

Thinking about “Who Benefits?” and “What Is the Benefit?” and Trying Out an Idea as Quickly as Possible

Masayuki Terada

**Senior Manager, X-Tech Development
Department, NTT DOCOMO**

Overview

Technological innovation has made it possible to generate, collect, and store a vast and diverse range of data. There are efforts being made to analyze data that have been overlooked in the past and use them for business. NTT DOCOMO is also using big data from a new perspective to solve social issues. We interviewed Masayuki Terada, who is engaged in the research and practical application of population statistics using mobile network data called Mobile Spatial Statistics and traffic-jam prediction using artificial intelligence called Traffic Congestion Forecasting AI, about the current progress of research and development and the thrill of being a researcher and developer.

Keywords: population statistics, traffic-jam prediction, differential privacy



Research and practical application of technologies for creating, using, and protecting statistical data

—Please tell us about your current activities concerning research and development.

I am currently involved in the research and practical application of the following three themes: creation of statistics from large-scale data, use of statistics for social prediction, and protection of privacy concerning statistics.

Our research and development (R&D) results regarding the creation of statistics from large-scale data are applied to various services. An example is “Mobile Spatial Statistics” [1], which has been used

for reporting the increases or decreases in the number of people at terminal stations during the novel coronavirus pandemic (**Fig. 1**). From operational data collected from our mobile phone network, the population distribution of where and how many people are located, by age, gender, and place of residence is estimated hourly and almost in real time throughout Japan. During this estimation process, the privacy of our customers is ensured by removing and aggregating personal identifiers as well as through privacy-preserving processing.

Since 2013, Mobile Spatial Statistics has been used by the government for urban planning and disaster countermeasures and by the private sector for store development and analysis of commercial spheres. When this service was first launched, it did not provide

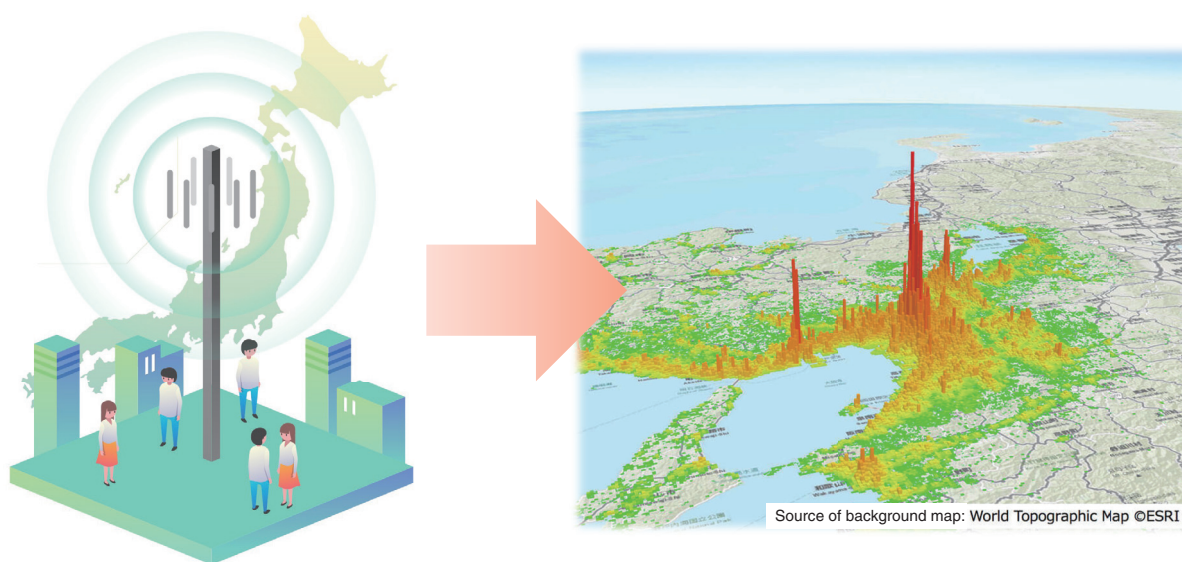


Fig. 1. Mobile Spatial Statistics.

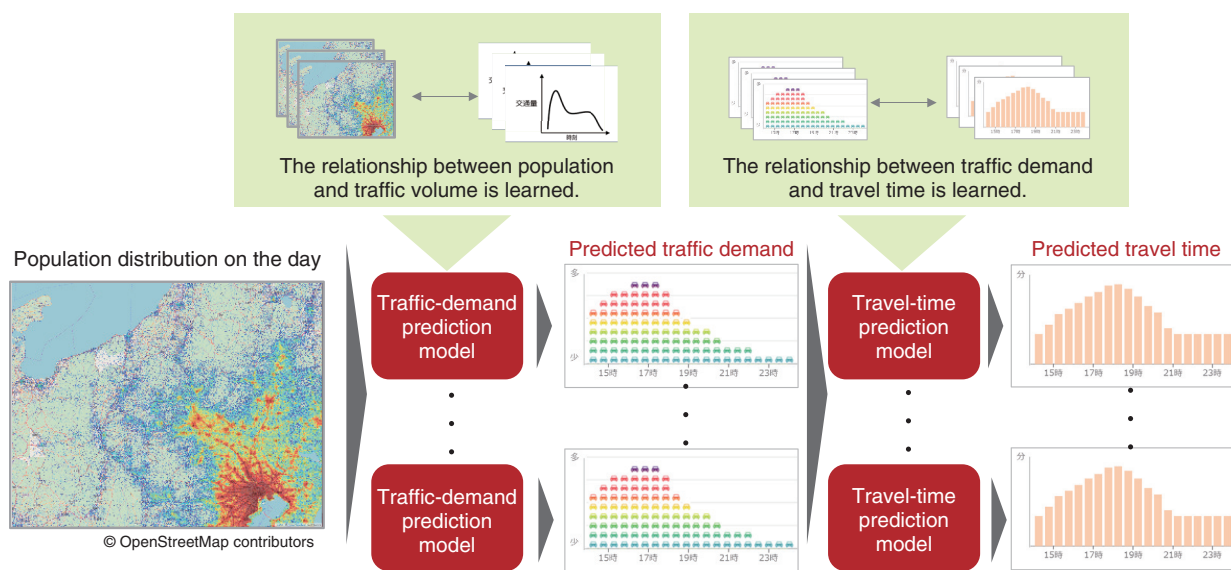


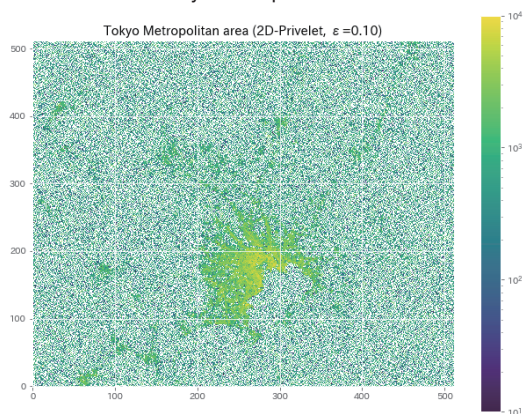
Fig. 2. Traffic Congestion Forecasting AI.

statistics in real time; however, from 2020, it has been providing statistics in near real time in the form of population data from the previous hour or so.

An example of statistics utilization for social prediction is the Traffic Congestion Forecasting AI [2] using Mobile Spatial Statistics (**Fig. 2**). Since social and economic activities are carried out by people, if we can observe the movements of people “now,” we

will be able to predict the “future” of social phenomena and economic trends. Traffic Congestion Forecasting AI is a world-first application of this principle to predicting traffic congestion. This technology differs from long-term congestion forecasts based on seasonal, day-of-week, and holiday patterns (such as the Japanese Obon festival and New Year’s holiday) and short-term forecasts using probe information

Result of applying the current method (two-dimensional Privelet method) to population mesh data around the Tokyo metropolitan area



Result of applying our method (neural network-Wavelet method)

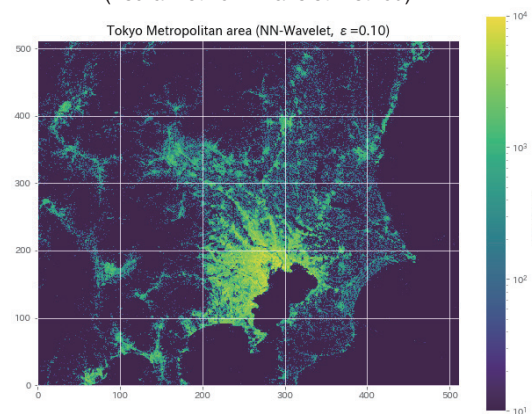


Fig. 3. Applying differential privacy to large-scale tabular data.

from in-vehicle GPS (Global Positioning System), etc. That is, a model is first created by enabling artificial intelligence (AI) to learn sets of past population distributions and traffic volume. Then, the population distribution at a certain time of the day in question is input into the created model, and traffic congestion up to about 10 hours ahead of that time can be predicted.

Traffic Congestion Forecasting AI is being used in a joint experiment with East Nippon Expressway Co., Ltd. (E-NEXCO) to provide traffic-prediction information for the Tokyo Wan Aqua-Line Expressway (a toll way that traverses Tokyo Bay) upstream route (in the Kawasaki direction) and the Kanetsu Expressway upstream route (Numata to Nerima), and that information is distributed daily for widespread use at around 2:00 pm on E-NEXCO's road-information website "Drive Plaza." The use of Traffic Congestion Forecasting AI will make comfortable driving possible without encountering traffic jams. Moreover, if more people use it to avoid traffic jams, traffic jams will decrease or be eliminated through traffic dispersion.

Regarding protection of privacy concerning statistics, we are focusing on applying *differential privacy* [3] to large-scale tabular data (Fig. 3). Since the data handled by NTT DOCOMO are often related to personal information, its handling requires strict legal restrictions and social responsibility. Accordingly, privacy-preserving technology, that is, technology that enables data to be used securely while protecting privacy, is essential. Differential privacy is a frame-

work for comprehensively guaranteeing the safety of privacy-protection technologies, and its application is being researched and developed by Google and Apple. In the U.S., it was announced that differential privacy was applied to the 2020 US Census.

To make Mobile Spatial Statistics more secure and useful, we have been investigating how to apply differential privacy to large-scale geospatial data such as Mobile Spatial Statistics and censuses. The results of our research have been reflected in the concept of privacy protection concerning Mobile Spatial Statistics, and we have returned those results to society in the form of academic papers. In 2020, I was appointed as a special professor at the Statistical Research and Training Institute of the Ministry of Internal Affairs and Communications, where I have been investigating how to apply differential privacy for official statistics in Japan.

—How did you find these research themes?

As the saying goes, "Necessity is the mother of invention." If there is a technology that you need but cannot find it in the world no matter how hard you look for it, it will likely become the theme of your R&D. A typical example of mine is applying differential privacy to large-scale data. Since privacy protection is a fundamental issue in regard to Mobile Spatial Statistics, I thought that a strong security guarantee, such as differential privacy, would be needed. However, no matter how many papers I flipped through at the time, I could not find an

appropriate method to satisfy that need. That's when I started to think, "If a method isn't available, I'll develop it myself."

The story is a little different for Traffic Congestion Forecasting AI. One of the reasons for developing it was that I hate traffic jams and want to avoid them. I heard that E-NEXCO is very concerned about the traffic congestion on the Tokyo Wan Aqua-Line Expressway, and I thought that if we could accurately predict the time when traffic congestion would occur, many people, including myself, would avoid that time, and the congestion would be eased. Therefore, I conducted a simple experiment using the real-time version of Mobile Spatial Statistics that I was working on. I was surprised at the accuracy of the prediction results. Therefore, we recollected the data properly and presented the experimental results to E-NEXCO. They showed strong interest, which led to the above-described joint experiment.

However, these stories were only the beginning of our R&D. I owe a lot to the enthusiasm of the people of E-NEXCO for solving traffic jams, the expectations and support of the corporate sales and business departments of our company from the experimental stage, and the constant technical improvements made by my team members and partner companies. I firmly believe that the current form of Traffic Congestion Forecasting AI would not have been possible without their efforts.

Spare no effort to make it easy

—What are the important perspectives in R&D?

The basic research phase, which is conducted to discover new knowledge, and the applied research phase, for creating new products and services, somewhat differ. It is important to determine how applied research results will contribute to society if the research is successful and who is happy with what; in other words, it is important to be able to answer the questions "Who benefits?" and "What is the benefit?"

I think that in many cases, the important mission of R&D efforts in corporations is to create new business and enable profits through the practical application of research results. However, if you jump at a theme just because it looks like it will be good for business, you will tend to follow the lead of other companies. This situation may be an eternal dilemma for those involved in R&D in corporations.

Under these circumstances, it is important to find a winning "one step ahead" theme and focus on it. To

find such a theme, the perspective of "Who is happy with what?"—that is, the self-questioning of "Who benefits?" and "What is the benefit?"—is useful. At the stage of searching for a theme, you may not be able to see a concrete market yet, and it may not be clear how to make a business out of a theme; however, if that theme has the potential to make someone happy, in other words, if it benefits someone, you should be able to build some kind of sound business model. Additionally, if you can create a group of people in the business department who are interested in and share the value of a theme by explaining "Who benefits?" and "What is the benefit?" you are aiming for, you can create a profitable business model together with them. I believe that this practice will lead to a broader business than you could have imagined on your own.

—Would you tell us what you have valued as a researcher and developer?

I value the saying "Spare no effort to make things easy." It sounds paradoxical, but I try to think on a daily basis about how I can make things easier. For example, in the case of Traffic Congestion Forecasting AI, if we rely on various types and sources of data for traffic-jam predictions, the design of the system will become increasingly complicated, making it more difficult to improve and maintain; in turn, the implementation and operation of the system will become more difficult. Therefore, we have been designing the system on the basis of using only the population-distribution data of the day and not referring to external data, such as weather information, as much as possible. As we proceeded with our research, we were tempted to use any type of data that could be used to improve accuracy in the immediate future; however, the increase or decrease in the number of people due to external factors, such as the weather, is already included in the population-distribution data, so I asked our team members to be patient and focused on refining the accuracy of the algorithm using the population-distribution data alone.

Partly because of reflections on my past, I try to think carefully about when to put research results into practical use and launch onto the market. I was involved in research on electronic tickets and vouchers in my previous job at NTT's research laboratories around 2000. That research was part of an ambitious project marking the beginning of fintech, and it had both a distribution function of virtual currencies, such as Bitcoin and Ethereum, and real payment functions

such as electronic payment. The project was highly regarded internationally and the Internet Engineering Task Force (IETF), which formulates standards for the Internet, established three Request for Comments (RFCs: drafts of standard specifications given by the IETF) in 2003; unfortunately, it was not put to practical use. This project was based on the premise that the world would come to a point where people would always carry an electronic terminal, such as a smartphone, connected to the Internet. However, the first iPhone was not released until 2007, so we had to wait four years after the RFCs were established. The world had not yet caught up with our premise. About 10 years after the establishment of the RFCs, when I had almost forgotten about the project, I was interviewed by a magazine on the theme of “Advanced technology that comes too soon.” This experience made me keenly aware of the necessity of identifying current trends, including peripheral technologies, when selecting R&D themes. Since then, we have been trying to present our research results to the world in step with the progress of peripheral technologies and the social environment while simultaneously presenting our future visions.

Keep as many ideas and tools as possible for nimble testing

—You are searching for research themes and taking on the challenge of development while considering various elements, right?

It is not always easy to find a theme that satisfies certain conditions, that is, if a theme matches current trends, one can gain understanding, solve social issues, and respond to our mantra of “Who benefits?” and “What is the benefit?”; therefore, it is necessary to try out a number of different themes. It is said that “Product development is a process of selecting three out of a thousand,” meaning only three out of a thousand ideas become reality. And I believe that the key is how quickly you can try out an idea once you have come up with a potentially valuable one. I try to keep as many ideas and tools on hand as possible for this purpose.

I mentioned earlier that our Traffic Congestion Forecasting AI started with a “simple experiment.” I was not originally an expert in AI technologies, such as machine learning; even so, machine-learning technologies had been commoditized, and easy libraries had emerged, and my experience of playing around with them to see how easy they were to use came in

handy. If I hadn’t had that experience or known that such libraries existed, I wouldn’t have had the motivation to “give it a try,” and maybe Traffic Congestion Forecasting AI wouldn’t be around now.

The experts I meet at conferences and lectures, as well as the customers I meet through requests from corporate sales colleague to accompany them, are also very helpful in finding themes and tuning the direction of our R&D. For example, when I talk to people about Mobile Spatial Statistics, it seems that those who are seriously considering using it tend to place more importance on the reliability and accountability of the statistics than on the numerical specifications. To meet these expectations, we are enhancing the brand image of Mobile Spatial Statistics by not only pursuing higher specifications but also on the basis of reliability and security.

—Please give a few words to our junior researchers and developers.

There is a saying, “Standing on the shoulders of giants.” It means that our own R&D is based on the accumulation of the results of our predecessors. I think the most exciting part of working in R&D is to be able to contribute to further development of society in the future by putting even a thin layer of your own achievements on top of those of your predecessors. Focused on the three themes that I am currently working on, i.e., creating, using, and protecting statistical data, I hope to put a thin layer of skin on the shoulders of giants and make them as tall as possible. Moreover, I’d be very happy if someone comes along in the future who can stand on top of these giants and make them even taller.

Having said that, I realize that finding a new R&D theme can be difficult. It’s not as if you can find a good theme if you moan a lot and think hard. Most of you probably have some kind of theme that you are working on now. Although it is important to thoroughly finish with a theme and release the results to the world, after you finish with the current theme, you may not find something interesting right away. Therefore, in addition to working on your current theme, I encourage you to take the time to talk with people in various fields, and if you find something interesting from those talks, look into it, become involved in it, and try out your ideas. Even if that process doesn’t directly result in a new theme, you can gain experience, and if you accumulate a large amount of such experience, it may help you find a theme for the future or give you a hint to solve a difficult problem

concerning your current theme from a different perspective.

Your boss and seniors have probably had the experience of finding new themes in this way, namely, experimenting and playing around with ideas under the umbrella of the current theme. While drawing on their knowledge, I hope that you will also value the pleasurable aspects of your work outside your core business of R&D.

References

- [1] Mobile Spatial Statistics (in Japanese), <https://mobaku.jp>
- [2] E-NEXCO Drive Plaza, “Demonstration Tests of Traffic Congestion Forecasting AI” (in Japanese), https://www.driveplaza.com/trip/area/kanto/traffic/ai_traffic_prediction.html
- [3] M. Terada, “What Is Differential Privacy?”, Systems, Control and Information, Vol. 63, No. 2, pp. 58–63, 2019 (in Japanese). https://www.jstage.jst.go.jp/article/isciesci/63/2/63_58/_pdf/-char/ja

■ Interviewee profile

Masayuki Terada

Senior Manager, X-Tech Development Department, NTT DOCOMO, INC.

He received a B.E. and M.E. in engineering from Kobe University, Hyogo, in 1993 and 1995 and Ph.D. in engineering from the University of Electro-Communications, Tokyo, in 2008. He joined NTT in 1995 and moved to NTT DOCOMO in 2003. He is currently engaged in R&D on technologies for privacy protection, population estimation from mobile phone networks, and traffic prediction from population data. He received the IPSJ Outstanding Paper Award and the IPSJ Yamashita SIG Research Award from the Information Processing Society of Japan (IPSJ) in 2015. He also received the IPSJ Industrial Achievement Award in 2019. He is a member of IPSJ and the Institute of Electronics, Information and Communication Engineers.