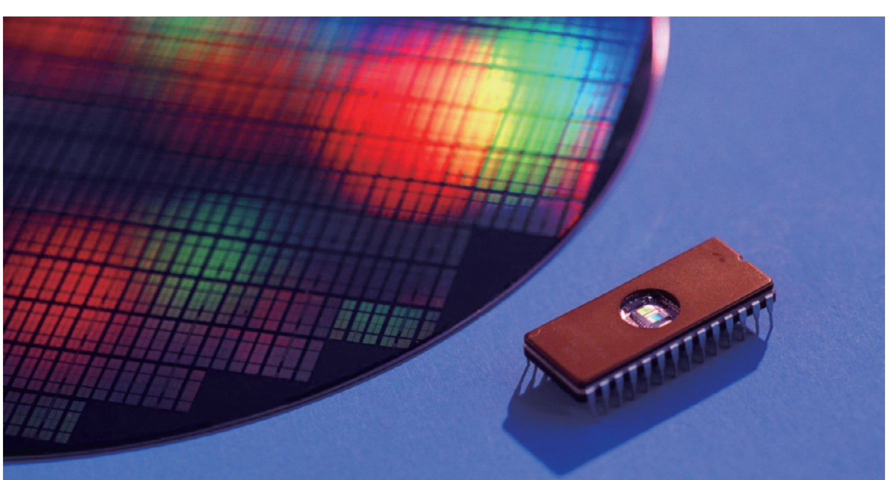
2. Diversification of processors

In some cases, even the power of the GPU is not adequate, so a processor called the field-programmable gate array (FPGA) is used. While the logic on regular processors is printed at the time of manufacture and cannot be changed, FPGA processors can be rewritten later and as many times as necessary. With FPGAs, it is also possible to manufacture only a few units of processors specialized in specific applications. In the financial industry, FPGAs are used to perform high-speed algorithmic trading of stocks and currency exchange, with the logic being changed as needed and tuned at the high-speed level of milliseconds to chase profits. In the competitive area of AI, dedicated logic is implemented on an increasing number of FPGAs in an attempt to achieve speed and efficiency that exceed those of GPUs. In addition, efforts are underway to install several thousand FPGAs at a time to improve the overall processing ability of datacenters.

A trend to use smartphone processors is also attracting attention. The processor of a smartphone uses multiple cores with different abilities to enable the phone to go to sleep or shut down in a short period of time, improving the efficiency of electric power usage. Low electric power usage and robust processing ability at an advanced level are spreading the use of smartphone processors to drones, connected cars, and other IoT devices. In the future, the use of smartphone processors is expected to further streamline the operation of datacenters.

Some companies design, manufacture, and use their own dedicated processors for specific purposes. One such company developed its own processor that specializes in deep learning and that is supporting the fast development of AI today. This specialized processor is currently installed in datacenters to streamline operations. It is a low-power-consuming processor based on the knowledge that even if the computation of floating-point arithmetic is less accurate than normal, it can be used in deep learning. The latest GPUs have also adopted this approach of lowering the computation accuracy of floating-point arithmetic to achieve higher speed and improved efficiency. Deep learning is a particularly competitive area, and it is driving increased competition in the development of specific-purpose processors.

The important characteristic of this new calculation ability is the development of software that utilizes the features of a processor. In particular, parallel processing would not exist without the support of software. The difficulty level is high to develop this type of software, which makes it challenging for the processor to reach its potential. Processor manufacturers are enhancing software parts called libraries to make it easier for software developers to harness the processor power.



3. Diversification of architecture

Like processors, system architectures are also diversifying. Edge computing is an architecture that uses huge numbers of IoT devices in the field, that is, the edge, to make distributed processing more flexible and dynamic. When immediate processing is necessary in the field, or accumulation and compression of collected data are required, the edge processes these tasks while coordinating with the center as needed. Application to automatic operation is expected soon, and development is ongoing along with the use of next-generation 5G (fifth-generation mobile communications) networks.

The blockchain, a distributed ledger technology, is also attracting attention as an infrastructure technology that supports Bitcoin virtual currency. The blockchain enables distributed data to be shared on the network while preventing data falsification. It is innovative in that it does not require the building of a robust and centralized database. Consequently, systematization of complex areas that have traditionally been impossible to do on a cost-effective basis is now possible. In addition, funds transfer and payment uses, financing, contract management, and the government's notarization service are undergoing proof-of-concept trials. Bitcoin has been operating ever since its release, and new implementations of high-quality blockchain technology are expected in the future.

4. Utilization of flexible options

To bring business ideas that constantly use new technologies to fruition, it is essential to master the use of a system infrastructure that includes diversified and complicated processors and architectures. However, this kind of system infrastructure does not need to be owned. Clouds are taking the place of diversified infrastructures at a rapid pace. For example, users of the GPU, which is required for machine learning by AI, can select from available menus on the cloud. Even the special FPGA hardware is on clouds.

Mechanisms that never existed before, such as the blockchain, are now provided by SaaS (software as a service). Whenever a new architecture appeared in the past, the barrier that faced an engineer was the preparation of the environment in which it operated. Now this barrier is considerably lower. New streamlining methods also exist, for example, server-less architectures, that only clouds can generate. As a result, it is now possible to launch a worldwide data processing service at low cost and without considering infrastructure preparation or performance design. It is also easier than before to migrate large amounts of data from an on-premise environment to a cloud and to retrieve it or migrate it to another cloud. Clouds enable the user to acquire any amount of data at any time, and to withdraw it at any time. In addition to being of great benefit to users, clouds bring flexible options in the use of system infrastructures to business.

5. Changes required of users

Users need to acquire more knowledge and know-how to be able to handle more diversified and complicated infrastructures. Processors today achieve high speeds by cooperating with software. However, once a performance problem occurs, it is difficult to repair it by simply scaling up or scaling out. Although a GPU's software library reduces the burden on developers, its performance may be compromised without tuning it in a way that considers the library features. Even though the FPGA is on clouds and is operable on a browser, it is essential to have the necessary knowledge to implement the logic unique to this hardware. Because clouds present flexible and diversifying options, it is not easy to review the entire system lifecycle and select suitable cloud services, along with an option to leave part of the system on premises. As a consequence, it is necessary to develop human resources who can actively adapt to changes and today's technological trends, which go beyond the framework of traditional software and hardware, and who can reexamine system infrastructure.

Rena Oi

Assistant Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

She received a B.A. in sociology from Keio University, Tokyo, in 2008. After joining NTT DATA in 2008, she worked on the development of a customer contact system. She joined the NTT DATA Technology Foresight team in 2017.

Experience Design Innovation

Rena Oi

Abstract

Development of the API (application programming interface) economy and UX (user experience) design are simplifying the creation and continuous evolution of innovative services. Propagation of the Internet of Things will drastically change interactions between humans and systems, resulting in more natural and freer user experiences.

Keywords: API economy, UX design, Internet of Things

1. Penetration of API economy

Recent years have seen a significant change in the way services are generated. Ideas people have for businesses are being transformed into actual businesses at an extremely fast pace. The company that grew rapidly by being the first to market a ride share service, where private car owners provide rides to customers in their own cars, becoming a kind of taxi service, is a prime example of how fast an idea can become a business. At the onset, these companies outsourced everything except for their smartphone application (app) and related matters. They utilized existing services that they did not own even for elements that normally make up the core of a service, such as a navigation system, credit card payment service, and telephone and messaging services.

The cooperative use of different application programming interfaces (APIs), referred to as API cooperation, is what made this possible. The API economy, which is this simple linkage of existing services via the Internet, enables fast launching of services. In addition, as the ride share service earned a good reputation and became recognized as a stable service, it has in turn been used by other services through API cooperation. For example, the ride share service is now embedded as a button on hotel and restaurant websites and offers a one-click, transportation linkage service to hotel guests and restaurant customers.

Using information technology (IT) to launch a business and to gather a variety of elements necessary for that business is now a common practice. Even small startups can build a new business by combining existing services. Companies can combine services by calling them via the Internet, and release their products with a short lead time, testing their value in the real world. The speed of launching a business is not the only thing that has changed. The subsequent, continuous improvement of user experience (UX) is also accelerating thanks to the use of IT. Through API cooperation, the ideal UX design process, which constantly improves services throughout the lifecycle, including the operation phase, is now being realized faster and more easily.

2. Acceleration of the API economy on Internet of Things platforms

The spread of the Internet of Things (IoT) is upgrading the way business is generated to the next level. The IoT platforms that companies now provide combine an integrated management system that includes networks that correspond with each of the various IoT devices, a server-side system that links with these devices, and a system for processing and analyzing the collected information. For example, these platforms manage IoT devices installed in manufacturing machines or in vehicles and collect information on operational status, fuel consumption, and parts status. Analyzing this huge amount of information means that abnormalities can be detected in devices, and the appropriate time for replacement can be determined. In some cases, the labor status of workers is collected and tallied to be used in balancing the workload.

A move to make IoT platforms open is also expected.

Companies keep many of the details regarding their IoT platform proprietary. The streamlining effect achieved by installing an IoT platform and using the information acquired from it definitely stays within a company. In the future, new businesses will be generated by disclosing the information collected on the IoT platform and the bundles of IoT devices that the companies manage, thereby offering them to the outside world. Trials are underway by some companies to collect information on electric power usage data from households that gave consent, analyze it, and use it in cross-industrial marketing.

At the household appliances trade show held in January 2017, the potential of the IoT platform was shown. As many as 700 IoT products were exhibited with the implementation of API cooperation with a cloud-based voice interaction interface. For example, the smart speaker, which uses natural conversations to order products and services via the Internet, is now a popular product that ships 5 million units per year in North America. The voice interaction interface, which has already been implemented extensively in households and has accumulated an adequate track record in voice recognition, is provided as an API cooperation service. Linking household appliances such as refrigerators and vacuum cleaners to existing IoT devices such as automobiles has made it possible to uniformly upgrade these household appliances to voice-supported services. Furthermore, this voice conversation interface is already linked to 4000 services, including all types of product sales, food delivery reservations, lodging reservations, and taxi calls. There is a high probability that these services will be linked with other services in the future to create platforms that provide increasingly streamlined customer experiences.

3. UX innovation by IT

The service industry has entered the stage where users can pursue a more natural UX, thanks to the power of IT. The rapid evolution of AI has enabled the development of the voice recognition interface and space recognition by computer vision, as well as the integration of actual and virtual space by augmented reality (AR). The latest IT technologies also have the potential for generating an entirely new UX.

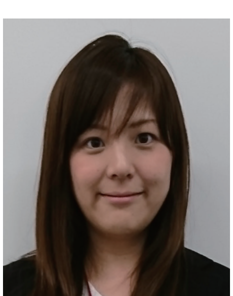
Of particular note is the supermarket without cashiers, which was publicly demonstrated by a large North American e-commerce company. This is an example where IT technology completely rewrote the traditional UX. This store contains many IoT devices, cameras, and sensors. Customers are authenticated by waving a barcode over the gate when entering the store, and their movements are continuously tracked by IoT devices. This tracking also identifies the motion of hands stretching toward a shelf to get a product, hesitation, and even returning the product to the shelf. Customers simply get what they need and exit the store. Smartphones then show the items selected and complete the payment. The UX is that you simply get what you need and go home. This may well be a new UX format that is difficult to imagine for someone who is used to conventional supermarkets.

4. Continuation of UX innovation

UX innovation will probably affect IT technology itself. For example, the operation of a smartphone, which is now a starting point for different kinds of services, will be innovated. Is it really intuitive to stare at a screen full of icons, which act as metaphors for the actual world, select the icon for the target operation, and touch it? Icons (i.e., apps) are independent of each other, their linkage is thin, and even information sharing is difficult. A person who has just started to use a smartphone has a long learning curve until he or she can use it smoothly. This fact alone may prove that a smartphone is an incomplete interface. Instead of depending on the flat screen of a smartphone, it may be necessary to make use of virtual reality (VR), AR, and projection mapping for a more complete interface. Use of natural human gestures is also possible.

However, the fact that humans must actively talk to a machine to interact with it may be unnatural in itself. A voice conversation interface that does not recognize that it is talking to a person without certain keywords is not natural. Current interfaces are triggered by unnatural behaviors that are far from intuitive ones such as looking, speaking, and grabbing movements. Thus, there is still room to innovate the UX in many IT interfaces.

The UX innovation by IT is different from unregulated installation of technologies. Businesses established on the Internet reflect the dynamic attitude that something is worth trying if it is technically feasible. However, such dynamism may decline after negative user feedback. For example, the supermarket without cashiers could adopt all technologies already used by Internet stores. The minute the customer picks up a product, an advertisement using AR might come on, informing the customer of a 2-for-1 deal. In addition, user reviews of that product might be listed. For a customer who always hesitates when reaching for a cupcake, it may be possible to create a mechanism where music is played that heightens the customer's appetite and urge to buy while appropriate images are shown. Are these ideas really appropriate? If a true UX cycle is achieved that repeats improvement based on user feedback, these ideas could be replaced with more appropriate mechanisms. It is assumed that continuous and serious UX innovation with users in mind is the proper collaborative design for both humans and IT.



Rena Oi

Assistant Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

She received a B.A. in sociology from Keio University, Tokyo, in 2008. After joining NTT DATA in 2008, she worked on the development of a customer contact system. She joined the NTT DATA Technology Foresight team in 2017.

Hardware Acceleration Technique for Radio Resource Scheduling in 5G Mobile Systems

Yuki Arikawa, Takeshi Sakamoto, and Shunji Kimura

Abstract

This article presents a hardware acceleration technique for the scheduling process in ultra-high-density distributed antenna systems for fifth-generation (5G) mobile communications systems. In 5G systems, the overall system throughputs for a huge number of combinations of antennas and user equipment (UE) for communications have to be calculated in the scheduling process. To speed up the calculation, this acceleration technique calculates the throughputs of each UE simultaneously. Experimental results show that the acceleration technique calculates the system throughput approximately 60 times faster than without the acceleration. As a result, the acceleration technique improved the throughput by about 73% for a system with 32 antennas and 256 UEs. The hardware acceleration technique therefore enables a future practical 5G system.

Keywords: 5G mobile communications systems, resource scheduling, hardware acceleration

1. Introduction

In mobile communications systems, resource scheduling assigns user equipment (UE) to each antenna for downlink transmission. Scheduling that decides the optimal combination of antennas and UEs is needed for efficient communications and improved overall system throughput [1]. In systems preceding fifth-generation (5G) mobile communications systems, this assignment has been executed by software-based processing due to the small number of antennas, as shown in **Fig. 1(a)**, and the small number of possible combinations of antennas and UEs.

For 5G systems, researchers have been studying flexible antenna deployment such as localized massive multiple-input multiple-output (MIMO) and distributed massive MIMO [2]. In this article, we focus on distributed massive MIMO (i.e., distributed antenna systems). In distributed antenna systems, as shown in **Fig. 1(b)**, a huge number of antennas are deployed at ultra-high density in order to increase the overall system throughput [3, 4]. The number of possible combinations reaches approximately 10^{76} in a system with 32 antennas and 256 UEs, which is based on a 5G-system model in the Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS) project [5].

To obtain the appropriate combination from this explosive increase in the number of possible combinations, the scheduler approximately searches for the appropriate combination [6]. In general, an approximate search can approach the appropriate combination as the number of searched combinations increases. However, it will be difficult to increase the number of searched combinations by using software-based processing within the limited scheduling period of 1 ms [7] because of the limitation in the number of CPU (central processing unit) cores.

To overcome this issue, we devised a hardware acceleration technique that enables the scheduler to accelerate the scheduling process in ultra-high-density distributed antenna systems, and a search process for quickly obtaining the right combination. The details of the search process and the hardware acceleration technique are respectively described in sections 2 and 3.

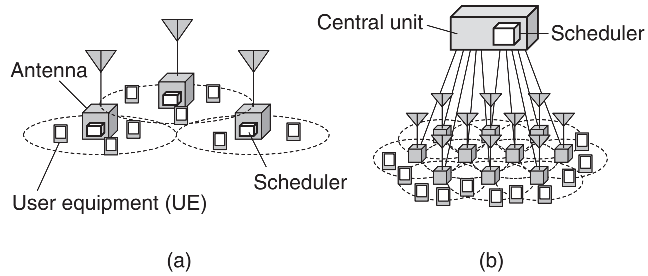


Fig. 1. Illustrations of antenna deployment (a) before the 5G system and (b) in the 5G system (ultra-high-density distributed antenna systems).

2. Search process

The search process approximately decides the combination by iterating processes for improving the combination so that the system throughput increases. In general, the UEs having the highest throughput are simultaneously chosen for all antennas when the scheduling searches for a combination. In the search process, the UE having the highest system throughput is chosen only for one antenna. Then UEs are chosen for the other antennas. This choosing of the UE is carried out one by one by other antennas. In this way, the best UEs are assigned to antennas so that the system throughput always increases. The combination can be approximated to the better combination by iterating this assignment of UEs, and the system throughput is improved.

The procedure for deciding the combination in the search technique is shown in **Fig. 2**. First, all antennas are set to *blank*, which means the radio transmission is stopped. Then, one of the antennas is selected. In the case shown in Fig. 2(a), antenna A is selected. In order to select the UE to which antenna A should transmit data, the system throughputs under this condition are calculated. In this case, the system throughputs of (UE#1, Blank, Blank), (UE#2, Blank, Blank), and (UE#3, Blank, Blank) are calculated, and these three system throughputs are compared. In this example, (UE#1, Blank, Blank) has the highest system throughput. Consequently, UE#1 is provisionally selected for antenna A, and the combination is updated.

Next, antenna B is selected. In the case shown in Fig. 2(b), in order to select the UE for antenna B, the system throughputs are calculated taking the interference power from antenna A and B into account. The throughput of UE#1 may change because the interference power from antenna B changes. When UE#4 is

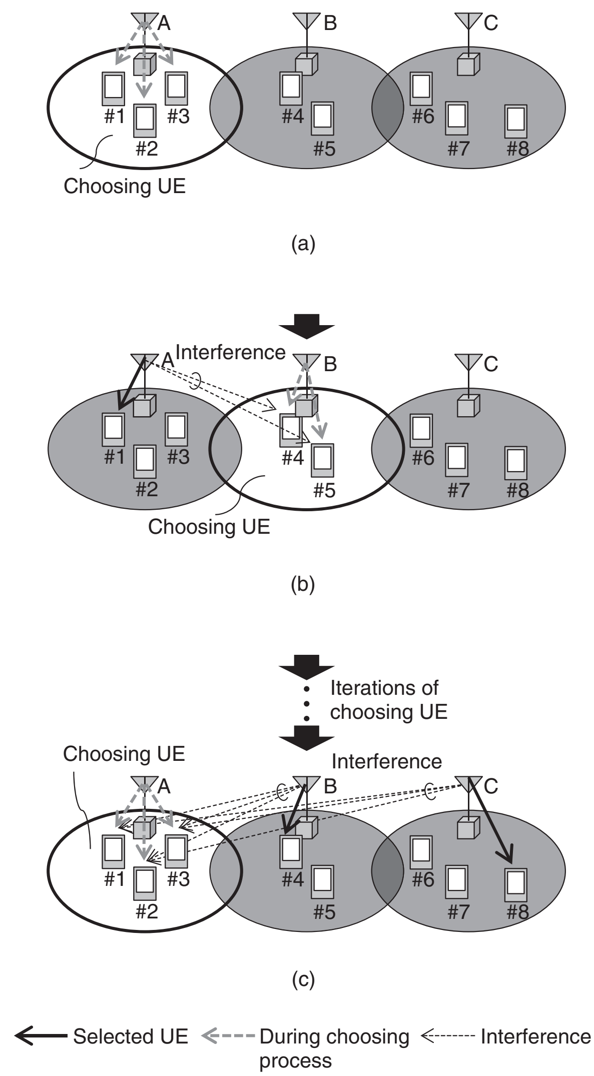


Fig. 2. Procedure for deciding the combination in the search process.

chosen, the system throughput is calculated by summing the throughputs of UE#1 and UE#4. With the same calculation as above, the system throughput is calculated by summing the throughputs of UE#1 and UE#5. These two system throughputs are compared, and the highest one is obtained when antenna B transmits data to UE#4 in this example. Thus, UE#4 is provisionally selected for antenna B, and the combination is updated. This technique enables the scheduler to take inter-cell interference power into account when scheduling decides the combination for the antennas.

In this technique, all antennas are selected one by one. This UE selection for each antenna is carried out sequentially to other antennas. In the case shown in Fig. 2(c), antenna A is selected again, and the throughputs of UE#1, UE#2, and UE#3 are calculated again because the interference powers from antenna B and antenna C change.

In this way, the combination is always updated so that the system throughput increases. This increases the system throughput monotonically as the number of iterations increases.

3. Hardware acceleration technique

The hardware acceleration technique accelerates the calculation of the system throughput for each combination during the search procedure. The search procedure finds the optimal combination to achieve higher system throughput by iterating the processing for improving the combination. As shown in **Fig. 3**, the conventional scheme (without acceleration) requires a longer processing time to obtain the optimal combination. In contrast, the system throughput is quickly improved in the acceleration technique. In this way, the combination that achieves higher system throughput is obtained within the required period.

The hardware acceleration technique for the scheduling process is shown in **Fig. 4**. In this technique, the search process is executed by a dedicated circuit in order to increase the number of searched combinations, which improves the scheduling performance. The software sets possible UEs for each antenna in the hardware accelerator before starting the search process. The details of the hardware acceleration technique are described below.

The flow of the search process performed by the hardware accelerator is shown in **Fig. 5**. First, the combinations of antennas and UEs are generated. Next, the system throughput is calculated by summing the throughputs of all the UEs in the generated combinations. Then, the combination for which the scheduling achieves higher system throughput is decided by comparing the system throughput for each combination. These three steps are iterated until the scheduling period expires so that the optimal combination can be obtained.

We investigated the processing time for each step in order to clarify the steps that should be accelerated in the above search process. The investigation by software-based processing revealed that the system-throughput

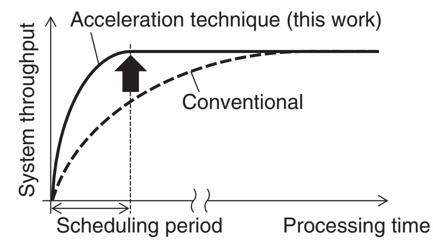


Fig. 3. Concept for achieving higher system throughput.

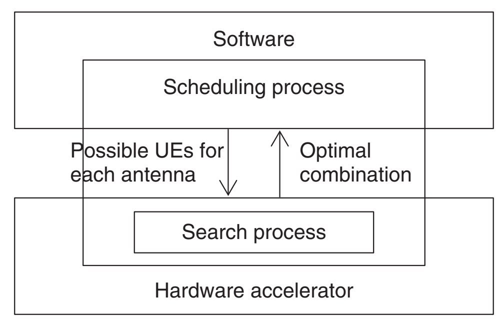


Fig. 4. Hardware acceleration for the scheduling process.

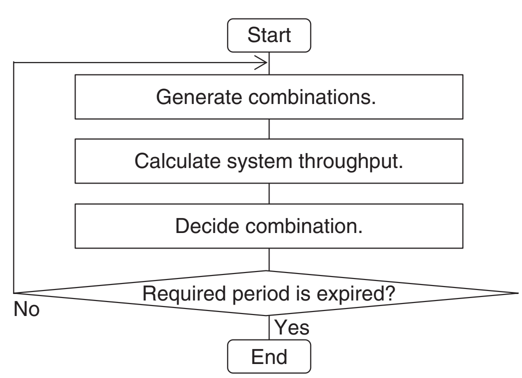


Fig. 5. Flow chart of scheduling process.

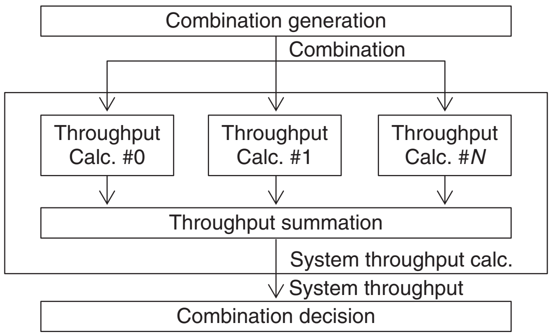


Fig. 6. Circuit block diagram.

calculation accounts for more than 90% of the processing time to execute the search process. On the basis of the results, we devised a parallel and pipeline processing technique to accelerate the system-throughput calculation.

A block diagram of the circuit is depicted in **Fig. 6**. The circuit comprises three parts: a combination-generation part that outputs the combination of antennas and UEs, a system-throughput-calculation part, and a combination-decision part that decides the combination by comparing the system throughput for each combination. Our proposed technique consists of two kinds of processing: parallel processing to calculate the throughput for all UEs simultaneously and pipeline processing to obtain the system throughput for generated combinations at every clock cycle.

The system-throughput-calculation part consists of multiple throughput-calculation blocks that output the throughputs of the UEs in parallel. The throughput-calculation blocks are provided with the same number of antennas. The throughput-summation block outputs the system throughput by summing the throughputs of the UEs. The throughputs of the UEs are simultaneously calculated at the throughput-calculation blocks. Hence, the circuit executes the search process at high speed. Furthermore, the circuit scale can be minimized by optimizing a parallel number for the same number of antennas.

The timing chart of the system-throughput-calculation block is shown in **Fig. 7**. In this block, the received signal power to interference power and noise ratio (SINR) is calculated. Next, the calculated SINR is converted to the throughput. Then the throughputs of the UEs are summed. These steps are independent of the preceding and the following combinations. Therefore, these steps are executed in the pipeline. This enables the scheduler to obtain the system throughput for generated combinations at every clock cycle.

4. Performance evaluation

We carried out experimental measurements and system-level simulations in order to evaluate the performance of the acceleration technique and the search process.

To verify the number of searched combinations within the scheduling period of 1 ms, we measured the processing time spent for the search process. The proposed technique was implemented on a field-programmable gate array (FPGA) (Xilinx Zynq-7045) at the clock frequency of 100 MHz. The processing time was measured with the FPGA. The processing time without acceleration was measured with a general-purpose processor (Intel Core i5) at the clock frequency of 2.67 GHz.

We carried out system-level simulations to evaluate the performance of the proposed scheme. The performance was evaluated in practical conditions based on the small-cell scenario in LTE (Long-Term Evolution) specifications [8]. The simulation conditions are listed in **Table 1**. The simulation conditions were based on an assumption of ultra-high-density distributed antenna systems, so 32 antennas were uniformly distributed in a circle with a radius of 155 m. The minimum distance between antennas was 20 m.

The results of the performance evaluation are given in **Table 2**. The processing time per searched combination measured with acceleration was 10 ns. The processing time without acceleration was 596 ns. The circuit executed the search process about 60 times faster than without acceleration. These results indicate that the number of searched combinations within the scheduling period of 1 ms using the proposed technique was 10^5, and the number was 1679 in the processing without acceleration. These results show that the number of searched combinations with the proposed technique is about 60 times larger than without the acceleration.

Furthermore, we carried out system-level simulations to evaluate the system throughput. The acceleration technique improved the system throughput by about 73% when there were 32 antennas and 256 UEs. Consequently, the proposed technique enables the scheduler to obtain the appropriate combination in ultra-high-density distributed antenna systems.

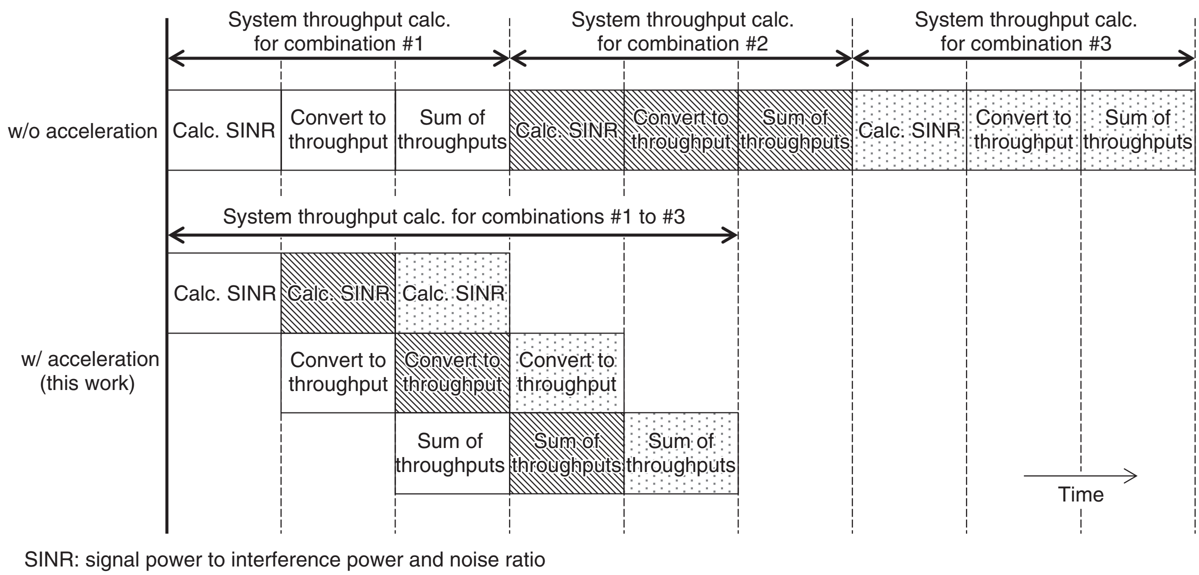


Fig. 7. Timing chart.

Table 1. Simulation conditions.

Parameter	Value
System bandwidth	10 MHz
Carrier frequency	3.5 GHz
Transmission power	30 dBm
Number of antennas and UEs	32; 256
Antenna deployment	Uniform in a circle with a radius of 155 m
Minimum inter-antenna distance	20 m
Antenna height	10 m
UE distribution	Uniform in a circle with a radius of 155 m
Antenna height at UE	1.5 m
Traffic model	Full buffer
Propagation loss	Line of sight: 22.0 log10(R) + 38.9 dB Non-line of sight: 36.7 log10(R) + 36.8 dB For 3.5 GHz, R in meter
Fading model	Rayleigh fading
Shadowing standard deviation	Line of sight: 3 dB Non-line of sight: 4 dB
Noise figure	9 dB
Thermal noise power density	–174 dBm/Hz

5. Summary

In this article, we proposed a hardware acceleration technique that accelerates the search process in the scheduling of ultra-high-density distributed antenna systems. Our technique consists of parallel processing to calculate the throughputs of each UE simultaneously and pipeline processing to obtain the system throughput for combinations at every clock cycle. As a result, it performs the search process 60 times faster than processing without acceleration. The proposed technique enables the scheduler to substantially increase the number of searched combinations. Consequently, the proposed technique enables the

Table 2. Results of performance evaluation.

	w/o acceleration	w/ acceleration (this work)
Processing time (ns)	596	10
Number of searched combinations	1679	100,000
System throughput (Mbit/s/REG*)	37	64 (73% improvement)

* resource element groups.

scheduler to obtain the appropriate combination in ultra-high-density distributed antenna systems. With the acceleration techniques, the scheduler improved the system throughput by about 73% when there were 32 antennas and 256 UEs. The scheduling with the proposed technique therefore enables a practical 5G system.

Acknowledgments

This article includes part of the results of "The research and development project for realization of the fifth-generation mobile communications system" commissioned by The Ministry of Internal Affairs and Communications, Japan.

References

[1] 3GPP TS36.300: "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Overall Description," V13.3.0, Apr. 2016.

[2] S. Suyama, J. Mashino, Y. Kishiyama, and Y. Okumura, "5G Multi-antenna Technology and Experimental Trials," Proc. of 2016 IEEE International Conference on Communication Systems (ICCS), pp. 1–6, Shenzhen, China, Dec. 2016.

[3] NTT DOCOMO, "DOCOMO 5G White Paper, 5G Radio Access: Requirements, Concept and Technologies," July 2014.

[4] T. Seyama, M. Tsutsui, T. Oyama, T. Kobayashi, T. Dateki, H. Seki, M. Minowa, T. Okuyama, S. Suyama, and Y. Okumura, "Study of Coordinated Radio Resource Scheduling Algorithm for 5G Ultra High-density Distributed Antenna Systems," Proc. of 13th IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS 2016), S3-5, Tokyo, Japan, Aug. 2016.

[5] ICT-317669 METIS project, "Simulation Guidelines," Del. D6.1, Oct. 2013.
https://www.metis2020.com/wp-content/uploads/deliverables/METIS_D6.1_v1.pdf

[6] Y. Arikawa, K. Kawai, H. Uzawa, and S. Shigematsu, "Practical Resource Scheduling in Massive-cell Deployment for 5G Mobile Communications Systems," Proc. of the 2015 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), pp. 456–461, Bali, Indonesia, Nov. 2015.

[7] 3GPP TS36.321: "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification," V12.5.0, Mar. 2015.

[8] 3GPP TR 36.872: "Small Cell Enhancements for E-UTRA and EUTRAN - Physical Layer Aspects," V12.1.0, Dec. 2013.

Trademark notes

Yuki Arikawa

Research Engineer, Fixed Mobile Convergence Device Development Project, NTT Device Innovation Center.

He received a B.E. and M.E. from Waseda University, Tokyo, in 2008 and 2010. Since joining NTT in 2010, he has been studying device architecture for 10G-EPON and 5G mobile communications systems. He is currently in charge of the research and development project for realization of the 5G mobile communications system commissioned by The Ministry of Internal Affairs and Communications, Japan. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Takeshi Sakamoto

Senior Research Engineer, Project Leader of Fixed Mobile Convergence Device Development Project, Metro-Access Network Device Project, NTT Device Innovation Center.

He received a B.E. and M.E. in electronic engineering from Kyoto University in 1994 and 1996. He joined NTT Opto-electrical Laboratories in 1996. He is currently engaged in developing digital devices for optical access networks. He is a member of IEICE.

Shunji Kimura

Senior Research Engineer, Supervisor, Project Manager of Metro-Access Network Device Project, NTT Device Innovation Center.

He received a B.E. and M.E. in electrical engineering in 1989 and 1991, and a Ph.D. in electronics, information, and communication engineering in 1997, all from Waseda University, Tokyo. In 1991, he joined NTT LSI Laboratories, where he worked on technology for designing and evaluating analog front-end integrated circuits for 40-to-100-Gbit/s-class optical transmission systems. In 2003, he was assigned to NTT Access Network Service Systems Laboratories, where he was engaged in the development of the next-generation optical access networks, 10G-EPON and NG-PON2. He is currently a project manager at NTT Device Innovation Center and is leading the development of metro-access network devices. He received the 1994 Japan Microwave Prize at APMC94, the 1996 Young Engineer Award from IEICE, and Best Paper Awards at OECC2010 and COIN2010. He is a senior member of IEEE (Institute of Electrical and Electronics Engineers) and IEICE.

Network Resource Management Technology

Shingo Horiuchi, Kazuaki Akashi, Masataka Sato, and Tadashi Kotani

Abstract

The era of software-defined networks is approaching, and NTT is working on network management and operation architecture for various types of networks and their topologies, including current physical and virtual networks. This article presents the technology used in network resource management architecture for comprehensive network operation based on unified information models.

Keywords: network resource management, operations support system, unified information model

1. Introduction

In recent years, the business situation of telecommunication operators has been changing around the world. This change has been most remarkable in two key domains: the network domain and the business domain.

In the network domain, much progress has been made in virtualization technology, which enables network functions to be implemented by software, and network construction, which will be achieved using low-cost, general-purpose servers and switches. This is in contrast to conventional networks that make use of dedicated communications equipment, and thus, there has been concern that the separation of hardware and software by virtualization technology will drive up the cost of operations.

Meanwhile, in the business domain, telecom carriers are looking to transform themselves from simple circuit service providers to business co-operators based on the B2B2X (business-to-business-to-X) model that provides customers with services via other service providers. The goal here is a service-creating collaboration with all kinds of service providers spanning diverse business fields. This means that telecom carriers themselves must be flexible for various types of businesses, which, in turn, means that telecom operations must be able to provide services that can combine a variety of networks and other services in a flexible manner.

Telecom operations include the work of accepting a service order from a customer, configuring network equipment based on service requirements, and providing that communications service to the customer. Such work can be made more efficient through the use of an operations support system (OSS). An OSS is equipped with a database for storing the management information needed for each type of network communications method and functions for managing that information (**Fig. 1**).

In constructing an OSS, the following requirements must be met to prevent operations from becoming overcomplicated through virtualization and to provide support for diverse types of businesses:

(1) It must be possible to add network equipment and network functions such as firewalls.

(2) It must be easy to provide new services by combining network and service elements.

The approach up to now has been to develop and construct an OSS specifically for each communications service and its network provided by the telecom carrier. This development and construction format, called a *silo* approach, can incur considerable development costs with regard to requirement (1) above [1].

To satisfy requirement (1), efforts have been made to unify the information model managed by the OSS. In particular, an information framework called the

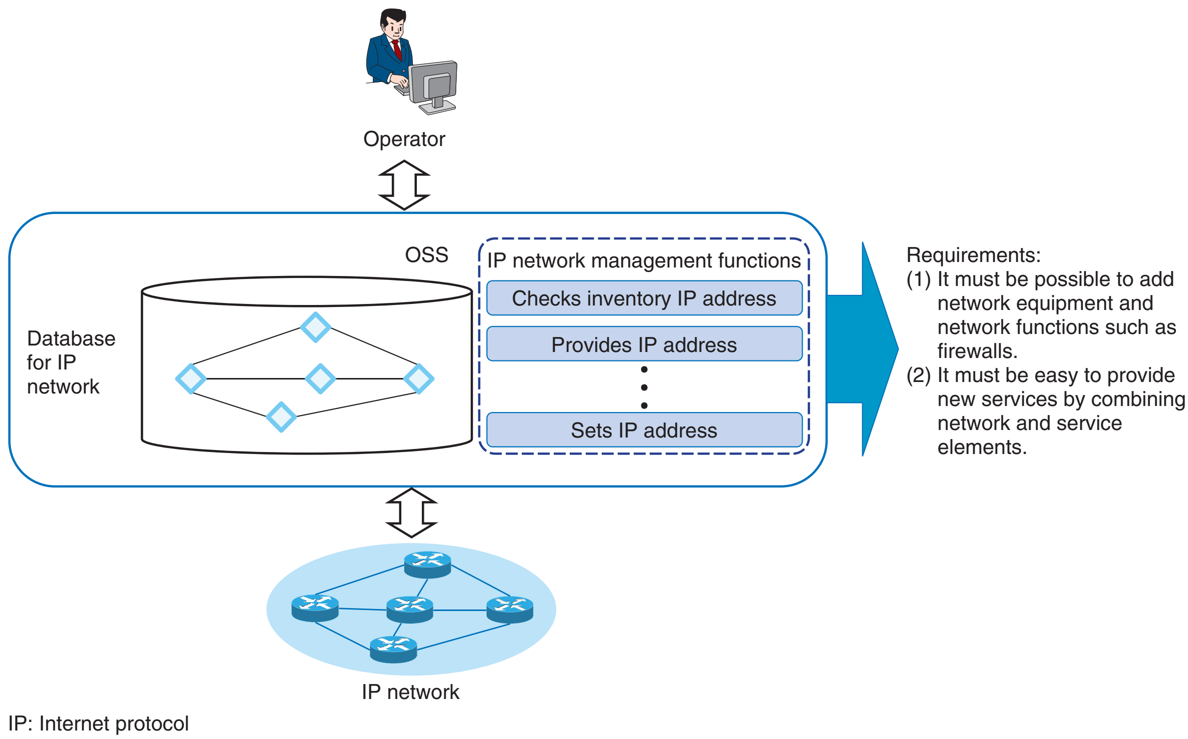


Fig. 1. Telecom operation and requirement for OSS.

Shared Information/Data Model (SID) has been discussed at TM Forum [2]. However, simply adopting a uniform information model will not completely eliminate the effects of adding new network equipment and changing management attributes on the OSS database and management functions. There is a need for a mechanism that can flexibly add equipment- and network-specific attributes to the OSS.

Next, with regard to requirement (2), it has been difficult with the conventional silo-type OSS to provide users with communications services in a highly convenient format that combines different services as needed. Consequently, to satisfy this requirement, it must be possible to uniformly manage information in the vertical direction from the physical equipment making up the network to the layers of communications protocols used on that equipment, and simultaneously, to manage information in the horizontal direction from the access network to the core network in an end-to-end manner.

2. Network resource management technology

In response to the issues described above, we have been researching a network resource management architecture that can provide flexible support for diverse types of networks [3]. This architecture is outlined in **Fig. 2**. It incorporates configuration-management functions based on generic management targets (entities) as a mechanism for constructing the database necessary for network management. It also features an external mechanism for prescribing individual equipment and communications-protocol characteristics as a *specification* separate from the OSS program. This information can be prescribed together with inter-layer relationships in the vertical direction to enable integrated management in both the vertical and horizontal directions. In this way, a mechanism is achieved for changing the operation of the OSS program based on injected specifications and the relationships between those specifications.

This architecture makes it possible to build a network management database independent of any particular network. It also possesses extendibility, enabling individual network characteristics to be appropriately expressed.

This technology enables a correspondence to be defined between SID logical resources used in this architecture as entities and individual pieces of network equipment used as network elements. The

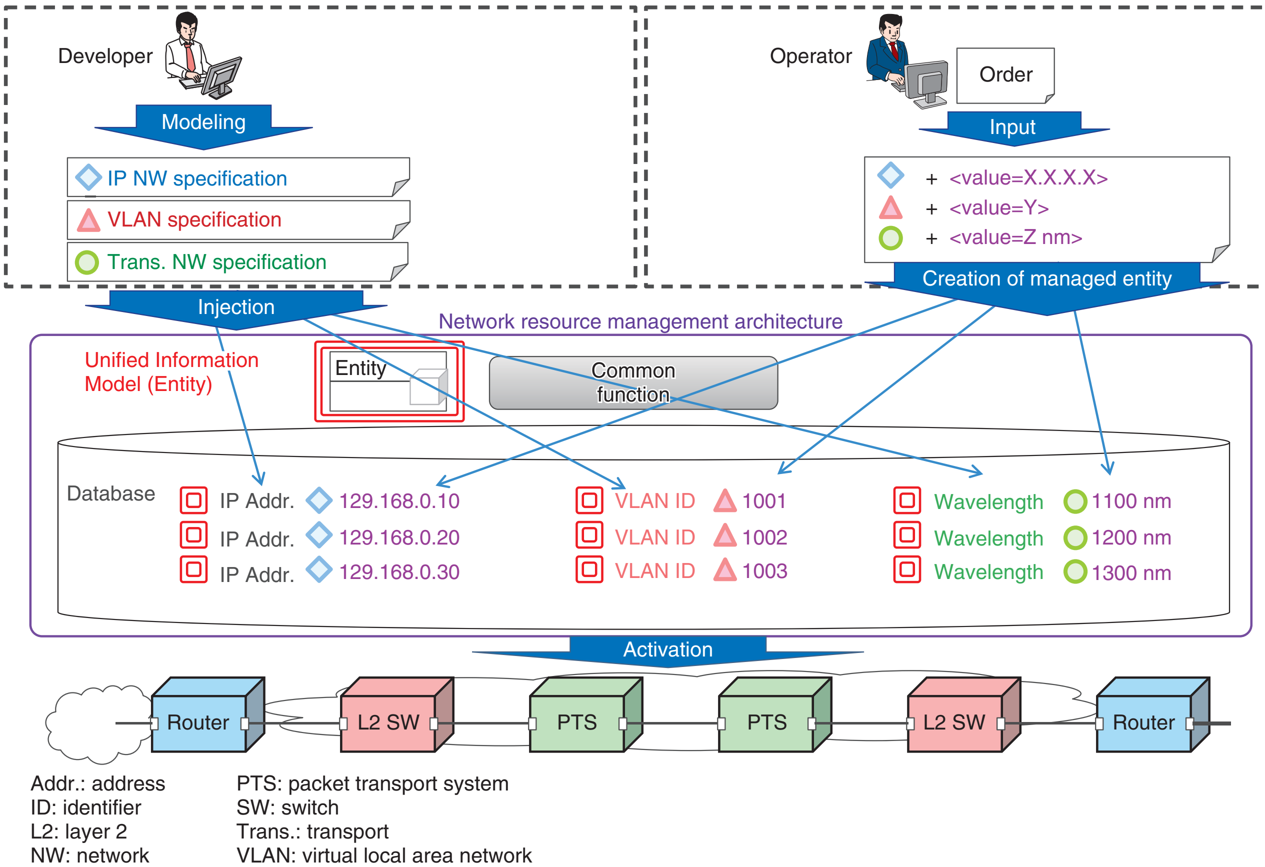


Fig. 2. Network resource management architecture.

information expressed here as SID-logical-resource entities complies with Recommendation ITU-T* G.800 (a compilation of uniform constructs such as definitions and signals for expressing the information transfer capability of a communications network), which means that transfer and coding functions for frames and packets in the equipment for each type of communications protocol are treated as concepts that can be converted into data. Furthermore, to achieve a protocol stack based on the OSI (Open Systems Interconnection) Reference Model, SID logical resources are prescribed in such a way that the concept of information transfer capability can be defined in a recursive manner. This makes it possible to express inter-layer relationships in the vertical direction.

Among SID logical resources, the following three entities are typical of those adopted by this architecture: Termination Point Encapsulation (TPE: a termination point on a communications protocol layer), Network Forwarding Domain (NFD: the domain expressing the connection relationship between TPEs and enabling information transfer on each layer), and Forwarding Relationship Encapsulation (FRE: information-transfer path generated on NFD).

The management information necessary for managing multilayer communications protocols can be expressed by combining these generic entities (**Fig. 3**).

This architecture includes a mechanism for adding characteristics of each communications protocol to the above generic entities and for storing protocol-specific information. The SID defines a Specification class that prescribes an entity specification and a SpecCharacteristic class that prescribes the attributes of that specification. It also defines a CharacteristicValue class with respect to an SID entity that prescribes attribute values expressing characteristics related to each protocol. In this architecture, using these Specification, SpecCharacteristic, and CharacteristicValue

* ITU-T: The Telecommunication Standardization Sector of the International Telecommunication Union.

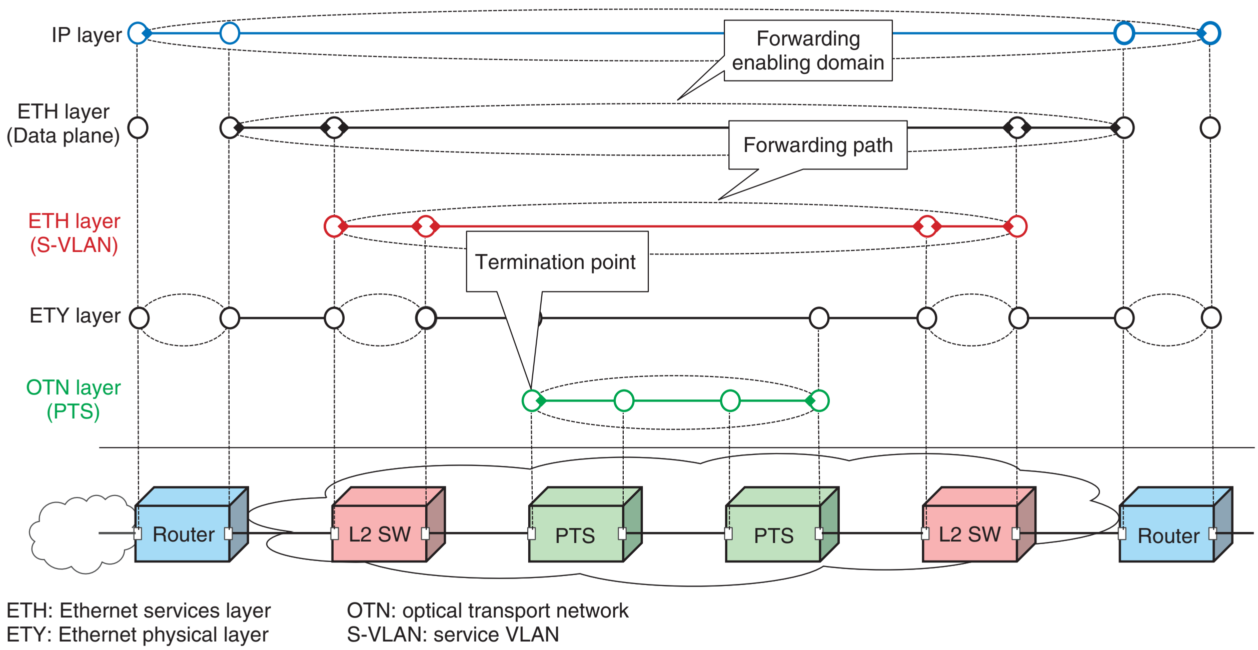


Fig. 3. Logical resource entities.

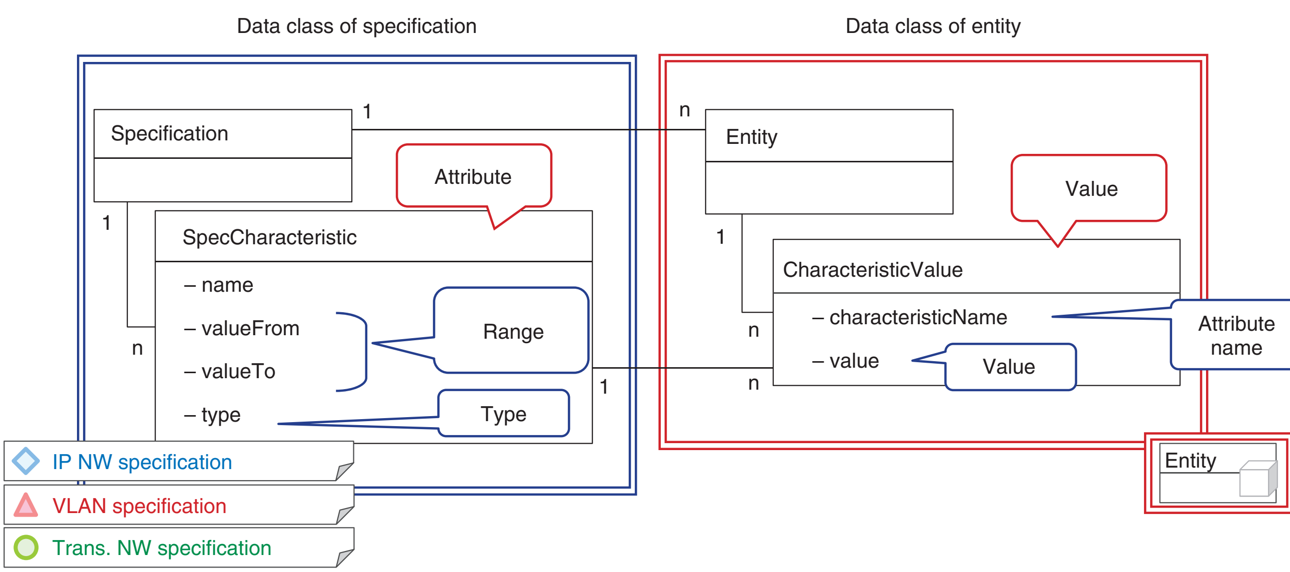


Fig. 4. Examples of entity and specification.

classes makes it possible to give common, generic entities characteristics that differ for each protocol, as shown in **Fig. 4**.

An example of implementing the architecture described above and generating the entities to manage an Internet protocol (IP) network is shown in **Fig. 5**. The process of developing an OSS service/network consists of defining the IP network information of an entity as a Spec for managing an IP network and registering that Spec in the proposed architecture. For example, to manage the IP address and subnet mask of a point (TPE) in an IP network, it would be

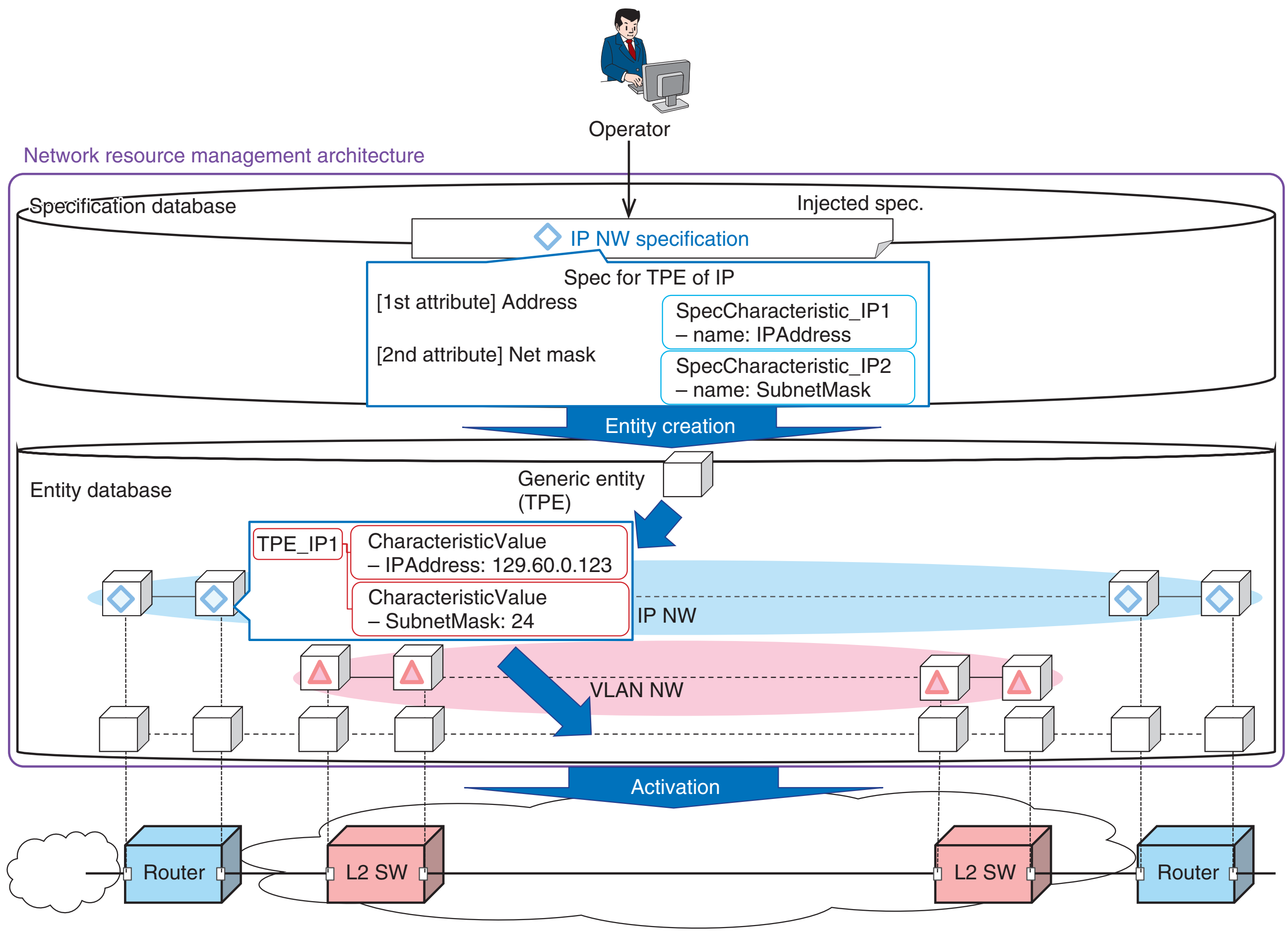


Fig. 5. Example of IP network.

necessary to define a Spec that adds the IP address (name) as the first attribute (SpecCharacteristic) and the subnet mask (name) as the second attribute (SpecCharacteristic). In addition, the range taken by the IP address or subnet mask is defined beforehand by SpecCharacteristic valueFrom and valueTo, while the format of the IP address or subnet mask value is likewise defined beforehand by the type of format it is. In the above way, SpecCharacteristic enables extendibility without having to directly give attribute information to a 'point (TPE) entity.' This is a key feature of the proposed architecture.

To manage an IP network when providing a service using an OSS, the service-providing operator instructs the system to generate a TPE entity in the network, applies a Spec to the TPE entity for use by the IP network, and generates data that associates an IP address (129.60.0.123) and subnet mask (24) with the entity as attributes.

Another example that shows the flow of generating an entity of an Ethernet virtual local area network (VLAN) in this architecture is shown in **Fig. 6**. To manage a VLAN in the process of developing a service/network, the service developer defines and registers a Spec for the entity configuring the VLAN network, similar to the flow in the IP network example given above. The Spec defines VLAN (name) as the first attribute (SpecCharacteristic) and the range allowed by VLAN ID as SpecCharacteristic valueFrom and valueTo. Defining Spec in this way enables a service-providing operator to generate TPE (TPE_VLAN1) for storing VLAN management information (VLAN: 1234) that associates a CharacteristicValue used by the VLAN network with 'point (TPE) entity' used in common by the IP network.

As shown by the above example, this architecture enables the development of an OSS capable of integrated management of multiple layers by preparing a

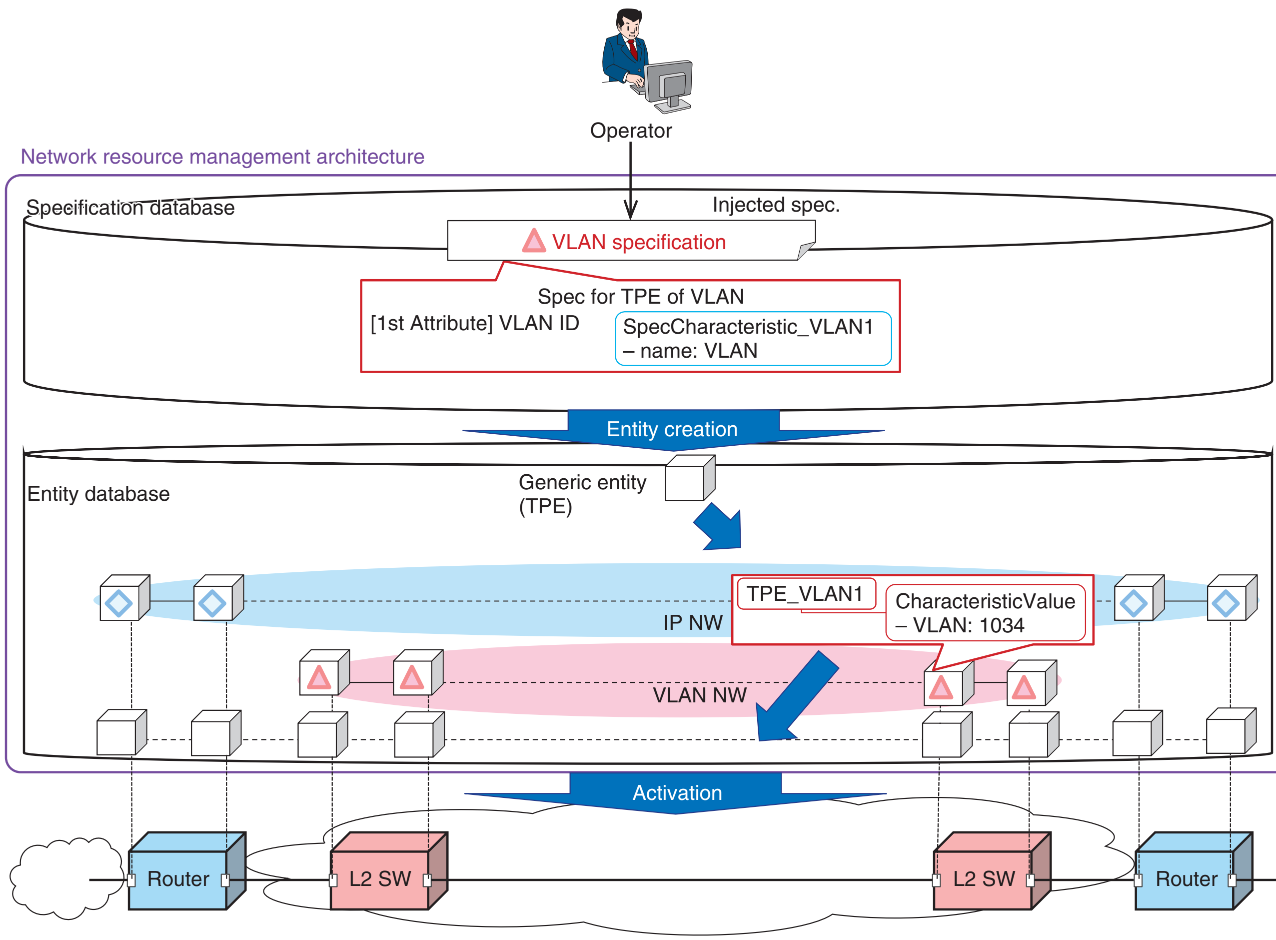


Fig. 6. Example of VLAN network.

Spec for each type of communications protocol and combining those Specs as needed.

3. Future plans

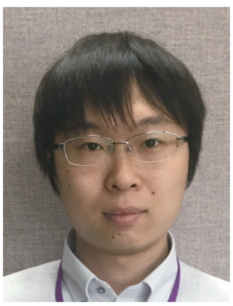
This article introduced the network resource management technology under development at NTT Access Network Service Systems Laboratories for comprehensively managing diverse types of networks. Our plan in the near future is to conduct trial operations using a small-scale commercial network to test the effectiveness and the merits of this technology.

References

[1] TM Forum site, http://www.tmforum.org/
[2] TM Forum Information Framework, https://www.tmforum.org/information-framework-sid/
[3] M. Sato, K. Akashi, S. Horiuchi, and T. Kotani, "Study of Variable Management Architecture for Diverse Networks," IEICE Tech. Rep., Vol. 116, No. 324, ICM2016-31, pp. 37–42, 2016.

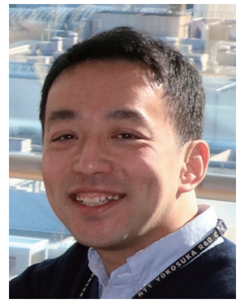
Shingo Horiuchi
Senior Research Engineer, Access Network Operation Project, NTT Access Network Service Systems Laboratories.

He received a B.E. and M.E. in engineering from the University of Tokyo in 1999 and 2001. He joined NTT Access Network Service Systems Laboratories in 2001. He has been researching and developing access network operation systems. He has also been involved in standardization efforts for operations support systems in TM Forum as a member of the Zero-touch Orchestration, Operations & Management (ZOOM) Project since 2014. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).



Kazuaki Akashi
Engineer, Access Network Operation Project, NTT Access Network Service Systems Laboratories.

He received a B.E. and M.E. in information and communication engineering from the University of Electro-Communications, Tokyo, in 2009 and 2011. He joined NTT Access Network Service Systems Laboratories in 2011. He has been researching and developing access network operation systems. He is a member of IEICE.



Masataka Sato
Senior Research Engineer, NTT Access Network Service Systems Laboratories.

He received a B.E. and M.E. in engineering from Tokyo Institute of Technology in 2002 and 2004. He joined NTT Access Network Service Systems Laboratories in 2004.

He has been researching and developing access network operation systems. He is a member of IEICE.

Tadashi Kotani
Senior Research Engineer, Supervisor, Access Network Operation Project, NTT Access Network Service Systems Laboratories.

He received a B.S. and M.E. in mechanical engineering from the University of Electro-Communications, Tokyo, in 1992 and 1994. Since joining NTT in 1994, he has mainly been engaged in researching and developing a network operations support system for the access network and wide area Ethernet network. He is a member of IEICE.

Practical Implementation of a New Clock Supply Module Supporting Telephone System Communications and Leased Line Communications for Corporate Customers

Takaaki Hisashima, Takeru Sakairi, Kaoru Arai, Hidetsugu Murayama, Osamu Kurokawa, and Katsutoshi Koda

Abstract

NTT has achieved high-accuracy frequency synchronization for communications using time division multiplexing, which is applied to telephony, leased lines, and other applications. Recently, the demand for high-accuracy time synchronization as well as high-accuracy frequency synchronization has increased, particularly in mobile systems. We introduce here the New Clock Supply Module (New CSM) enabling high-accuracy frequency and time synchronization on packet transport networks.

Keywords: clock supply module, time synchronization, phase information transmission

1. Introduction

Services such as fixed-line telephony and digital leased lines are currently provided on networks using time division multiplexing (TDM). TDM requires that the frequency of each system on the network be aligned with a high degree of accuracy for digital demultiplexing. At present, this is achieved by constructing a clock path network using a clock supply module (CSM), providing frequency synchronization over the entire network. This means that services requiring frequency synchronization must be continuously connected to a CSM and clock path network.

In existing clock path networks, the frequency information is transmitted using a synchronous digital hierarchy (SDH) transmission system between the master clock, which is a high-accuracy frequency source for the highest-hierarchy building, and a CSM installed in each building (SDH network), as shown in **Fig. 1**. Additionally, a redundant configuration using N (normal) and E (emergency) paths is used in order to achieve high reliability. If frequency information on the N path is disrupted, it will automatically be switched to the E path, providing stable frequency synchronization. Consequently, support for various transmission systems is essential in order to efficiently construct a clock path network with route distribution. Additionally, a migration to packet transport technologies on the transmission system also requires support in clock path networks.

Furthermore, in recent years, we have seen an increase in the number of services that require time synchronization of systems within a network, meaning an expansion in the application of network synchronization technologies. For example, the field of mobile communications requires time synchronization

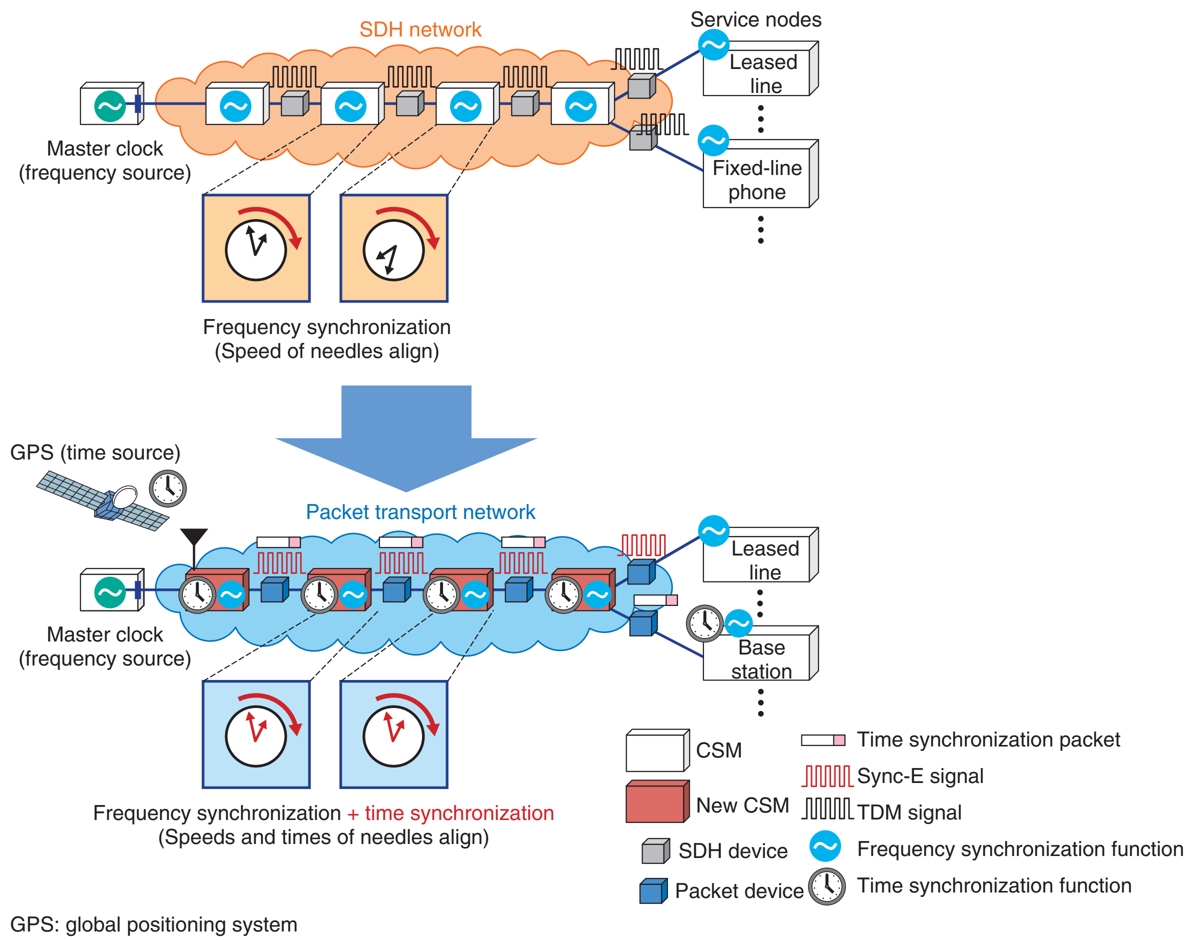


Fig. 1. Frequency/time synchronization using clock path network.

with a margin of error within the double-digit nanoseconds in order to provide support for broadband services using 5G (fifth-generation mobile communications networks) [1].

Against this background, we have developed a New Clock Supply Module (New CSM) that can transmit frequency information and phase information on the packet transport network and provide high-accuracy time synchronization.

The New CSM comprises four units—a base unit, an extended unit, a relay unit, and a phase adjustment unit—as shown in **Fig. 2**. The base unit is synchronized with the master clock (frequency source), GPS* (time source), and the upper CSM, transmitting frequency information and time information. The extended unit extends the capacity of service nodes such as switching equipment and leased line equipment. The relay unit extends the distance over which frequency information and time information can be transmitted. The phase adjustment unit makes it possible to efficiently replace an existing CSM with the New CSM. The distinctive technologies in these units provide phase information transmission, high-accuracy time synchronization, and automatic replacement functions. We explain these in more detail in the following sections.

2. Issues

First, we describe some issues that need to be resolved regarding the technologies mentioned in the previous section.

2.1 Necessity of phase information transfer in addition to frequency information

Methods of transmitting frequency information on

* GPS: global positioning system

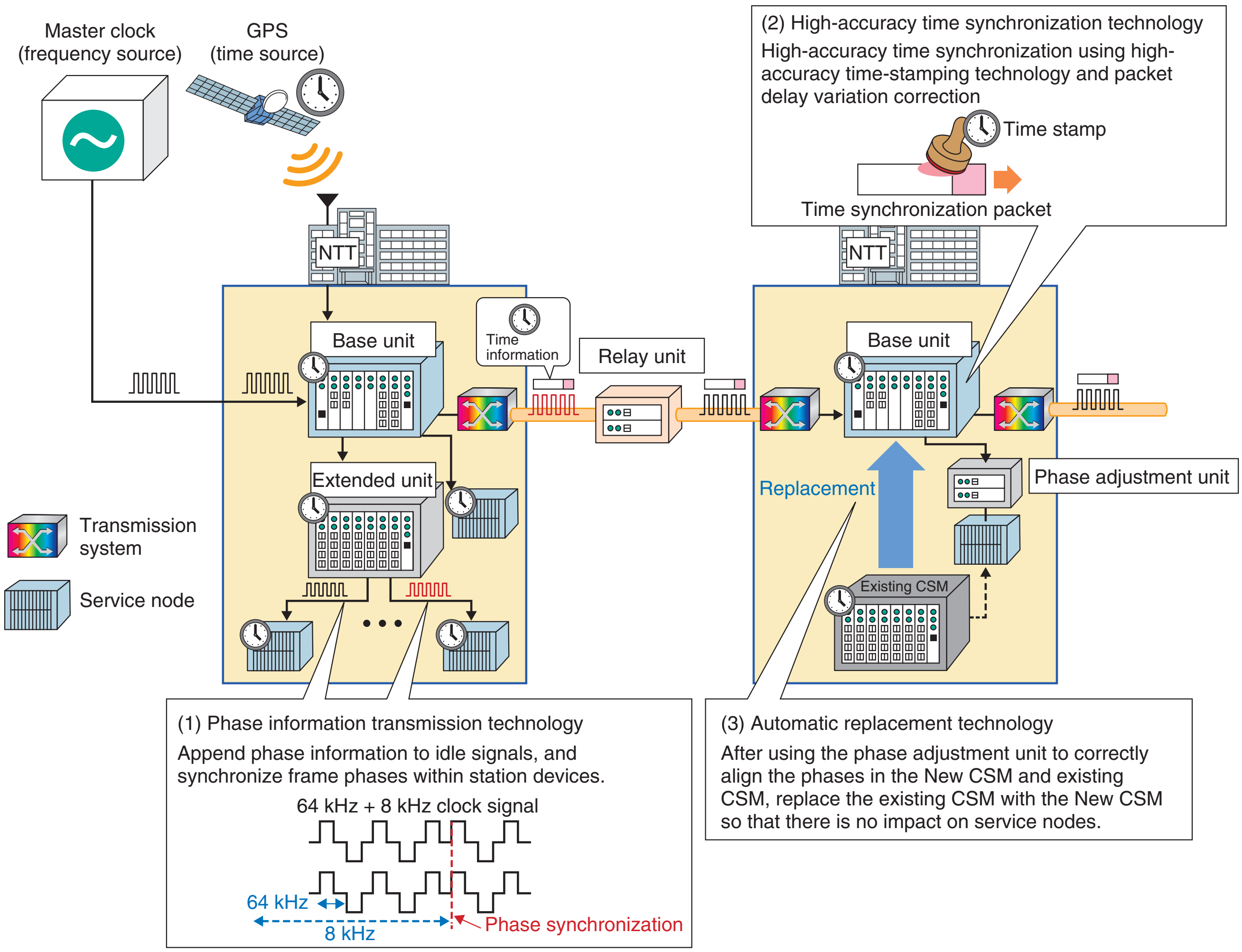


Fig. 2. New CSM configuration.

packet transport networks are standardized as Synchronous Ethernet (Sync-E) technology [2–4]. Sync-E technology enables transmission of frequency information without the effects of electrical noise. However, when some devices communicate with other devices in the same building, some service nodes that receive frequency information require that frame header timing be aligned. Therefore, it is necessary to synchronize phases as well as frequencies. In packet transport networks, when phase information is transmitted using packets, the effects of other packets will generate time fluctuations in transmission and reception cycles, preventing high-accuracy phase information transmission.

2.2 Support for high-accuracy time synchronization

Time synchronization methods are standardized as Precision Time Protocol (PTP), a time synchronization protocol that can be used on packet transport and other networks [5]. PTP provides bidirectional exchange of time synchronization messages in networks consisting of a master node and a slave node, and it corrects the time offset of a slave node to the master node. Time synchronization messages store time-stamp information detailing transmission/reception times, and PTP uses this time-stamp information to calculate the time offset. Time-stamp information is assigned using hardware processing within the devices. However, there may be an error between the original time-stamp information and the assigned time depending on the hardware processing performance.

This error in the time-stamp information has a direct impact on the accuracy of time synchronization and therefore needs to be reduced.

Additionally, as one of the synchronization conditions, PTP must presuppose an equivalent time delay in both directions. Accordingly, utilizing PTP in a configuration in which master and slave nodes are connected using transmission systems may adversely affect the accuracy of time synchronization. This is because random time delay variations may occur as a result of packet processing conflicts within transmission systems. Accordingly, high-accuracy time synchronization requires a reduction of the time error resulting from delay variations.

2.3 Rapid replacement of existing CSM with New CSM

Construction of a clock path network using the New CSM requires replacement of the existing CSM. If this process is supported only by changing the connection of the clock transmission cable, the frequency synchronization can be achieved between the existing CSM and the New CSM. However, since it is not possible to accurately match the phase information before and after the replacement, the phase jumps at the service node receiving the phase information, which may affect the quality of service. When replacing an existing clock system of the previous generation with a current CSM, operators carry out phase alignment manually by monitoring waveforms on an oscilloscope. However, a quicker, more accurate way to align phases is required when replacing several thousand existing CSMs with the New CSM.

3. Technical points in resolving issues

Here, we discuss the technical requirements to resolve the above issues.

3.1 Phase information transmission technology

In packet networks, transmission of frame signals requires that noncontiguous idle signals be sent whenever frame signals are not sent. Idle signals can be transmitted without being dependent on the transmission/reception status of packets. Therefore, the New CSM focuses on these idle signals. Using a phase information transmission method that replaces some idle signals with other idle signals (phase idle signals) that do not violate rules enables the transmission of accurate phase information.

3.2 High-accuracy time synchronization technology

To resolve the problematic time errors caused by the time-stamping error and the delay variation from linked system devices, the New CSM implements a time-stamping method using a Vernier scale and a packet delay variation correction method through the accumulation of statistical distribution data to achieve high-accuracy time synchronization.

(1) Time-stamping method using Vernier scale

The time-stamping error is affected by the internal clock frequency of the device. For example, if the internal clock frequency is 100 MHz, a time-stamping error of a maximum of 10 ns may be generated. Since there is a physical limit to increasing the internal clock frequency, it is inevitable that a certain amount of time-stamping error will occur. In response to this problem, the New CSM drastically reduces the time-stamping error by using two internal clock frequencies. This is a method called the Vernier scale, in which two internal clocks are used as a main scale and subscale. Altering the interval between these scales enables values to be read with high resolution (**Fig. 3**). With this method, time-stamping errors can be reduced to within 0.5 ns.

(2) Packet delay variation correction method through accumulation of statistical distribution data

Reducing time synchronization errors resulting from delay variations generated in a transmission device requires correction of the reception time of time synchronization packets, so that delays in both directions are equivalent. Averaging the reception time of time synchronization packets has been the conventional way to correct the reception time. This method is effective if there is a constant average delay in both directions. However, because the traffic of user packets changes from moment to moment, the average value of the bidirectional delay time does not necessarily match. Accordingly, the New CSM measures the actual delays in multiple delay measurement packets. Then it statistically ascertains the delay variation characteristics, enabling correction of delay variations in time synchronization packets (**Fig. 4**).

3.3 Automatic replacement technology from existing CSM to New CSM

To achieve accurate phase matching between the existing CSM and New CSM in a short time period during the replacement process, the New CSM has an automatic phase adjustment function and performs appropriate phase difference correction by running

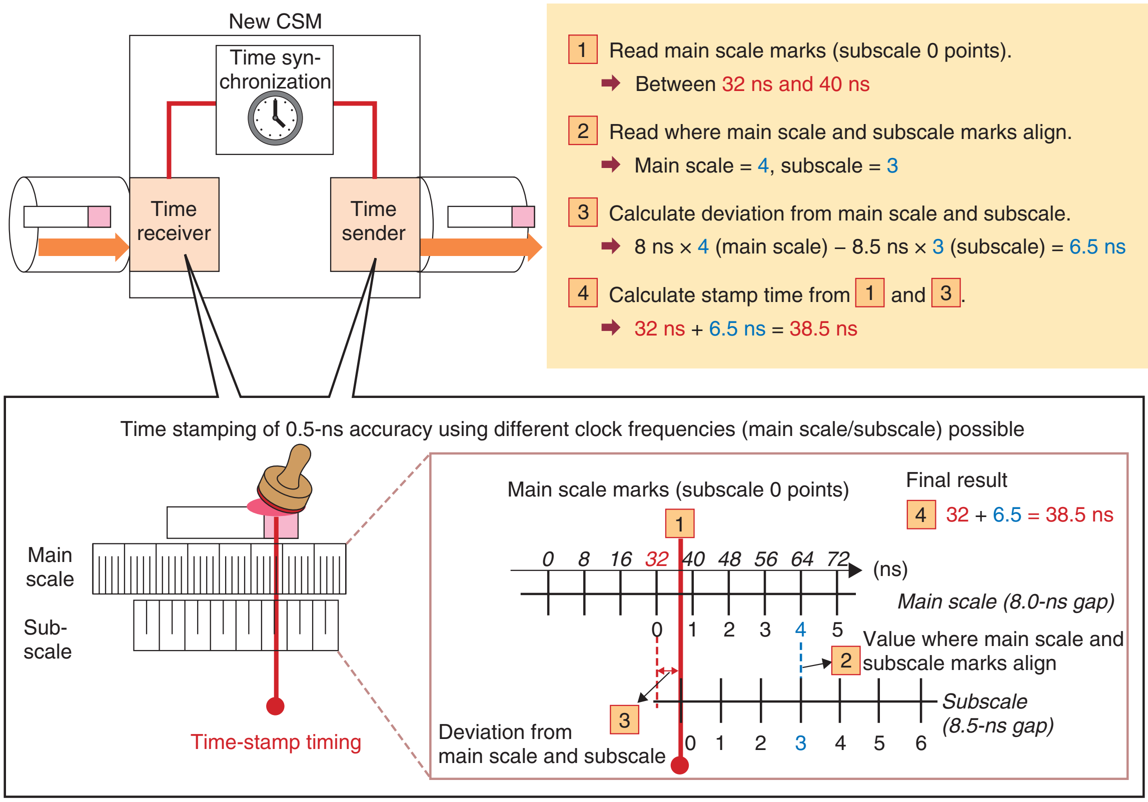


Fig. 3. Time stamping using Vernier scale.

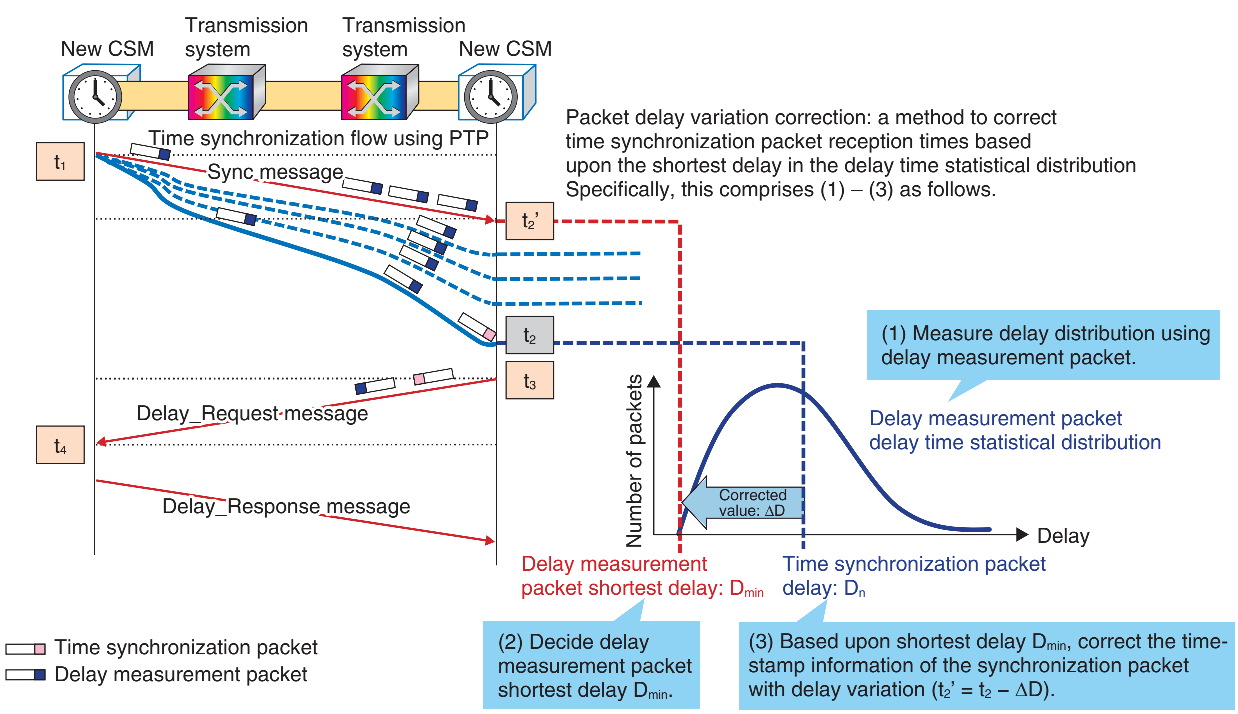


Fig. 4. Packet delay variation correction method.

algorithms using parameters such as the waveform characteristics of New CSM and the current CSM, the connection cable length, and the amount of internal delay. Additionally, the accuracy of the work is increased during the parameter input by simplifying the setting work, visualizing the setting procedures, and installing control applications that enable warning notifications in the event of problems. With this function, it is possible for the operator to achieve phase alignment on the order of nanoseconds simply by inputting the necessary parameters.

4. Future prospects

We introduced the New CSM developed by NTT Network Service Systems Laboratories that includes technologies for phase information transmission, high-accuracy time synchronization, and automatic replacement of CSM. Going forward, we plan to offer trouble-free, ongoing assistance targeting the introduction and operation of the New CSM, and we will carry out research and development to meet higher accuracy requirements.

References

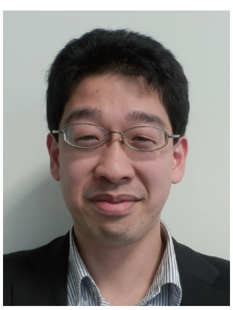
[1] ITU-T (International Telecommunication Union - Telecommunication Standardization Sector) Rec. G.8271/Y.1366: "Time and Phase Synchronization Aspects in Packet Networks," 2012.

[2] ITU-T Rec. G.8261/Y.1361: "Timing and Synchronization Aspects in Packet Networks," 2008.

[3] ITU-T Rec. G.8262/Y.1362: "Distribution of Timing Information through Packet Networks," 2007.

[4] ITU-T Rec. G.8264/Y.1364: "Timing Characteristics of Synchronous Ethernet Equipment Slave Clock (EEC)," 2008.

[5] IEEE Std. 1588: "IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems," 2008.



Takaaki Hisashima

Research Engineer, Transport Network Innovation Project, NTT Network Service Systems Laboratories.

He received a B.S. and M.S. in chemical sciences from Osaka Prefecture University in 2006 and 2008. He joined NTT WEST in 2008, where he worked on the development of a security system at a research and development (R&D) center at NTT WEST. He has been in his current position since July 2014 and is engaged in developing network systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE).

Takeru Sakairi

Researcher, Transport Network Innovation Project, NTT Network Service Systems Laboratories.

He received a B.S. in engineering from Tokyo University of Agriculture and Technology in 2008, and an M.S. in engineering from Tokyo Institute of Technology in 2010. He has been a researcher at NTT Network Service Systems Laboratories since 2010. He is engaged in R&D of a packet transport multiplexing adaptor for a dedicated service and clock supply module.

Kaoru Arai

Researcher, Transport Network Innovation Project, NTT Network Service Systems Laboratories.

He received a B.S. and M.S. in applied physics from Tokyo University of Science in 2010 and 2012. Since joining NTT Network Service Systems Laboratories in 2012, he has been involved in R&D of packet network systems such as clock supply systems. He also contributed to standardization efforts in ITU-T Study Group 15. He is a member of IEICE.

Hidetsugu Murayama

Senior Research Engineer, Transport Network Innovation Project, NTT Network Service Systems Laboratories.

He received a B.E. and M.E. in photonics from Keio University, Kanagawa, in 2002 and 2004. He joined NTT in 2004 and was involved in R&D on Internet protocol networks and dedicated service networks at NTT and NTT EAST. He has been in his current position since July 2016 and is involved in R&D of network synchronization systems.

Osamu Kurokawa

Senior Research Engineer, Transport Network Innovation Project, NTT Network Service Systems Laboratories.

He has been in his current position since August 2014 and is engaged in R&D of a packet transport multiplexing adapter for dedicated services, an Ethernet private line system, and a clock supply module.

Katsutoshi Koda

Director, Transport Network Innovation Project, NTT Network Service Systems Laboratories.

He received a B.S. and M.S. in mechanical sciences from Tokyo University of Science in 1987 and 1989. He joined NTT in 1989 and began working on the development of transport systems at the Network System Development Center. He has been in his current position since July 2014, where he is managing the development project for the Ethernet private line system. He also manages the R&D of future optical networking and transport network management technology.

Standardization Trends for Future High-speed Passive Optical Networks

Ryo Koma, Jun-ichi Kani, Kota Asaka, and Ken-Ichi Suzuki

Abstract

Standardization organizations are conducting studies on optical networks with higher-than-ever speeds as part of efforts to continue the development of access networks. This article briefly reviews the general concept of the latest high-speed passive optical network system called NG-PON2 (Next-Generation Passive Optical Network Stage 2) and introduces current standardization trends for future high-speed optical access systems.

Keywords: passive optical network, NG-PON2, 100G-EPON

1. Standardization of optical access systems

Passive optical networks (PONs) are being widely utilized to provide optical broadband services to many users in a cost-effective manner. They consist of a point-to-multipoint topology in which an optical line terminal (OLT) is placed at the central office and connected to multiple optical network units (ONUs) placed at each user's home through a feeder fiber and an optical splitter. Current PON systems employ time division multiplexing and time division multiple access (TDM/TDMA) because the OLT and part of the feeder fiber can be shared by many users. In such PONs, each ONU communicates with the OLT in different time slots assigned by the OLT.

Standardization of 1-Gbit/s to 10-Gbit/s-class PONs has already been completed by both the Institute of Electrical and Electronics Engineers (IEEE) and the International Telecommunication Union - Telecommunication Standardization Sector (ITU-T). Likewise, IEEE has completed the standardization of GE-PON (Gigabit Ethernet PON) and 10 Gigabit Ethernet PON (10G-EPON), and ITU-T has done so for Gigabit-capable PON (G-PON), 10 Gigabit-capable PON (XG-PON), and 10 Gigabit-capable symmetric PON (XGS-PON).

ITU-T has also standardized the 40-Gbit/s-class PON system called the Next-Generation Passive Optical Network Stage 2 (NG-PON2) for a further increase in total capacity. The basic approach of NG-PON2 is time and wavelength division multiplexing (TWDM)-PON, which is a hybrid of conventional TDM/TDMA and wavelength division multiplexing (WDM) technologies and can increase the aggregated PON rate beyond 10 Gbit/s. This article briefly reviews the basic concept of NG-PON2—the latest high-speed PON system—and introduces the current standardization trends for future high-speed optical access systems being discussed in IEEE, ITU-T, and the Full Service Access Network (FSAN) Group, which has contributed to the standardization of G-PON, XG(S)-PON, and NG-PON2.

2. NG-PON2

An example of the 40-Gbit/s-class optical access system called NG-PON2 is shown in **Fig. 1**. Standardization of this system was completed in July 2015 by ITU-T. While previous PON systems mainly offer broadband services for residential users, NG-PON2 systems are expected to accommodate both business and residential users; they will also accommodate mobile antennas.

As mentioned in the previous section, TWDM-PON is adopted as the primary technology. NG-PON2 supports PtP (point-to-point) WDM overlay, in

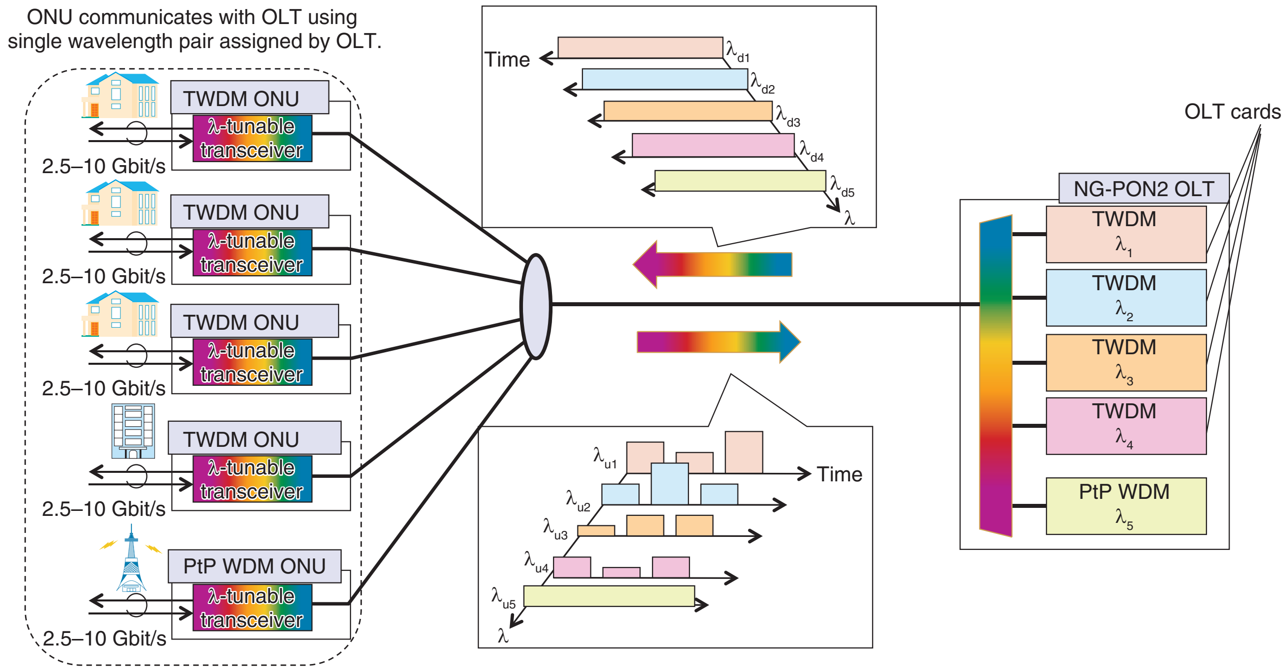


Fig. 1. NG-PON2 system overview.

which each ONU communicates with the OLT using an assigned wavelength channel that is occupied by the ONU. There are four, optionally eight, multiplexed wavelengths for both upstream and downstream communications. In an NG-PON2 network, there are three line rates (per wavelength): symmetrical 10 Gbit/s, 2.5 Gbit/s (upstream) and 10 Gbit/s (downstream), and symmetrical 2.5 Gbit/s. When symmetrical 10 Gbit/s is utilized, the NG-PON2 system offers a symmetrical maximum transmission capacity of 40 Gbit/s. The maximum split ratio and transmission distance are specified as 1:256 and 40 km, respectively.

One of the main advances of the NG-PON2 system is its use of wavelength tuning in the ONU. In-service ONU wavelength tuning enables the NG-PON2 system to offer advanced network functions. For example, dynamic wavelength assignment offers ONU traffic load balancing by adjusting the numbers of ONUs assigned to each wavelength, as well as protection of OLT cards, in which the malfunction of an OLT card triggers immediate ONU wavelength reassignment.

3. IEEE 100G-EPON

To elucidate the latest trends in high-speed PON systems that offer a total capacity over 40 Gbit/s, we introduce the standardization activities of IEEE for a 100-Gbit/s-class Ethernet PON system. In December 2015, IEEE confirmed the establishment of the IEEE 802.3ca Task Force (TF), which started work on standardizing 100G-EPON with the aim of completing it by 2020 [1].

An example of the 100G-EPON system under discussion in 802.3ca TF is shown in **Fig. 2**. It is assumed that 100G-EPON will adopt WDM technology, just as with NG-PON2. The consensus for the basic 100G-EPON is that the line rate per wavelength and the maximum number of multiplexed wavelengths will be 25 Gbit/s and four wavelengths, respectively. In addition, it has also been decided that non-return-to-zero modulation (used in previous PON systems) will be adopted at the transmitter.

To offer an incremental upgrade of maximum transmission capacity from 25 Gbit/s to 100 Gbit/s for downstream and 10 Gbit/s to 100 Gbit/s for upstream, seven types of ONUs are expected to be used in the 100G-EPON system: 25/10G-ONU, 25/25G-ONU, 50/25G-ONU, 50/50G-ONU, 100/25G-ONU, 100G/50G-ONU, and 100/100G-ONU. Additionally, IEEE 802.3ca TF will begin studying optional configurations of 50G-EPON and 100G-EPON, where the line rate per wavelength is 50 Gbit/s, and the

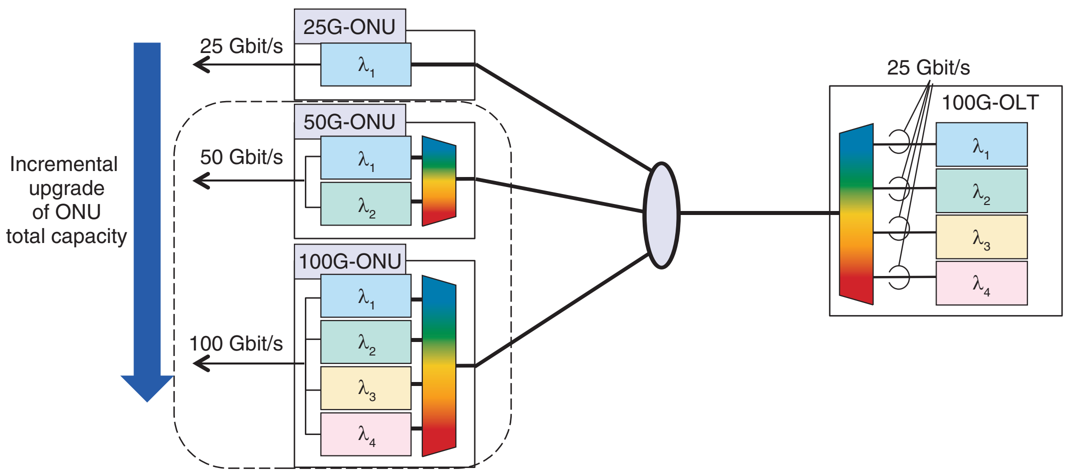


Fig. 2. 100G-EPON system overview.

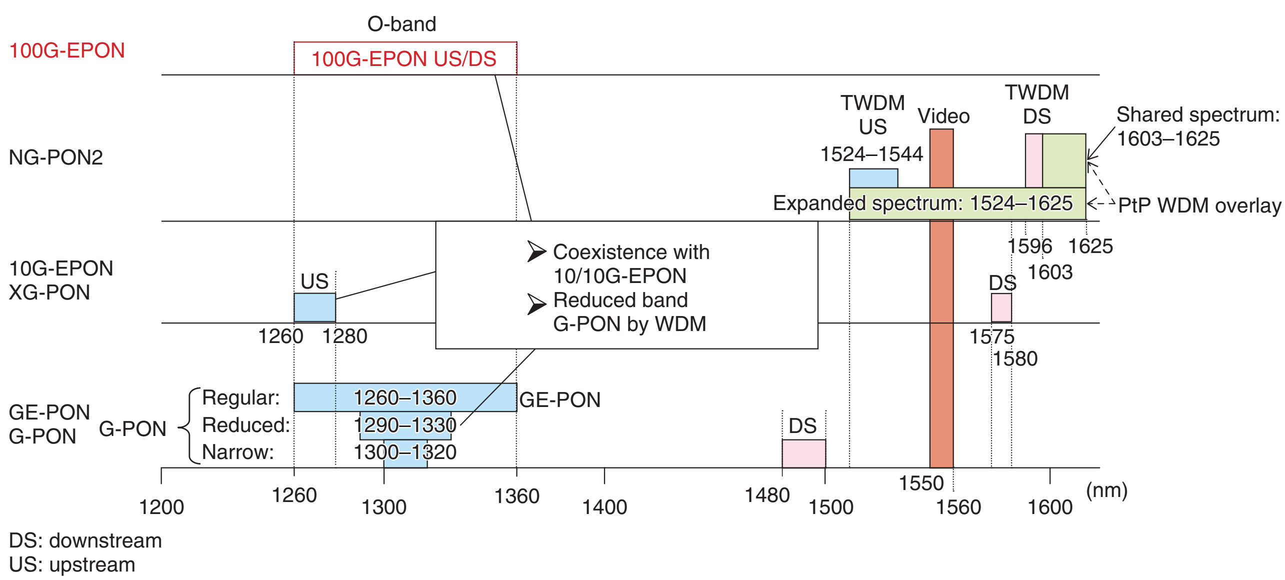
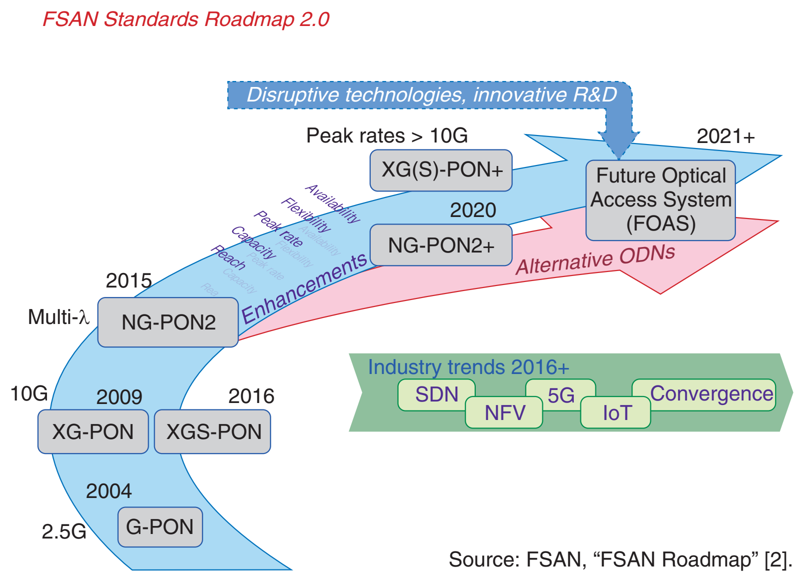


Fig. 3. A wavelength plan for 100G-EPON.

maximum number of multiplexed wavelengths is two.

A conceptual image of the incremental ONU upgrade in 100G-EPON downstream is also shown in Fig. 2. Three types of ONUs are illustrated: a 25G-ONU, 50G-ONU, and 100G-ONU. Each ONU receives downstream signals using assigned wavelength channels (the line rate per wavelength is 25 Gbit/s); 25G-ONU receives λ_1, 50G-ONU receives both λ_1 and λ_2, and 100G-ONU receives all wavelength channels $\lambda_{1\text{–}4}$. A total capacity of 50 Gbit/s and 100 Gbit/s is respectively achieved by channel bonding the two and four wavelengths received.

A 100G-EPON wavelength plan assumed by IEEE 802.3ca TF is shown in **Fig. 3**. For comparison, wavebands of current PON systems are also summarized. Since increasing the signal bandwidth makes optical fiber dispersion much worse, the use of the O-band (1260 nm to 1360 nm, around the zero dispersion wavelength) for both upstream and downstream has been agreed on to suppress the penalty of signal waveform distortion caused by optical fiber dispersion.



Source: FSAN, "FSAN Roadmap" [2].

5G: fifth-generation mobile communications network
NFV: network functions virtualization
IoT: Internet of Things
ODN: optical distribution network
SDN: software-defined networking

Fig. 4. FSAN standards roadmap.

Note that the waveband is also assigned to conventional 1-Gbit/s and 10-Gbit/s-class PON systems. For the present, a precise wavelength plan for 100G-EPON is still under study in IEEE 802.3ca TF. However, it has been agreed that 100G-EPON will co-exist with 10G-EPON and G-PON using a set of reduced wavelengths.

There are two main options to offer co-existence with the use of WDM; one option is for the first 25G-ONU wavelength channel to use the waveband at 1310 nm with a 20-nm width and to co-exist with 10G-EPON. The other option is for the first 25G-ONU wavelength channel to use a 20-nm-wide waveband at 1270 nm and to co-exist with G-PON using a set of reduced wavelengths. IEEE 802.3ca TF is aiming to conclude the IEEE standard 802.3ca by April 2020 and will focus on discussing details of the specifications for the physical layer, ONU control plane, and transmission convergence layer in the near future.

4. Further enhancement of XGS-PON and NG-PON2 by FSAN/ITU-T

FSAN and ITU-T recently initiated a study on enhancing the capability of 10-Gbit/s per wavelength class PON systems such as NG-PON2 and XGS-PON. Other targets include further improvement in the line rate per wavelength beyond 10 Gbit/s as one of the study options. FSAN released a standardization roadmap in November 2016, shown in **Fig. 4** [2]. This roadmap predicts that the future optical access system will have to offer not only higher maximum transmission capacity but also several new attributes such as long reach and high splitting ratios, system reconfigurability and flexibility, and high availability of network equipment. The emergence of high-speed mobile services such as 5G (fifth-generation mobile communications network) and the development of software-defined networking and network functions virtualization technologies are also expected to push these trends.

At the same time, ITU-T Study Group 15 (SG15) has commenced work on "G.sup.HSP: G. Supplement High-Speed PON," a technical supplement on the current technologies available for improving the PON system line rate per wavelength to beyond 10 Gbit/s [3]. FSAN member companies are offering their contributions to ITU-T SG15 to promote the development of the future optical access system.

5. Summary

This article reviewed the latest standardization trends to achieve further increases in the bandwidth of PON systems. IEEE, FSAN/ITU-T, and the Broadband Forum have initiated studies into PON convergence in order to prevent conflicts in the standardization activities for high-speed PON systems. We expect these activities to lead to the further evolution of optical access systems and sustain the continuous development of broadband services.

References

[1] IEEE P802.3ca 100G-EPON Task Force, http://www.ieee802.org/3/ca/

[2] FSAN, "FSAN Roadmap," http://www.fsan.org/roadmap/

[3] ITU-T, G.sup.HSP, http://www.itu.int/itu-t/workprog/wp_item.aspx?isn=13463

Ryo Koma

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

He received a B.E. from Tokyo University of Science, Chiba, in 2010, and an M.S. from the University of Tokyo in 2012. In 2012, he joined NTT Access Network Service Systems Laboratories, where he has been researching digital signal processing based PON systems.

Jun-ichi Kani

Senior Research Engineer, Supervisor (Distinguished Researcher), Group Leader, Optical Access System Project, NTT Access Network Service Systems Laboratories.

He received an M.E. and Ph.D. in applied physics from Waseda University, Tokyo, in 1996 and 2005. He joined NTT in 1996. He has been with NTT Access Network Service Systems Laboratories since 2003, where he has been engaged in research, development, and standardization of optical communications systems for metro and access applications.

Kota Asaka

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

He received a B.S. and M.S. in electrical engineering from Waseda University, Tokyo, in 1996 and 1999, and a Ph.D. in physics from Kitasato University, Tokyo, in 2008. In 1999, he joined NTT Photonics Laboratories, where he conducted research on photonics integrated circuits and long-wavelength VCSELs (vertical cavity surface emitting lasers). From 2005–2008, he researched optical coherence tomography by using SSG-DBR-LDs (super-structure-grating distributed Bragg reflector laser diodes) at Kitasato University. From 2009–2012, he worked on developing low-cost and small optical subassemblies for optical access networks. He has been with NTT Access Network Service Systems Laboratories since 2012. He served as the working group secretary of TC86/SC86C/WG4 of the International Electrotechnical Commission, and he has been involved in international standardization efforts for next-generation optical access systems in ITU-T SG15/Q2 and the FSAN Group since 2013. Dr. Asaka is a member of IEEE Communications Society and the Institute of Electronics, Information and Communication Engineers (IEICE).

Ken-Ichi Suzuki

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

He received a B.E. and M.E. in electronic engineering from Utsunomiya University, Tochigi, in 1988 and 1990, and a Ph.D. in information science and technology from Hokkaido University in 2009. He joined NTT in 1990, where he has been working on the research and development of optical communications systems. He currently leads the Access Systems Technology Group that is investigating optical access related services/technologies including FASA (Flexible Access System Architecture) developed by NTT. He has been a vice chair of IEEE 1904 Access Networks Working Group and a director of the Optical Access Test Implementation Liaison Committee at the HATS (Harmonization of Advanced Telecommunication Systems) conference of Japan. Dr. Suzuki is a member of OSA (The Optical Society), IEEE, and IEICE.

Failure Analysis of Internet Protocol Terminal Equipment

Abstract

This article describes a fault that occurred in the transmission of voice signals in a customer's Internet protocol telecommunication system and the analysis carried out to identify the problem. This is the forty-second article in a series on telecommunication technologies. This contribution is from the Network Interface Engineering Group, Technical Assistance and Support Center, Maintenance and Service Operations Department, Network Business Headquarters, NTT EAST.

Keywords: IP phone, PBX, VoIP gateway

1. Introduction

The Network Interface Engineering Group investigates the causes of puzzling and troublesome phenomena in Internet protocol (IP) network services such as the FLET'S HIKARI NEXT fiber-optic broadband service and also considers countermeasures. We do this by capturing and collecting IP packets in the Ethernet section dividing terminal equipment and telecommunication equipment and analyzing the data in those packets.

A typical customer premises network is configured with a voice over Internet protocol (VoIP) gateway for converting IP signals to analog or integrated services digital network (ISDN) signals for accommodation by terminal equipment. If a problem should occur in voice or facsimile (fax) communications on these circuits, it becomes necessary to analyze the data collected by measurement equipment.

We introduce here a case study in which we applied our analysis methods to the occasional occurrence of a one-way-voice problem in extension calls between two hubs on a private branch exchange (PBX)-accommodated extension telephone. A one-way-voice problem is when a voice signal is transmitted in only one direction.

2. Case study overview and investigation method

The customer had installed a PBX in hub A and a VoIP gateway in hub B and was making extension calls between these hubs via the FLET'S VPN WIDE virtual private network service. In this configuration, a one-way-voice problem was periodically occurring in extension calls between these two hubs.

The extension telephone on which the one-way-voice problem occurred was not a specific telephone, and it is known that the problem did not reoccur after remaking the call directly after the problem occurrence. The measures usually taken by maintenance personnel in the field to identify problems were to exchange telephones, exchange the PBX package, or collect data through packet capture at the time of problem occurrence and perform primary analysis. These measures, however, failed to find the cause of the problem, so the Technical Assistance and Support Center was called upon to conduct an investigation.

To collect and analyze data at the time of problem occurrence, we captured packets in the Ethernet section at the point marked with "●" in **Fig. 1**.

3. Analysis of collected data

The results of analyzing Real-time Transport Protocol (RTP) packets at the time of problem occurrence are shown in **Fig. 2**.

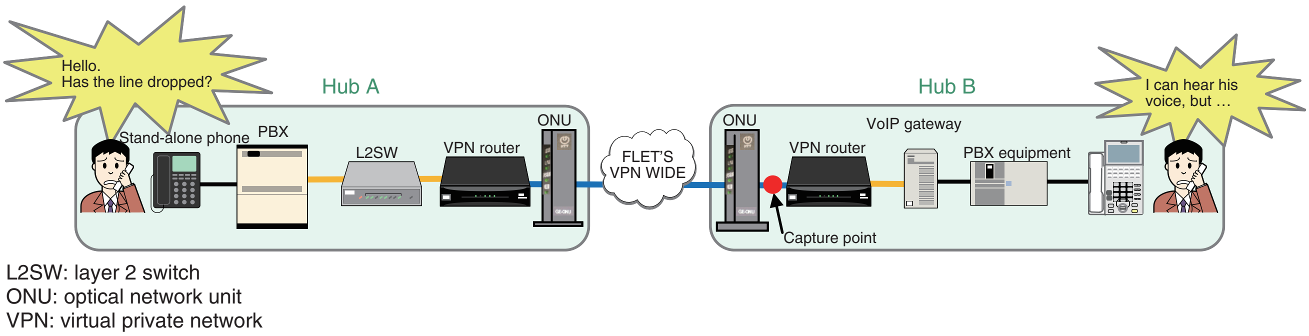


Fig. 1. Connection configuration and data collection point.

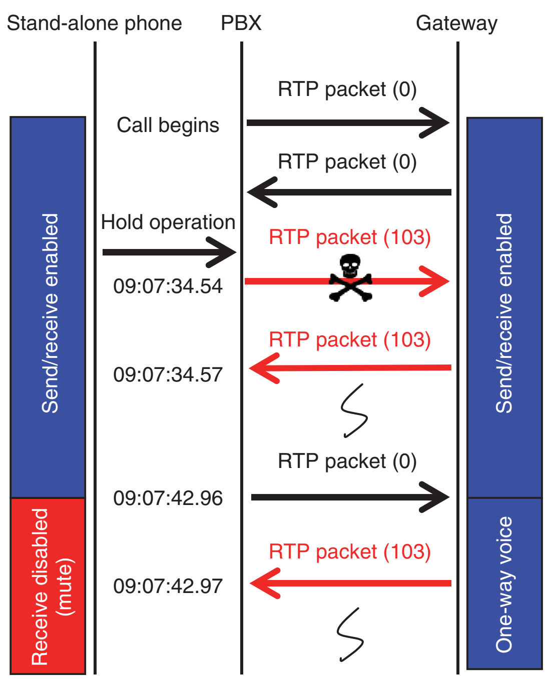


Fig. 2. Results of RTP packet analysis.

3.1 Analysis results 1: Change in payload type

The analysis of captured data revealed that the payload type (PT)*1 of RTP packets suddenly switched to "103" (Dynamic RTP-103) at the time of problem occurrence.

The customer's system applies PT "0" indicating G.711PCMU—a pulse code modulation standard for voice communications—to voice calls, but at the time of problem occurrence in a certain voice call, the PT of RTP packets in both directions suddenly switched to "103" beginning at 9:07:34. Later, however, the PT of RTP packets only in the direction from hub A to the VoIP gateway returned to "0," at which time the one-way-voice problem occurred.

To identify the purpose of PT "103," we checked with personnel in charge of PBX development and learned that it is a PT used in fax communications. We were told that the PBX switches the PT to "103" on receiving a fax signal from a lower-level terminal.

3.2 Analysis results 2: Frequency components included in hold tone

We decoded the captured data collected in the Ethernet section into analog audio and analyzed the result using open source audio analysis software (WaveSurfer). We found that a frequency component of 1100 Hz, namely, a calling (CNG) signal (**Fig. 3**) used in fax communications, appeared in the audio signal directly before the problem occurred (**Fig. 4**).

This frequency component was included in the hold tone (tune: "Let It Be") output by the extension telephone (stand-alone phone) accommodated by the PBX. An actual CNG signal is specified as the transmission of a 1100-Hz tone for a period of 0.5 s ±15% (0.425–0.575 s), but the analyzed audio included the transmission of a 1100-Hz tone for a period of 0.7 s, which was outside the CNG specified value. Nevertheless, we still considered the possibility that the PBX erroneously recognized the audio signal as a fax CNG signal and switched the PT to "103" accordingly, so we continued with our analysis.

Although the VoIP-gateway side also switched to PT "103" in step with the PBX side, no CED signal (called station identification: reply signal to CNG) arrived at the PBX, so the PBX switched back to PT "0." However, the VoIP-gateway side did not

*1 Payload type (PT) in RTP indicates the digital encoding method applied to audio.

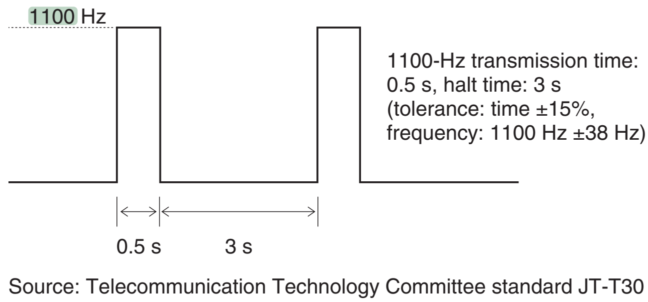


Fig. 3. Format of CNG signal.

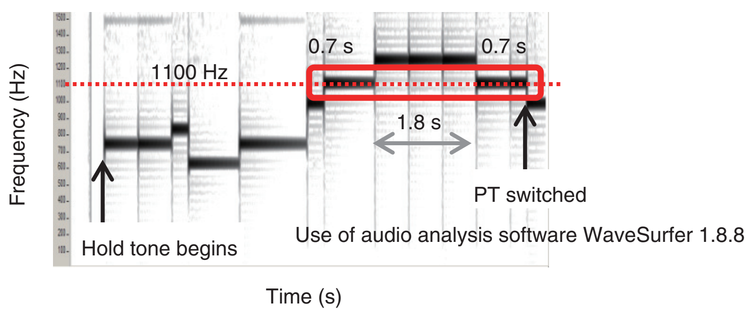


Fig. 4. Results of analyzing audio components.

subsequently switch back*2 the PT, which seems to be the reason why voice transmission occurred only in the direction from the PBX to the gateway (one-way-voice state).

4. Countermeasures

We proposed the following three countermeasures to suppress PT switching in the PBX:

Countermeasure 1: Set the PBX so as not to switch to PT "103" used for fax communications.

Countermeasure 2: Change the source of the hold tone.

Countermeasure 3: Exchange the PBX package to one with no function for switching to PT "103" for fax communications.

Of these countermeasures, we considered that countermeasure 1 would be the least burdensome for the customer and therefore implemented it to check its effectiveness. We did this by connecting the stand-alone phone actually used by the customer to a test PBX and noted that no one-way-voice problem occurred when transmitting the above hold tone during a call. On the basis of this result, we were able to confirm that the problem occurred because of the switch to PT "103" for fax communications even when the PBX was not conducting fax communications and that countermeasure 1 was effective in solving the problem.

5. Summary

The one-way-voice problem described in this report occurred for the following reason. On observing the transmission of an 1100-Hz audio component for 0.7 s closely resembling the CNG signal used in fax communications within the audio of the customer's extension call, the PBX erroneously recognized that frequency component as a CNG signal and switched the PT of the RTP packets to "103" used in fax communications.

This problem was solved by setting the PBX so that it would not switch the PT to "103" for fax communications.

*2 No-switch-back operation is a PBX specification (confirmed by PBX development personnel).

6. Conclusion

In IP network services, a system configuration that includes VoIP gateways and PBXs may use equipment with analog and ISDN interfaces. Solving problems that arise in such a system requires measurement equipment and analysis skills appropriate for each type of interface.

The case study presented here showed how analyzing both analog signals and Ethernet data led to a solution. Going forward, the team at the Technical Assistance and Support Center aims to contribute to solving troublesome problems in network services by analyzing signals and data in various types of interfaces using diverse tools and techniques.

External Awards

IPSJ Computer Science Research Award for Young Scientists

Winner: Toshiyuki Kurabayashi, NTT Software Innovation Center
Date: July 31, 2017
Organization: Information Processing Society of Japan (IPSJ)

For "Automatic High Coverage Test Data Generation for Integration Testing by Dynamic Symbol Execution and Search-based Software Testing."
Published as: T. Kurabayashi, X. Zhang, and H. Tanno, "Automatic High Coverage Test Data Generation for Integration Testing by Dynamic Symbol Execution and Search-based Software Testing," Proc. of IPSJ/SIGSE Software Engineering Symposium 2016, pp. 147–152, Aug./Sept. 2016 (in Japanese).

Best industry paper award

Winner: Shinobu Saito and Yukako Iimura, NTT Software Innovation Center; Aaron Massey, University of Maryland, Baltimore; Annie Antón, Georgia Institute of Technology
Date: September 6, 2017
Organization: The Institute of Electrical and Electronics Engineers (IEEE)

For "How Much Undocumented Knowledge is there in Agile Software Development?"
Published as: S. Saito, Y. Iimura, A. Massey, and A. Antón, "How Much Undocumented Knowledge is there in Agile Software Development?," The 25th IEEE International Requirements Engineering Conference, Lisbon, Portugal, Sept. 2017.

Papers Published in Technical Journals and Conference Proceedings

Understanding Semantic Structure of Tables with a Hybrid Deep Neural Network Architecture

K. Nishida, K. Sadamitsu, R. Higashinaka, and Y. Matsuo

Proc. of AAAI-17 (Thirty-First Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence), pp. 168–174, San Francisco, CA, USA, February 2017.

We propose a new deep neural network architecture, TabNet, for table type classification. Table type is essential information for exploring the power of Web tables, and it is important to understand the semantic structures of tables in order to classify them correctly. A table is a matrix of texts, analogous to an image, which is a matrix of pixels, and each text consists of a sequence of tokens. Our hybrid architecture mirrors the structure of tables; its recurrent neural network (RNN) encodes a sequence of tokens for each cell to create a 3d table volume like image data, and its convolutional neural network (CNN) captures semantic features, e.g., the existence of rows describing properties, to classify tables. Experiments using Web tables with various structures and topics demonstrated that TabNet achieved considerable improvements over state-of-the-art methods specialized for table classification and other deep neural network architectures.

---

200 Gbit/s 16QAM WDM Transmission over a Fully Integrated Cladding Pumped 7-Core MCF System

C. Castro, S. Jain, Y. Jung, E. De Man, S. Calabrò, K. Pulverer, M. Bohn, J. Hayes, S.-ul Alam, D. J. Richardson, K. Takenaga, T. Mizuno, Y. Miyamoto, T. Morioka, and W. Rosenkranz

Proc. of the Optical Fiber Communication Conference and Exhibition (OFC) 2017, Th1C. 2, Los Angeles, CA, USA, March 2017.

A complete, realistic integrated system is investigated, consisting of directly spliced 7-core multi-core fiber, cladding-pumped 7-core amplifiers, isolators, and couplers. The system is demonstrated in a 16QAM (quadrature amplitude modulation) C-band wavelength division multiplexing scenario over 720 km.

In-service Crosstalk Monitoring for Dense Space Division Multiplexed Multi-core Fiber Transmission Systems

T. Mizuno, A. Isoda, K. Shibahara, Y. Miyamoto, S. Jain, S. U. Alam, D. J. Richardson, C. Castro, K. Pulverer, Y. Sasaki, Y. Amma, K. Takenaga, K. Aikawa, and T. Morioka

Proc. of OFC 2017, M3J. 2, Los Angeles, CA, USA, March 2017.

We present in-service inter-core crosstalk monitoring for multi-core fiber transmission systems. We transmit 54-WDM (wavelength division multiplexing) PDM-16QAM (polarization-division multiplexed 16 quadrature amplitude modulation) signals over 111.6-km 32-core dense space division multiplexed transmission line incorporating cladding-pumped 32-core MC-EYDFA (multi-core erbium/ytterbium-doped fiber amplifier), and demonstrate –30 dB crosstalk monitoring without affecting transmission performance.

High-capacity Dense Space-division Multiplexed Transmission Systems Using Single-mode Heterogeneous Multicore Fibre

T. Mizuno

Proc. of ISUPT 2017 (8th International Symposium on Ultrafast Photonic Technologies), Invited talk, Session 4, p. 32, Winchester, UK, July 2017.

Space-division multiplexing (SDM) is one of the most promising technologies for future ultra-high capacity optical transport networks. This invited talk will focus on dense space-division multiplexing (DSDM) transmission technology using a single-mode heterogeneous multicore fibre (MCF) with more than 30 spatial channels. We review our crosstalk-managed 32-core DSDM transmission line employing heterogeneous design with a square lattice arrangement. We also review our recent demonstration of high-capacity and long-distance transmission including the 32-core DSDM unidirectional transmission over 1600 km, the inline amplified DSDM transmission with a 32-core cladding-pumped multi-core Er/Yb doped fibre amplifier, and the 1 Petabit/s inline-amplified DSDM transmission over 200 km. Part of this research utilized results from the EU-Japan coordinated R&D project on "Scalable And Flexible optical Architecture for Reconfigurable Infrastructure (SAFARI)" commissioned by the Ministry of Internal Affairs and Communications (MIC) of Japan and EC Horizon 2020.

High-capacity Dense Space Division Multiplexed Multicore Fiber Transmission

T. Mizuno, K. Shibahara, T. Kobayashi, and Y. Miyamoto

Proc. of OSA Advanced Photonics Congress 2017, NeTu2B.3, New Orleans, LA, USA, July 2017.

We present dense space-division multiplexed transmission technology based on a single-mode multicore transmission line. Long-distance 1600-km transmission and 1-Pb/s high-capacity 205.6-km transmission over a crosstalk-managed heterogeneous 32-core fiber are reviewed.

Self-propelled Ion Gel at Air-water Interfaces

K. Furukawa, T. Teshima, and Y. Ueno

Scientific Reports, Vol. 7, pp. 9323-1–8, August 2017.

We report on a self-propelled gel using ionic liquid as a new type of self-propellant that generates a powerful and durable motion at an air-water interface. The gel is composed of 1-ethyl-3-methylimidazolium-bis(trifluoromethylsulfonyl)imide (EMIM-TFSI) and poly(vinylidene fluoride-co-hexafluoropropylene) (P(VDF-co-HFP)). A long rectangular ion gel piece placed on the interface shows rapid rotation motion with a maximum frequency close to 10 Hz, corresponding to a velocity over 300 mms^{-1} at the outmost end of the piece. The rotation continues for ca. 10^2 s, followed by a reciprocating motion ($<\sim 10^3$ s) and a nonlinear motion in long-time observations ($>\sim 10^3$ s). The behaviours can be explained by the model considering elution of EMIM-TFSI to the air-water interface, rapid dissolution into water, and slow diffusion in an inhomogeneous polymer gel network. Because the self-propellants are promptly removed from the interface by dissolution, durable self-propelled motions are observed also at limited interface areas close in size to the gel pieces. A variety of motions are induced in such systems where the degree of freedom in motion is limited. As the ion gel possesses formability and processability, it is also advantageous for practical applications. We demonstrate that the gel does work as an engine.
