野村総合研究所

Demonstration Project on Utilization of Privacy Information such as Location Information

Report Summary

March 2019

NOMURA RESEARCH INSTITUTE, LTD.

Contents

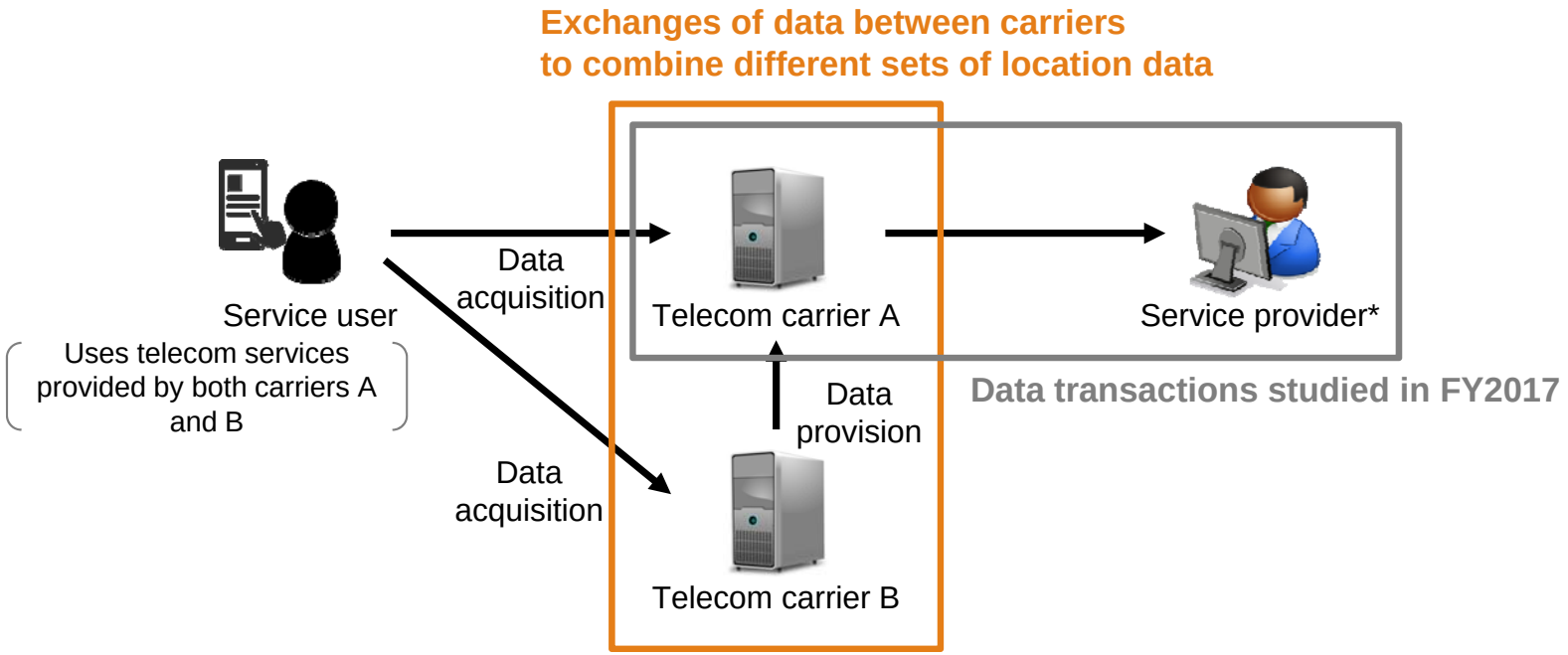
Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

1. Purpose of project

A study was conducted in FY2018 to facilitate exchanges of data (such as location data) between telecommunications carriers in a manner that balances data use with privacy protection.

- An ongoing study has been made of personal data (such as location data) since FY2017 with the aim of achieving a balance between data use and privacy protection. Subjects studied have included the involvement of the individuals whose personal data is being used, data privileges in relation to data distribution, and the division of responsibilities between the parties involved in managing data.
- In FY2018, a study was made of rules governing the sharing and provision between telecommunications carriers of location data handled by carriers. The focus of this study was on exchanges and use of data when a carrier provides location data obtained in the course of providing telecommunications services to another carrier, and the recipient of that data combines and uses that data with location data of its own.
- The study was conducted with the assistance of businesses that agreed with the objectives of this project and its value to society.

Data transactions covered by study



* Business providing services using location data, etc.

Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

A committee of experts and businesses was formed to conduct the study.

1. Elements of study
   - (1) Survey and analysis of privacy protection in relation to use of personal data in Japan and other countries
   - (2) Demonstration using model cases premised on data distribution between multiple carriers
   - (3) New privacy issues arising as a result of advances in IoT
2. Establishment of committee
   - A committee of eight experts specializing in privacy protection, privacy of communications, and information security was formed. This met four times and was attended by the secretariat of the Personal Information Protection Commission of Japan, telecommunications carriers, interested bodies, and businesses involved in the model demonstration.

Committee members, etc.

Members (* study leader)

Ryoji Mori*	Attorney-at-law, Eichi Law Offices
Yuriko Inoue	Professor, International Corporate Strategy, Hitotsubashi University
Tamayo Kimura	Association of Consumer Organizations (“Shufuren”)
Ichiro Satoh	Deputy Director & Professor, National Institute of Informatics
Katsumi Takahashi	Executive Research Scientist, NTT Secure Platform Laboratories
Shinji Terada	Senior Staff, Keio Research Institute at SFC / Executive Director, Mobile Content Forum
Toshiro Hikita	Senior Researcher, Toyota InfoTechnology Center
Tatsuhiko Yamamoto	Professor, Keio University Law School

Observers

- Secretariat of the Personal Information Protection Commission, Japan
- Telecommunications Carriers Association (TCA)
- Japan Data Communications Association
- NTT DOCOMO, Inc.
- KDDI Corp.
- Softbank Corp.
- NTT Broadband Platform, Inc.
- Wire and Wireless Co., Ltd.
- Mitsui Fudosan Co., Ltd.
- Odakyu Electric Railway Co., Ltd.

Secretariat

- Second Telecommunications Consumer Policy Division, Telecommunications Business Department, Telecommunications Bureau, Ministry of Internal Affairs and Communications (MIC)
- Nomura Research Institute, Ltd.

NRI Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

Rules on privacy protection in Japan and other countries were surveyed to identify areas contributing to improved data exchanges between telecommunications carriers.

- A survey of the literature identified matters that should be considered in the data provision contracts provided for in METI's Guidelines on Use of AI and Data when considering contracts between telecommunications carriers.

Literature surveyed

Country	Laws, policy documents, etc.	Details	Key points investigated
EU	ePrivacy Regulation (draft) *Draft revision as of September 2018	Law governing privacy protection in the field of telecommunications in the whole of the EEA*	• Handling of location data • Handling of IoT data
U.S.	California Consumer Privacy Act (CaCPA)	Law enacted in California to protect consumer privacy. Although state legislation, it has repercussions for the U.S. as a whole and beyond.	• Handling of location data • Rules on consumer requests to disclose or cease use
South Korea	Information and Communications Network Act and associated guidelines	Law on personal data protection and other matters in the field of telecommunications services in general in South Korea. Guidelines on non-identification of individuals are associated with this legislation.	• Scope of regulations • Provisions of contracts on B2B data provision
South Korea	Location Data Protection Act	Separate piece of legislation on the gathering and use of location data.	• Handling of location data • Scope of regulations
EU	Communication towards a common European data space	Policy document setting forth the European Commission's digital single market strategy on B2B data sharing, etc.	• Important principles of agreements and exchanges of non-personal data
Japan	METI Guidelines for Contracts on Use of AI and Data	Identifies factors to consider in B2B contracts on data use.	• Thinking on B2B data privileges • Division of responsibilities

Elements to add or consider when revising sample contract (shaded orange)

METI Guidelines data provision contracts	Sample contract prepared for FY2017 study
(1) Definition of data, etc.	Article 1. Definition of terms
(2) Content and methods of provision of data provided	Article 2. Specification of data covered Article 3. Provision, etc. of data
(3) Consent to use of data provided, etc.	Article 4. Restriction of purposes of use of data Article 5. Prohibition of identification of data
(4) Compensation and terms of payment	
(5) Non-guaranteeing of data provided	
(6) Restriction of liability, etc.	Article 6. Restriction of provision to third parties
(7) State of use	Article 8. Measures to ensure secure management
(8) Management of data provided	Article 8. Measures to ensure secure management
(9) Mitigation of damage	
(10) Confidentiality	Article 13. Confidentiality
(11) Handling of derived data, etc..	
(12) Period of validity	Article 16. Period of validity
(13) Disclaimer for force majeure	
(14) Cancellation	
(15) Measures after termination of contract	Article 7. Period of preservation and erasure of data
(16) Exclusion of antisocial forces	
(17) Remaining provisions	
(18) Prohibition of assignment of rights and obligations	
(19) Entire agreement	
(20) Governing law	
(21) Resolution of disputes	Article 18. Jurisdiction by agreement

*EEA：European Economic Area

Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

Detailed model cases of linkage, combination, and use of location data by multiple telecommunications carriers was developed to test acceptability, etc. to consumers.

- In this study, linkage of location data means provision to a third party.
- Combination of location data refers to the process where multiple telecommunications carriers obtain location data on the same consumer, one carrier provides data to the other carrier, and that carrier treats the location data provided and the location data that it acquired itself as data on a single service user by using identifiers as common keys.

How different sets of location data are combined
(in the case of data acquired via Wi-Fi services provided at commercial facilities and stations)

Combining location data using identifiers as common keys

Wi-Fi area	Time	Identifier
Shop A	14:15:53	7a-bb-2d-51-43-11
Shop A	14:15:56	c6-28-66-35-0c-97
Shop B	14:15:59	d8-88-23-a0-01-87
Shop B	14:16:00	d6-ad-33-46-43-91
Mall entrance	14:16:25	8e-25-56-62-4a-9a
Shop B	14:16:34	3e-6e-82-ae-7d-d3
Shop A	14:16:41	83-21-a3-5d-4d-d2
Shop A	14:16:50	c0-44-0e-da-0a-98
Mall entrance	14:17:07	8e-cd-11-a7-18-ec
Mall entrance	14:17:16	c6-28-66-35-0c-97
Mall entrance	14:17:29	5a-34-90-26-a1-27

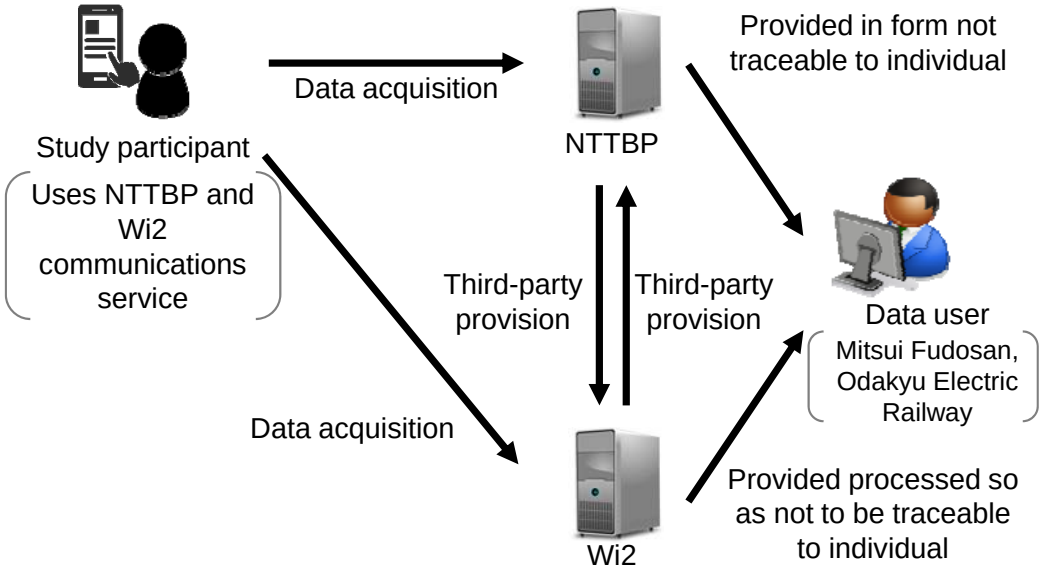
Identifier	Wi-Fi area	Time
7a-bb-2d-51-43-11	Station #1	12:04:01
8e-cd-11-a7-18-ec	Station #2	12:04:06
1c-8a-c9-21-5d-57	Station #2	12:04:11
c6-28-66-35-0c-97	Station #2	12:04:18
1b-2b-86-73-15-dd	Station #1	12:04:30
3e-6e-82-ae-7d-d3	Station #1	12:04:32
8e-cd-11-a7-18-ec	Station #1	12:04:06
1c-8a-c9-21-5d-57	Station #2	12:04:11
8e-cd-11-a7-18-ec	Station #1	12:04:06
c6-28-66-35-0c-97	Station #2	12:04:18
c6-28-66-35-0c-97	Station #1	12:04:18

Identifier	Time	Wi-Fi area	Time	Wi-Fi area
7a-bb-2d-51-43-11	12:04:01	Station #1	14:15:53	Shop A
c6-28-66-35-0c-97	12:04:18	Station #2	14:15:56	Shop A
3e-6e-82-ae-7d-d3	12:04:32	Station #1	14:16:34	Shop B
8e-cd-11-a7-18-ec	12:04:06	Station #1	14:17:07	Mall entrance
c6-28-66-35-0c-97	12:04:18	Station #2	14:17:16	Mall entrance

Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

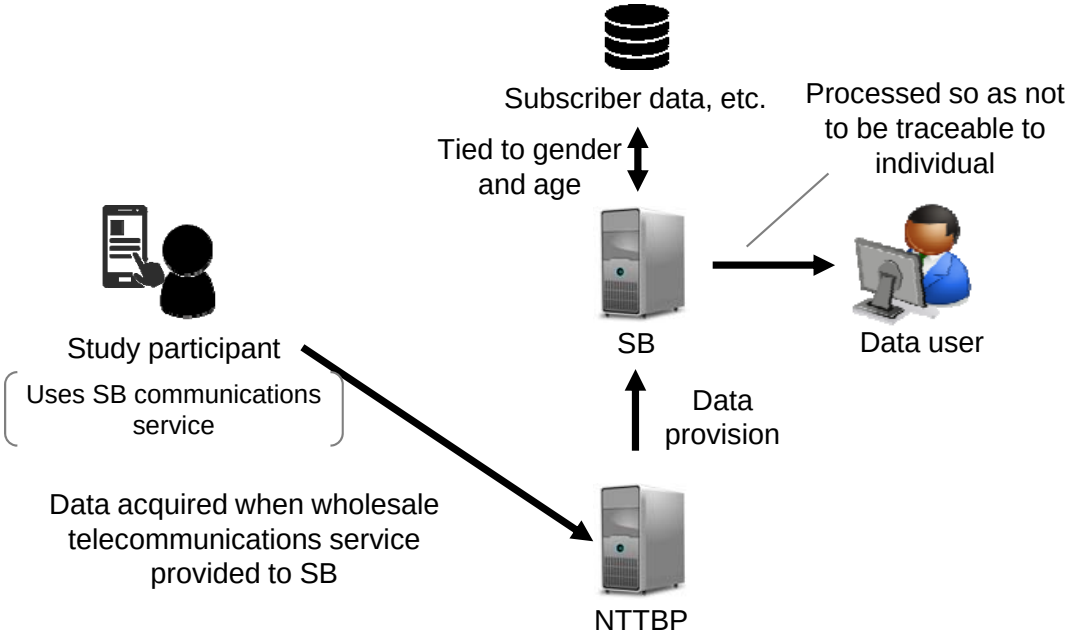
Outline of model cases (demonstration test)

CASE 1 Use of location data provided by third party
*Not traceable to subscriber data



- Study participant uses Wi-Fi service provided by NTT Broadband Platform (NTTBP) and Wire and Wireless Co., Ltd. (Wi2).
- The two carriers that acquired location data on the study participant through the provision of Wi-Fi services provide each other with the location data that they acquired on a third-party basis. They then combine the location data that they received from each other with the location data that they gathered themselves to obtain location data on the study participant over a wider area.* After processing the combined location data so that it cannot be traced to the individual, the data is used to analyze usage of facilities and public transport.
- The study also tested the effectiveness of use of advanced “secure computation” in conjunction with the linkage and combination of location data.

CASE 2 Use of data tied to subscriber data (gender, age) received from wholesale telecommunications service provider



- The study participants were Softbank Corp. (SB) subscribers. For the study, NTTBP provided SB with the location data that it obtained regarding the study participants in the area* in which SB provides Wi-Fi service receiving wholesale telecommunications service from NTTBP. Once the data has been linked to subscriber data (gender and age) by SB and processed so as not to be traceable to specific individuals, it is used to analyze usage of public transport.

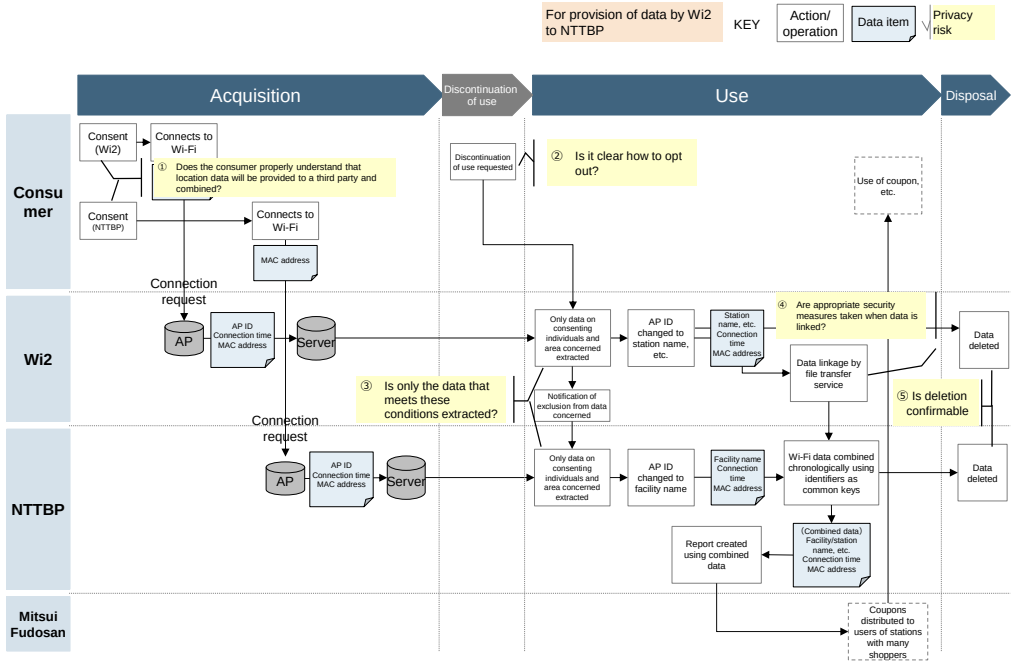
*Limited to area covered by demonstration project.

Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

Privacy risks in the data flows anticipated by the model cases were identified. Response strategies and challenges were then identified to assist exploration and generalization of more detailed responses.

- For the demonstration, location data on study participants in the area studied was acquired by multiple telecommunications carriers and then linked and combined by the carriers.
- As the data items and forms of data contained in the location data did not differ significantly between the participating carriers, there was no data that could not be combined for technical reasons.
- Members of the committee observed that the location data being linked should be hashed.[1] This is because this would not only facilitate alignment and combination of data, but also enhance secure data management during processing.

Data flows and privacy risks (Case 1)[2]



Detailed privacy risks and response strategies (case 1)[2]

Action/operation presenting risk		Privacy risk	Response strategy (top: response in demonstration test, bottom: generalized response and challenges)
Acquisition	Consent of consumer regarding use of Wi-Fi service	① Does the consumer properly understand that location data will be provided to a third-party telecom carrier and that the carrier will use it in combination with its own location data?	✓ Explained per the details of the model cases. Approx. 90% of acceptability survey respondents said that they understood. ✓ **One problem is that although consumers understand the explanation, many do not always check the detailed terms of service.**
Discontinuation of use	Request by consumer to discontinue use	② Is it clear to the consumer how he/she can request discontinuation of use?	✓ It was indicated that the research agency that recruited the study participants (Macromill) was the point of contact for opting out. ✓ Action needs to be taken by the carrier that gathers the data. A carrier that receives a request needs to give the data recipient information including the identifiers for the consumer concerned and the **data recipient needs to take the same steps to discontinue use.**
Use	Extraction of only data on consenting individuals and areas concerned	③ Is only data on consenting individuals in the area concerned extracted from location data log data, and is there no mingling with other data (on other areas and individuals who have not given consent)?	✓ A list of consenting individuals was prepared beforehand. This was shared by the two carriers before data was extracted. ✓ The carrier that acquires data needs to create a list, etc. that distinguishes whether consumers who use its services have **consented to data linkage and combination.**
	Data provision to third-party telecom carrier	④ Are appropriate security measures taken when data is provided to another telecom carrier?	✓ Data was provided using a file transfer service following information security procedures prescribed by the carrier itself. ✓ **Suitably secured means of data provision (leased line, etc.) may be used according to the volume of data provided, etc.**
Disposal	Deletion of data in accordance with consumer's request to discontinue use	⑤ Can it be guaranteed that data has been deleted by data recipients?	✓ Each carrier deletes data after the end of the project and reports that it has done so to the secretariat. ✓ Need to employ a method such as **submission in writing to the data provider that data has been deleted.**

[1] Hashing is the replacement of original data with "hashed values" using irregular fixed-length outputs obtained from original data using certain computational procedures and functions.
[2] With secure computation in Case 1. See report regarding Case 2.

Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

The study confirmed the acceptability of a business model based on linkage, combination, and use of location data. It also verified methods of informing and obtaining the consent of consumers when engaging in such business.

Acceptability to consumers of business model

- A business model based on linkage, combination, and use of location data was found to be somewhat acceptable to consumers. Participants were asked their opinions about a service involving the use of direct mail, coupons, and multiple sets of location data to provide transport information, etc. Regarding all services, between 30% and 60% said that they were happy for location data to be used to provide these services (see figure below).
- However, the initial difficulty encountered in recruiting the desired number of study participants suggests that consumers do not in general find the acquisition and use of location data very acceptable. Members of the committee consequently observed that, despite the limited area studied, the results regarding acceptability should be interpreted taking into account that some consumers are averse to the provision of location data for a certain period.

When receiving the following services, do you feel that is acceptable for businesses to share and use location data acquired through Wi-Fi services provided at facilities? (Please choose the most appropriate option.)

Question no.	Summary of service (business model)	Acceptability to consumers
#10	• Distribution of coupons that can be used at a commercial facility by people living in a particular area when it has been found that the facility is visited by many people from that area.	39% / 50% / 10% / 1% N= 431
#11	• Distribution of information on a sale at cosmetics shop B to users of clothes shop A when it has been found that many shoppers visit both shop A and shop B.	35% / 47% / 16% / 2% N= 431
#12	• Display at stations and on your smartphone, etc. of bus, train, and other information regarding congestion and waits for alternative means of transport, regardless of mode of transport, in the event of a disaster or sudden accident.	59% / 36% / 4% / 0% N= 431
#13	• Display on your smartphone of information on shop D when you alight at station C when it has been found that people who alight at station C are highly probably to visit shop D.	30% / 45% / 22% / 4% N= 431

Very acceptable / Somewhat acceptable / Somewhat unacceptable / Very unacceptable

Methods of informing and obtaining the consent of consumers

- Most consumers are capable of understanding the provision of location data to a third-party telecommunications carrier and the combination of that data with other data when the process is explained to them in the terms of service, etc. Approximately 90% of the respondents said that they understood the explanation* provided for the demonstration project.
- One problem, however, is that many consumers do not always check detailed terms of service.

Notification details for demonstration (Case 1)*

Demonstration test: Information on participation in study

This study is an MIC-backed project to investigate use of location data obtained from smartphones, etc. in order to balance the free distribution of data with privacy protection .

- During the study period, the carriers participating in the study (2) will gather Wi-Fi-related location data (3) from your smartphone when you use Wi-Fi services in the study area (1).

(1) Study area	(2) Carriers participating in study	(3) Location data gathered
• LaLaport Ebina,* LaLaport Yokohama • Odakyu/Sagami line stations	• NTT Broadband Platform • Wire and Wireless	• Names of facilities and stations where Wi-Fi was used (location data) • Device ID of smartphone used (MAC address) • Times of Wi-Fi use

*Location data may similarly be gathered at the ViNAWALK shopping center near LaLaport Ebina.

- The location data gathered will be shared by the participating carriers, processed so that it cannot be traced to particular individuals, and then used to analyze usage of facilities and means of transport.

Example of data after processing:
Dec. 10 (Mon.) LaLaport Ebina 20 visitors Visitor routes (Odakyu line: 15 Machida – Ebina, 5 Atsugi – Ebina)

- Calculation using data in encrypted form (technically known as secure computation) will be used as one method of processing.
- The location data gathered will be promptly deleted when the project ends (end March 2019).
- If you agree to take part but later change your mind, please contact the Macromill Questionnaire Secretariat.
  If you decide to withdraw from the study, your data will be excluded from the analysis if the analysis has not already been performed.

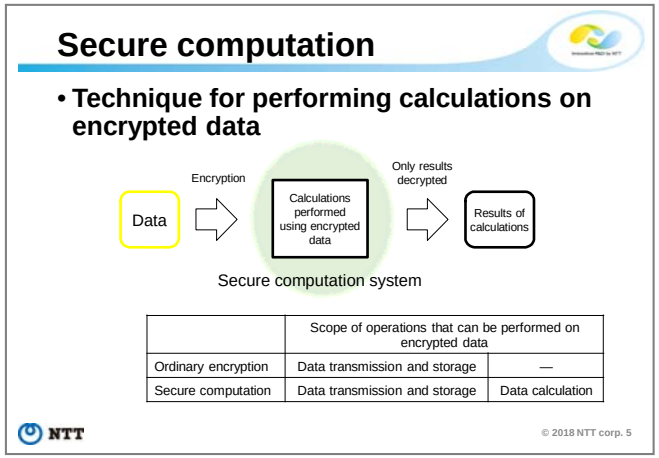
*See the report for details of the notice displayed on study participants' smartphones.

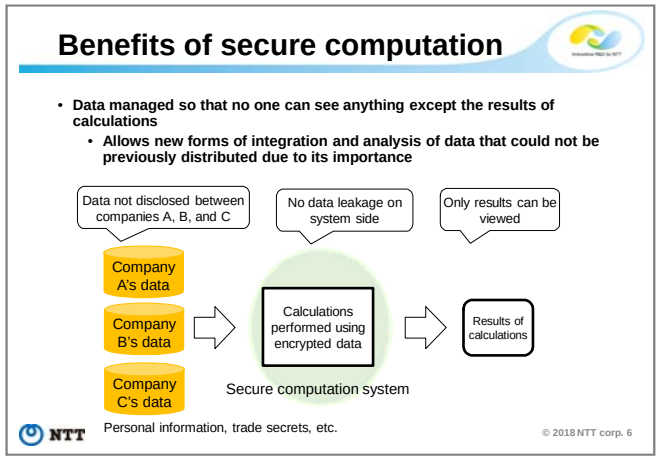
Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

The effectiveness of advanced data analysis techniques relating to secure data management was tested with a view to protecting the privacy of location data.

Summary of techniques tested

- Secure computation allowing analysis of data in encrypted form was tested.
- Assisted by NTT Secure Laboratories, the project employed secure computation. This uses secret sharing based on ISO/IEC 19592-2 as the data format.





Source: NTT Secure Laboratories, *Secure computation systems and principles*.

Results of tests

- Linkage and combination of location data using secure computation was confirmed.
- Processing the encrypted data by secure computation confirmed this technique's utility as a means of secure data management.
- As data is combined and analyzed while still in secure computation format, this technique should also help minimize data while avoiding the careless exposure of data content to operators during the process.
- It should be noted, however, that the above test results were obtained in regard to the secure computation used in the present study, and that the effectiveness of secure computation should be determined according to the specific technique used.

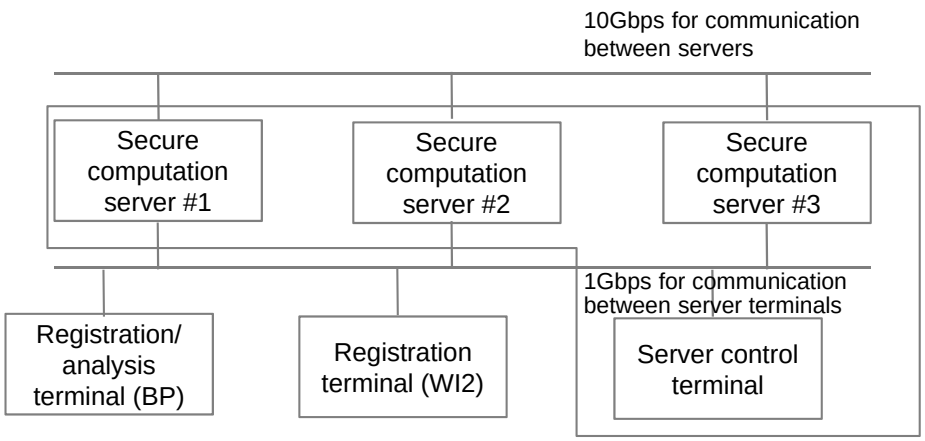
Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

Performance of secure computation

- The results of calculation of data using location data processed by the carriers participating in the demonstration project using generic spreadsheet software were similarly reproduced using secure computation.
- Similar results to those obtained using spreadsheet software were obtained in the test environment shown below. Processing time was 2.7 sec (average of five trials).
- The data used for processing consisted of approximately 15,000 records.

Test environment

System composition



Server, registration/analysis, and server control terminal environment	
OS	RHEL 7.2
CPU	Xenon 4 cores @ 3.5GHz, 2 sockets
Memory	64GB
HDD	SSD 200GB
NW	10Gbps (between servers)

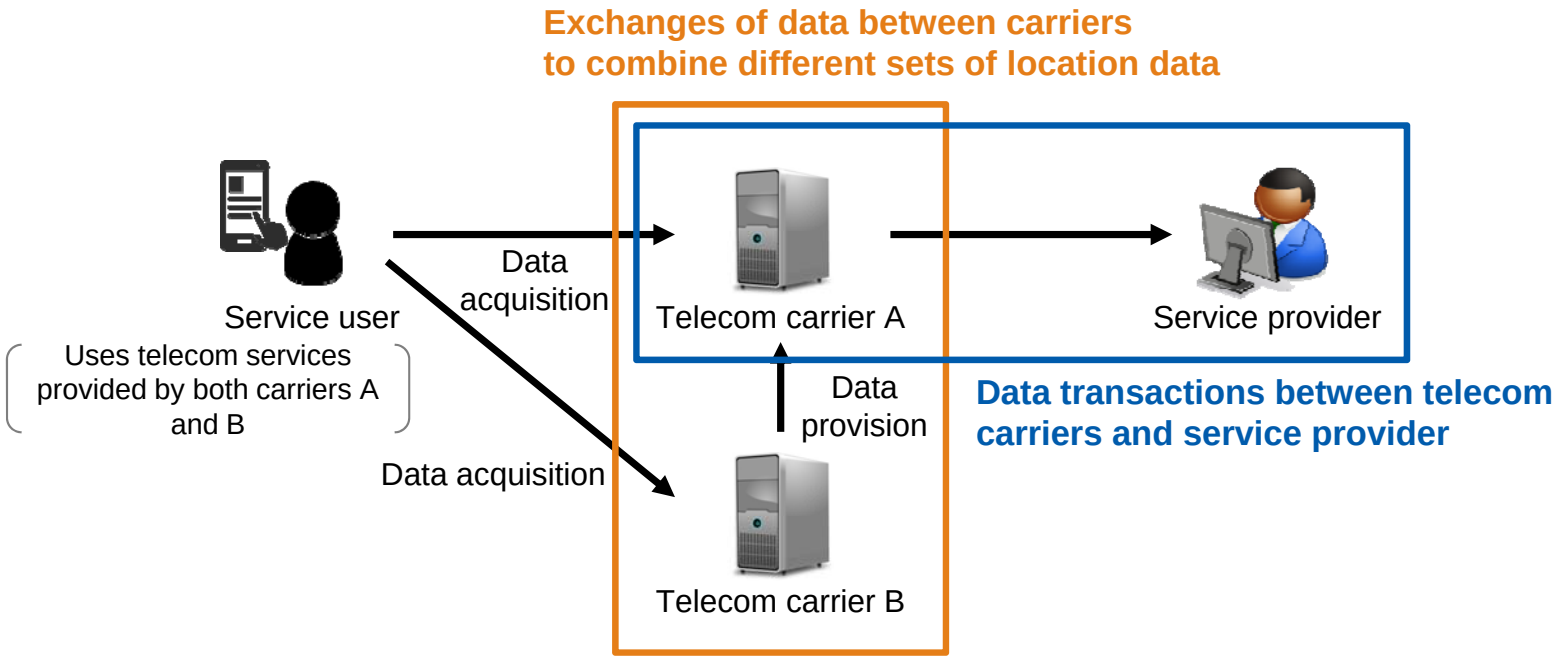
Registration/analysis and server control terminal environment	
OS	Windows7
CPU	Core I 5 2.3GHz
Memory	16GB
HDD	450GB
NW	1Gbps

Copyright(C) Nomura Research Institute, Ltd. All rights reserved.

The sample contract drawn up as a guide for location data transactions between telecommunications carriers was revised in light of the results of the present study.

- The sample contract was revised to provide for data transactions involving the third-party provision of location data acquired by a telecommunications carrier in the course of providing a telecommunications service to another carrier, and the combination and use of that data with location data acquired by the recipient carrier itself.
- The sample suggests provisions that may be considered when a contract on data transactions involving location data processed by telecommunications carriers is entered. The provisions contained in the sample contract are suggestions and may be appropriately modified to provide for other data transactions similar to those envisaged.
- For cases where location data is provided by the provider to the recipient of a wholesale telecommunications service and the location data received by the recipient is used in a form that is traceable to subscriber data, reference should be made to the points of note in the report as well as to the sample contract.

Data transactions anticipated by sample contract



Copyright(C) Nomura Research Institute, Ltd. All rights reserved.