Approach to Network Science— Solving Complex Network Problems through an Interdisciplinary Approach

Kohei Shiomoto

Abstract

As networks come to be used in diverse ways and become increasingly complex and massive in scale, it is becoming difficult to support networks using only existing network technologies aimed at achieving complete control of individual network elements. At NTT Network Technology Laboratories, we are researching and developing network science as an interdisciplinary approach that combines existing network technologies with new technologies from other fields. We are also carrying out research and development (R&D) on technologies that apply network science to enable service providers and end users to use networks in more intelligent ways. This article provides an overview of these R&D efforts at NTT.

Keywords: interdisciplinary approach, communication traffic management, communications-service quality management

1. Introduction

Broadband access has become widely available through both fixed-line and wireless means, and information and communication technology (ICT) has become a vital social infrastructure in our daily lives. The NTT Group has made global cloud services the cornerstone of its business operations and has simultaneously taken up the challenge of improving the competitiveness and earning power of its network services. Furthermore, as reflected by its Hikari Collaboration Model announced in May 2014 [1], the NTT Group seeks to stimulate the ICT market by assisting a wide array of players in creating new value and to contribute to solving social problems and fortifying Japan's industrial competiveness. In short, the NTT Group aims to build up the earning power of its network services while making extensive cost reductions and providing high-value-added services in collaboration with operators in a variety of industrial fields. At NTT Network Technology Laboratories, we are researching and developing network science to support the network infrastructure that the NTT Group needs to advance the above initiatives.

The issues to be resolved and goals to aim for in constructing the future network are shown in **Fig. 1**. Network usage patterns are changing as ICT becomes entrenched as a social infrastructure. Numerous video and movie services are now being provided, and mobile services over smartphones and tablets are proliferating. At the same time, the diversification of terminals looks to intensify in the years to come with the advent of wearable terminals and advanced sensors, and the 5th generation of mobile network technology appears to be ready for rollout in 2020. Improving not only transmission speeds but energy and spectrum efficiencies as well will enable users to enjoy high-definition video services in all kinds of places using all sorts of mobile terminals. In addition,
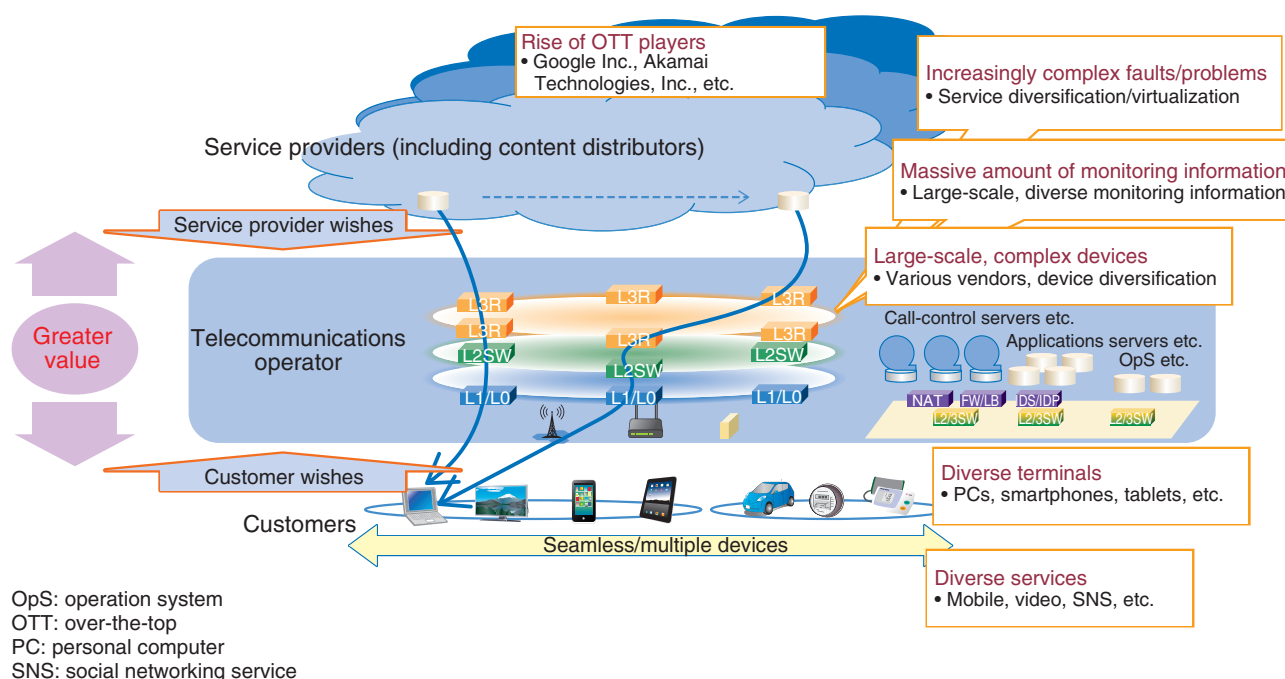
Fig. 1.   Issues and goals toward the future network (service co-creation network).

the diffusion of Internet of Things (IoT)[*1] technologies and machine-to-machine (M2M) communications will enable many and varied devices to be interconnected via the network (Fig. 1). The provision of services using the cloud will also expand, and services that cannot be imagined under present conditions will be introduced one after another in the future. This diversification of services will naturally be accompanied by a wide variety of traffic patterns. In an era dominated by cloud and mobile services, the network must be able to maintain a stable level of quality so that customers can enjoy a satisfactory quality of experience (QoE)[*2] just about anywhere.

Today's network consists of transfer/transmission facilities such as L3 (layer 3) routers, L2 switches, and L1/L0 transmission equipment, server facilities such as call-control servers and application servers, and other types of facilities and functions such as network address translators, firewalls, load balancers, and intrusion detection and prevention systems. As such, the configuration of the network is becoming increasingly complex and diversified (Fig. 1). As a consequence, the operation of the network consisting of such many and varied constituent elements is likewise becoming increasingly complex.

Furthermore, in addition to network operations, end-to-end operations including the cloud are also

important, and they are becoming a factor in this rise in complexity. Additionally, the introduction of virtualization technologies such as software-defined networking (SDN)[*3] and network function virtualization (NFV)[*4] will enable the design of logical configurations independent of physical configurations, although this can have the effect of making responding to problems all the more complicated.

The network structure is also becoming increasingly complicated. Operators of content delivery networks (CDNs) are constructing content delivery platforms that connect cache servers over the network on a global scale. Even service providers that are expanding their services throughout the world, for example, Google Inc., are beginning to construct networks. Service providers and CDN operators such as these are called *Hyper Giants*, and they can have a huge impact on traffic flow by controlling content

---

*1 IoT: Technology for giving all sorts of things communication functions to enable Internet/cloud connections, mutual control, etc.

*2 QoE: The level of quality experienced by a user of communication services.

*3 SDN: Generic term for technology that enables centralized management of network communication devices and alteration of network configuration/settings via software.

*4 NFV: A method for implementing network functions by software on general-purpose servers using virtualization technology.
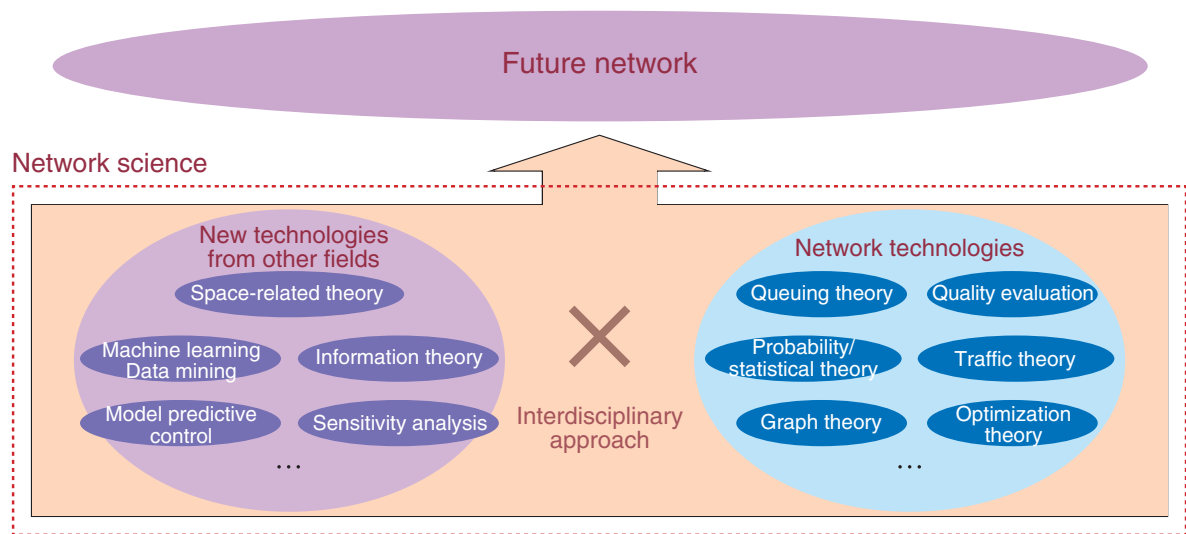
Fig. 2. Interdisciplinary approach to network science.

delivery on a global scale [2]. In fact, mutual interference between content control by Hyper Giants and network control by telecommunications operators is becoming a problem.

Furthermore, a variety of operators are becoming involved in the network. In addition to service providers that provide services to general customers and telecommunications operators that construct and operate networks, CDN operators that provide efficient delivery of content can be found between telecommunications operators and content providers. Telecommunications operators must be able to support service providers and CDN operators and not just general customers. It is essential that they provide high-value-added services in collaboration with all kinds of players.

Consequently, as network usage scenarios become increasingly diverse and as networks themselves become increasingly complex and massive as described above, it is becoming all the more difficult to cope with a network using only existing network technologies that aim to have complete control of individual network elements. At NTT Network Technology Laboratories, we are researching and developing network science as an interdisciplinary approach that combines existing network technologies with new technologies from other fields. We are also researching and developing technologies that apply network science to enable service providers, end users, and other operators to use the network in more intelligent ways.

2. Network science

The existing approach assumes that the target system is to be faithfully modeled and analyzed using stochastic and statistical techniques so that the individual elements making up the system can be well understood. In the existing network, however, obtaining complete control of individual elements making up the system is in itself becoming extremely difficult. With this in mind, we are researching and developing an interdisciplinary approach called network science that combines new technologies from other fields with existing network technologies (**Fig. 2**).

A variety of domains can be considered as new technologies from other fields, but at present, we are exploring the application of technologies such as space-related theory, machine learning and data mining, information theory, and model predictive control.

2.1 Space-related theory
Space-related theory deals with spatial information against the background of academic demands and social trends. Up to now, traffic theory has targeted stochastic behavior along a time axis, but in space-related theory, the focus is on a theory that targets stochastic behavior along a space axis. Integral geometry forms some of the foundation for this theory; given that the area and perimeter length of certain objects are known, the probability that their shapes overlap can be calculated. This principle can be used,

for example, to calculate the probability that spatially and stochastically distributed objects overlap [3]. Application examples of this theory include the analysis of measurements taken by sensors and the design and control of networks that are robust to natural disasters. Constructing an actual physical network involves designing a network highly resistant to disaster by calculating the probability of being hit by a disaster based on a hazard map showing the likelihood of such an event (an earthquake, for example). This theory can also be applied to network control in a virtual network in which the logical configuration of the network can be controlled independently of its physical configuration. For example, the ever-changing probability of being hit by a disaster can be calculated based on forecasts of typhoons, torrential rains, or other extreme weather conditions, so that network servers can be switched to an arrangement with a low probability of being affected by such a disaster.

At NTT Network Technology Laboratories, we are also studying base-station design using homology. The use of conventional mathematical programming to calculate an optimal arrangement of sector antennas can result in an explosive number of combinations that hinders problem solving. Homology-based modeling, however, can drastically reduce computational complexity in such cases, enabling a solution to be found [4].

2.2 Machine learning and data mining

The combined approach of machine learning and data mining enables the discovery of features and regularities in large volumes of data. For our purposes, the target of analysis is traffic collected from the network and data related to service quality. This type of data is not limited to numerical data. It can include, for example, atypical messages in logs generated by equipment and devices. Specifically, we use machine learning and data mining to analyze large volumes of data and discover the relationship between two types of numerical data (regression), to discover groups of data having similar properties from large volumes of data (clustering), and to sort data into predetermined categories based on training data (classification). We are also using machine learning and data mining to achieve advanced network operations [5].

2.3 Information theory

Information theory includes a technique called compressed sensing [6]. Because the elements of an observation vector can be expressed as the product of an unknown vector and a known matrix, compressed sensing can be used to infer the unknown vector. Specifically, if the structure of the known matrix is sparse—that is, if many of the elements of the matrix are zero—the unknown vector can be inferred. Compressed sensing is a technique originally developed in fields such as image processing and signal processing, but its application to network operations is progressing. Here, assuming that end-to-end quality on a variety of network paths can be measured, we consider a situation in which quality deterioration has been detected on several paths. Compressed sensing can then be used to determine which interval in the network is the source of this deterioration. In mobile communications, compressed sensing can be applied to techniques for isolating locations of quality deterioration [7].

2.4 Model predictive control

Model predictive control can be used in cases where predicting the response to a network control operation such as changing transmission paths is difficult. This technique was originally developed for controlling plant processes. Model predictive control iteratively solves an optimization control problem while shifting forward an N-step interval at every step of the process. It applies only the first operation step of the optimization input sequence to the control target (**Fig. 3**). An optimization control problem that takes N forward steps into account makes predictions N steps into the future and solves an optimization control problem with control conditions attached at every input step at that time. Model predictive control theory has so far been applied to traffic engineering, where it was found to be effective [8]. It is now being studied for use in resolving the problem of mutual interference in control operations between telecommunications operators and Hyper Giants.

3. Application of network science

With the aim of applying network science, we are researching and developing (1) network data analysis, (2) proactive network control, (3) QoE-centric operation, and (4) Disaster-free Networks, as summarized below.

3.1 Network data analysis

In network data analysis, we mainly use machine learning and data mining to formulate effective countermeasures to faults in network services. We study in particular each step making up a fault countermeasure, namely, fault detection, root cause
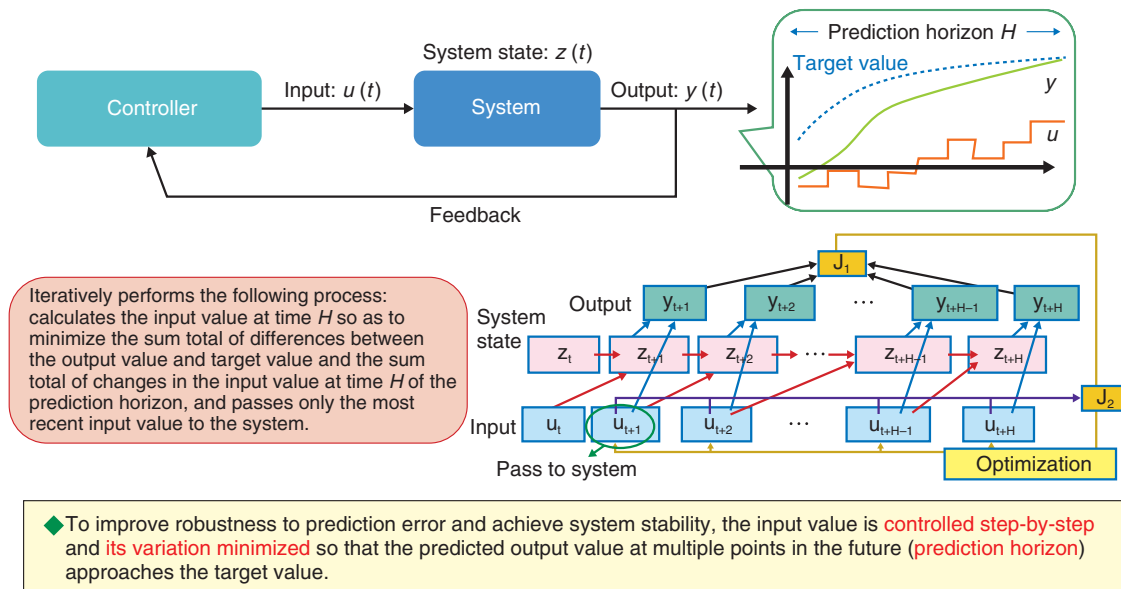
Fig. 3. Model predictive control.

analysis, and service restoration. First, for the fault detection step, we are developing technologies for visualizing service status, supporting visual monitoring, detecting silent faults, performing predictive detection, and other tasks. For the root cause analysis step, we are developing technologies for identifying the root cause and location of a fault and for determining the impact of a fault on services. Finally, for the service restoration step, we are developing technologies for formalizing and automating restoration tasks.

In developing these technologies, we apply machine learning and data mining, information theory (compressed sensing etc.), and optimization theory to diverse types of data related to network operations such as syslog[*5] data, Twitter[*6] data, numerical data, response histories, and workflow data [9]. For example, network devices generate syslog data in great quantities, but the messages that are output are in formats that differ from one device to another and from vendor to vendor. As a result, using this information in on-site maintenance operations can be difficult. However, machine learning and data mining technology can be used to discover regularities in such a huge quantity of syslog messages such as messages that tend to be issued simultaneously, and it can also be used to visualize network operations [10]. Additionally, this technology can be used to detect and comprehend service failures by analyzing Twitter

data. The idea here is to train the system beforehand using previous tweets posted at the time of service failures and to then identify tweets related to service failures during real-time monitoring of Twitter feeds. These failure-related tweets can then be extracted and used to clarify the nature of a service failure [11].

The network of the future will have an increasingly complicated and diverse configuration, and network operation will be all the more complex. Furthermore, if we include the cloud when talking about the network, we can expect even greater complexity. In light of this complexity, the detection of faults and failures, troubleshooting and factor clarification after detection, and failure recovery after troubleshooting will become increasingly difficult. Moreover, in addition to simply determining whether a service interruption has occurred, it will also be necessary to deal effectively with service flaws that are difficult to quantitatively evaluate such as deterioration in the customer's QoE. Such service flaws are sometimes called obscure faults whose reproducibility is low, and solving them may require a relatively long time as a result. The introduction of SDN, NFV, and other virtualization technologies is expected to make problem response even more complicated, so we can expect

*5 Syslog: Log information consisting of software/hardware operating conditions periodically recorded by network devices such as switches, routers, and servers.

*6 Twitter is a registered trademark of Twitter, Inc.

network data analysis to become increasingly important going forward.

3.2 Proactive network control

Proactive network control is a network control technique that optimizes the network configuration based on traffic predictions [12]. To make traffic predictions, it is important to clarify the latent mechanisms involved in generating traffic.

The ways in which users use the network are also diversifying, so it is also important to understand for what purposes users use the network in order to make reliable traffic predictions. Additionally, there is a need to be aware of mobile traffic arising from the increasing popularity of smartphones and tablets. Here, it is important to understand how users move through both cyberspace and physical space. We are developing technologies to solve the above problems by making use of space-related theory and machine learning and data mining.

Controlling traffic based on traffic predictions requires that the error arising in those predictions be minimized. However, when performing network control as in changing transmission paths, it is becoming increasingly difficult to predict the response to such a control operation. This difficulty is reflected by the mutual interference that can occur between content control by a Hyper Giant and network control by a telecommunications operator. A Hyper Giant may change the servers it uses to deliver content based on measurements of network quality. This can have the effect of changing traffic flow on the telecommunications operator's network and generating congestion. The telecommunications operator may then change network transmission paths to deal with that congestion, but from the Hyper Giant's point of view, this may mean a change in communications quality. It is in this way that mutual interference arises between content control by a Hyper Giant and network control by a telecommunications operator. It has consequently become difficult for a telecommunications operator to predict the result of its own control operation to change transmission paths. Model predictive control is considered to be an effective technique under such conditions.

3.3 QoE-centric operation

In the operation and management of network services, there is a growing need to consider the user's QoE in addition to simply determining whether a service interruption has occurred [13]. Although QoE is a subjective matter involving such characteristics as *beautiful* (e.g., for images) and *natural* (for sound), quantifying such subjective characteristics should enable QoE to be inferred based on indicators measured on the network side. The QoE so obtained could then be used in the operation and management of network services.

Until recently, the main approach to quantifying QoE has been to conduct experiments with multiple subjects seated in evaluation booths. However, it is now becoming important to ascertain user QoE in the field though the use of crowdsourcing with smartphones and other novel means. Furthermore, to facilitate estimation of QoE, we can model the correspondence between data that can be directly measured within the network such as throughput and delay and QoE related to video, voice, and data-communications services. As for video, we are researching and developing HTTP/TCP (Hypertext Transfer Protocol/Transmission Control Protocol)-based progressive-download video services that have recently become popular as well as 4K/8K high-definition video services. Next, in terms of voice, we are researching and developing mobile VoIP (voice over Internet protocol) applications such as LINE[*7] and Skype[*8] and IP-based voice services in a mobile environment such as voice over LTE (VoLTE)[*9]. Finally, in the area of data communications, we are targeting web browsing in our research and development (R&D) efforts.

As described above, QoE can be understood on the basis of data measured by the network and used in the operation and management of network services. At the same time, such data can be provided to service providers to assist them in adding value to their services. A quality-related application programming interface (quality API) is one means of providing such data, and we are conducting field tests in collaboration with service providers using such a quality API [14].

3.4 Disaster-free Networks

We are researching and developing Disaster-free Networks applying space-related theory [15]. A Disaster-free Network is a network whose design and control methods are robust to natural disasters and other calamities. Space-related theory can be used to calculate the probability that areas in which natural

---

*7 LINE is a trademark or registered trademark of LINE Corporation.

*8 Skype is a trademark of Skype Limited.

*9 VoLTE: Service that provides voice communications in packet form using Long Term Evolution (LTE).

disasters may occur overlap with the network. In short, when deploying network facilities, the probability that those facilities will be affected by a disaster can be calculated based on a hazard map that indicates the likelihood of earthquakes occurring. The results of those calculations can then be used to design a network highly resistant to disasters. Space-related theory can also be applied to network control in a virtual network focusing on the fact that the logical configuration of the network can be controlled independently of its physical configuration. For example, the ever-changing probability of being hit by a disaster can be calculated based on forecasts of typhoons, torrential rains, and other weather events so that the current arrangement of network servers can be switched to one with a low probability of being affected by such a disaster.

4. Conclusion

As network usage formats diversify and as the network itself becomes increasingly complex and massive, it is becoming all the more difficult to support the network by using only existing network technologies that attempt to achieve complete control of individual network elements. In the face of this problem, NTT Network Technology Laboratories is promoting R&D of network science as an interdisciplinary approach that combines technologies used in diverse fields. In the coming cloud era, we can expect the network to become an increasingly essential infrastructure. We can also expect to see quantum leaps in mobile technologies that will enable users to enjoy a wide range of services wherever the users may be. At the same time, we can envision the penetration of IoT and M2M technologies and the interconnection of many and varied devices over the network. Security too will become increasingly vital as the importance of the network as a social infrastructure grows. In addition, the introduction of virtualization technologies such as SDN and NFV will enable the design of logical configurations that are independent of physical configurations and will help satisfy the need for advanced network construction and operation techniques. Going forward, we aim to create a new academic discipline to support the further evolution of the network through the embodiment and realization of network science.

References

[1] NTT press release published on May 13, 2014.
http://www.ntt.co.jp/news2014/1405eznv/ndyb140513d_01.html

[2] C. Labovitz, S. I-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet Inter-domain Traffic," Proc. of ACM SIGCOMM (Special Interest Group on Data Communication) 2010, pp. 75–86, New Delhi, India, Aug./Sept. 2010.

[3] H. Saito, "Spatial Design of Physical Network Robust against Earthquakes," J. Lightwave Tech., Vol. 33, No. 2, pp. 443–458, 2015.

[4] Y. Ikeda, R. Kawahara, and H. Saito, "Homology-based Metaheuristics for Cell Planning with Macroscopic Diversity Using Sector Antennas," Proc. of IEEE ICC (International Conference on Communications) 2015, London, UK, Jun. 2015.

[5] K. Shiomoto, "Applications of Big Data Analytics Technologies for Traffic and Network Management Data—Gaining Useful Insights from Big Data of Traffic and Network Management," NTT Technical Review, Vol. 11, No. 11, 2013.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2013 11fa1.html

[6] D. L. Donoho, "Compressed Sensing," IEEE Trans. Inform. Theory, Vol. 52, No. 4, pp. 1289–1306, Apr. 2006.

[7] R. Kawahara and Y. Tonomura, "Mobile QoS Tomography Using Compressed Sensing," Proc. of ITC 2014 (26th International Teletraffic Congress), pp. 229–237, Karlskrona, Sweden, Sept. 2014.

[8] T. Otoshi, Y. Ohsita, M. Murata, Y. Takahashi, N. Kamiyama, K. Ishibashi, K. Shiomoto, and T. Hashimoto, "Traffic Engineering Based on Model Predictive Control," IEICE Trans. Commun., Vol. E98-B, No. 6, pp. 996–1007, 2015.

[9] K. Ishibashi, "Analytics-based Operation for Implementing Service Co-creation Networks," NTT Technical Review, Vol. 13, No. 9, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2015 09fa5.html

[10] T. Kimura, K. Ishibashi, T. Mori, H. Sawada, T. Toyono, K. Nishimatsu, A. Watanabe, A. Shimoda, and K. Shiomoto, "Spatio-temporal Factorization of Log Data for Understanding Network Events," Proc. of INFOCOM 2014 (33rd IEEE International Conference on Computer Communications), pp. 610–618, Toronto, ON, Canada, Apr./May 2014.

[11] K. Takeshita, M. Yokota, and K. Nishimatsu, "Early Network Failure Detection System by Analyzing Twitter Data," Proc. of IM 2015 (14th IFIP/IEEE Symposium on Integrated Network and Service Management), Ottawa, ON, Canada, May 2015.

[12] H. Honda, Y. Takahashi, and K. Ishibashi, "Proactive Network Control," NTT Technical Review, Vol. 13, No. 9, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2015 09fa4.html

[13] T. Hayashi, "QoE-centric Operation for Optimizing User Quality of Experience," NTT Technical Review, Vol. 13, No. 9, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2015 09fa3.html

[14] H. Yamamoto, T. Kawano, K. Takeshita, K. Sato, T. Hayashi, K. Miyazaki, S. Sato, B. Kihara, K. Mikami, K. Yonemura, S. Saito, and K. Akimoto, "Quality of Experience Optimization Technology (Quality API)," NTT Technical Review, Vol. 13, No. 6, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2015 06fa3.html

[15] H. Saito, "Basic Theory of Network Science—the Network, Spatial Characteristics, and Spatial Information," NTT Technical Review, Vol. 13, No. 9, 2015.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr2015 09fa2.html

Kohei Shiomoto
Senior Manager of Communication & Traffic Service Quality Project, NTT Network Technology Laboratories.

He received his B.E., M.E., and Ph.D. in information and computer sciences from Osaka University in 1987, 1989, and 1998, respectively. He joined NTT in 1989 and engaged in R&D of asynchronous transfer mode (ATM) traffic control and ATM switching system architecture design. During 1996–1997, he was engaged in research on high-speed networking as a Visiting Scholar at Washington University in St. Louis, MO, USA. During 1997–2001, he directed architecture design for the high-speed IP/MPLS (multiprotocol label switching) label switching router research project at NTT Network Service Systems Laboratories. He also conducted research on photonic IP router design and routing algorithms, and Generalized MPLS (GMPLS) routing and signaling standardization, first at NTT Network Innovation Laboratories and then at NTT Network Service Systems Laboratories, and he was involved in GMPLS standardization in the Internet Engineering Task Force. He led the IP Optical Networking Research Group in NTT Network Service Systems Laboratories from April 2006 to June 2011 and the traffic engineering research group at NTT Service Integration Laboratories from July 2011 to June 2012. Since July 2012, he has been leading the Communication & Traffic Service Quality Project of NTT Network Technology Laboratories. He is a fellow of the Institute of Electronics, Information and Communication Engineers (IEICE), a member of the Institute of Electrical and Electronics Engineers (IEEE), and the Association for Computing Machinery. He has been the Chair of various committees of the Asia-Pacific Board of the IEEE Communications Society, and he served as the Secretary for International Relations of the Communications Society of IEICE from June 2003 to May 2005. He has been involved in the organization of several international conferences including MPLS, iPOP, and WTC. He received the Young Engineer Award from IEICE in 1995 and the Switching System Research Award from IEICE in 1995 and 2000.