Applying Mobile Big Data for Transport Planning

Relationship between Big Data use and Personal Information Protection

Japan Transport and Tourism Research Institute

Research Fellow
Toshiaki MUROI
Top Questions on Big Data Utilization

① What is the Act on the Protection of Personal Information in Japan? (What is the definition of “Personal Information”?)

Introduce Act and definitions of personal information and “guideline”

② How has Japan’s carrier (such as NTT Docomo) cleared the challenge of protecting personal information when providing their data to third parties?

③ Have you ever had a problem handling the personal information in Japan?

Analyzes success and failure cases and their factors in Japan, and introduces the key points of success in Japan.
How can you comply with Act on the Protection of Personal Information?

Personal information, personal data, retained personal data (Article 2 paragraphs 1, 4 and 5)

The term “personal information”... shall mean information about a living individual which can identify the specific individual by name, date of birth or other description contained in such information. (including such information as will allow easy reference to other information and will thereby enable the identification of the specific individual.)

Details about this personal information are described in the guidelines set by the Ministry of Internal Affairs and Communications (MIC).
MIC Guidelines on the Protection of Personal Information in the Telecommunications Business

Guideline issued by the Ministry of Internal Affairs and Communications in Japan

Specifying basic matters that telecommunications carriers should observe for proper handling to protect personal information.

Guideline revision history
- Enacted in 1998
- Revised in 2005
- Revised in 2009
- Revised in 2010
- Revised in 2011
- Revised in 2013
- Revised in 2015
- Revised in 2017

Act on the Protection of Personal Information was revised in 2017.
Personal Information and Anonymously Processed Information

Original Data (including Personal Information)

ID Number | Gender | Date of Birth | Name | Phone number | Detailed location

Deletion and encryption of some data

ID Number | Gender | Date of Birth | Name | Phone number | Detailed location

Personal Information

Deletion, encryption, aggregation, and concealment of some data

ID Number | Gender | Date of Birth | Name | Phone number | Detailed location

Anonymously Processed Information

ID Number | Gender | Date of Birth | Name | Phone number | Detailed location

Tabulated & Concealed

(C) Mr. MUROI Toshiaki, Japan Transport and Tourism Research Institute, 2020
## Relationship between Personal Information and Anonymously Processed Information

1. **Operational data**
   - **Data with which people can be identified**: 
     - Mr. Tato Yamada

2. **Non-identification process**
   - **Data without personal ID**
     - **Somebody’s data**: 
       - Mr. X located in Zone A

3. **Tabulation process**
   - **Estimated population**
     - **Some group’s data**: 
       - 40’s, 30’s, … in Zone A etc.

4. **Concealing process**
   - **Mobile spatial statistics (NTT Docomo)**
     - **Data of a group including a specific number of people**

### Personal information
- ID
- Name
- Phone Number
- Gender
- D.O.B.
- Detailed location/time

### Anonymously Processed Information
- Encrypted ID
- Name
- Phone Number
- Gender
- Birth year
- Zone
- Around time

### Group information
- **Tabulated**
  - Zone A: 48 people
  - Zone B: 6 people

### Statistics (Non-personal information)
- **Concealed**
  - Zone A: 48 people
  - Zone B: No Data

---

(C) Mr. MUROI Toshiaki, Japan Transport and Tourism Research Institute, 2020
Two examples of using big data in Japan

**CASE 1**

**Mobile Spatial Statistics**
- Cell phone location information
- Number of customers: 79,203,000 (2018)
- Market share: 44.5% (2018)
- Research started in 2008
- Starting in 2013, this service has been implemented

**SUICA (Super Urban Intelligent CArd)**
- Information when passing the gate
- SUICA release quantity: 69,420,000 (2018)
- Number of passengers: 17,000,000/day (2018)
- Research started in 2009
- Announcing commercialization in 2013, Canceled due to public opinion

**CASE 2**
“Personal information” means information that can identify the specific individual.

The act of causing a loss of identification with anonymization is not the use of personal information.

Determined that the Act on the Protection of Personal Information will not apply because of loss of identification.

Prompt businesses to pay social consideration, not just comply with the Act.

Study Group met four times (from Sep. 24 to Dec. 10, 2009)

<table>
<thead>
<tr>
<th>Study Group meeting</th>
<th>Study Group meeting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Law scholar</strong></td>
<td>Masao Horibe</td>
</tr>
<tr>
<td>(Chair)</td>
<td>Prof. Emeritus, Hitotsubashi Univ.</td>
</tr>
<tr>
<td><strong>Information economy expert</strong></td>
<td>Akihiko Shinozaki</td>
</tr>
<tr>
<td></td>
<td>Prof., Kyushu Univ. Graduate School</td>
</tr>
<tr>
<td><strong>Statistics expert</strong></td>
<td>Takeshi Hiromatsu</td>
</tr>
<tr>
<td></td>
<td>Prof., Institute of Information Security</td>
</tr>
<tr>
<td><strong>Consumer perspective advocate</strong></td>
<td>Sawako Nohara</td>
</tr>
<tr>
<td></td>
<td>President &amp; CEO, IPSe Marketing, Inc.</td>
</tr>
<tr>
<td><strong>Lawyer</strong></td>
<td>Tsunemichi Yokoyama</td>
</tr>
<tr>
<td></td>
<td>Mori Hamada &amp; Matsumoto</td>
</tr>
</tbody>
</table>
Compliance with precautionary principle (example)

1. Promote public relations
   - Issue press releases and disclose information on the website to eliminate users’ anxiety and discomfort.

2. Secure transparency
   - Draw up guidelines on big data creation which clarify the purpose, basic principle, creation procedures, three-step processing, etc., and make them available on the website.

3. Secure opportunities for user involvement
   - Ensure that providers can accept, at anytime, users’ requests that base station location information should not be provided to third parties.

Independent Committee report (excerpt)

- Automated de-identification, aggregation and concealment may make personal identification impossible.

(1) Does not violate the Act on the Protection of Personal Information, (2) Comply with the precautionary principle, and (3) conclusion of the Independent Committee

(Telecommunications carriers will) provide third parties with data when approved by Ministry of Internal Affairs and Communications.
Diversity of mobile big data utilization fields

- Disaster prevention
- Public transportation
- Urban planning
- Tourism encouragement
- Event analyses
- Customer information

Effective data in various fields, not limited to transportation
cooperative research conducted by NTT DoCoMo

- **Disaster prevention**
- **Tourism encouragement**
- **Urban planning**

**To estimate of number of people who are unable to return home after disasters, Saitama Prefectural Government,**

**To build up actual performance**

**Tourism behavior research, Okinawa Prefectural Government,**

**Consumer behavior research, Kashiwa Local Government,**

Starting in 2013, this service has been implemented.

- **2010**
  - Cooperative research about Urban planning, The university of Tokyo,

- **2010**
  - Cooperative research about Disaster prevention, Kogakuin University,

- **2011**
  - Cooperative research about Disaster prevention, Kogakuin University,

- **2011**
  - Consumer behavior research, Kashiwa Local Government,

- **2012**
  - Consumer behavior research, Kashiwa Local Government,

- **2012**
  - To estimate of number of people who are unable to return home after disasters, Tokyo Metropolitan Government

(C) Mr. MUROI Toshiaki, Japan Transport and Tourism Research Institute, 2020
### Differences in handling data between the two companies

<table>
<thead>
<tr>
<th>(1) Compliance with laws and guidelines</th>
<th>NTT docomo</th>
<th>Both complied</th>
<th>JR</th>
<th>Both complied</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Prior explanation and information disclosure to users</td>
<td>○ Establishment of a Independent Committee and Dedicated website opened</td>
<td>× Hardly explained</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Dissemination of opt-out procedure</td>
<td>○ Instructions on website</td>
<td>× (at first) there was no announcement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Data accuracy is appropriately “rough”</td>
<td>○ Only hourly &amp; 500m mesh count results are provided</td>
<td>× Provide data without aggregation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Start use from public purpose</td>
<td>○ Start use for public purposes by academics and governments</td>
<td>× Start with private business use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Benefits returned directly to users</td>
<td>○ Providing wide benefits to society</td>
<td>× Unable to show benefits to general users</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There is a high demand for private operators to want to use mobile big data. However, data that the private sector wants to use mainly for marketing is generally a unit for each building or facility, such as each red dot shown in the figure. This is too detailed in terms of space and time, giving the impression that an individual has been identified and may be at high risk of social rebound.

<table>
<thead>
<tr>
<th>Matter</th>
<th>Private Use</th>
<th>Public Plan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Application</td>
<td>Marketing</td>
<td>Transportation Plan City Master Plan</td>
</tr>
<tr>
<td>Space and Time Resolution</td>
<td>Store, Building Narrow (50m²) Short (5min int.)</td>
<td>Zone Wide (500～1,000m²) Long (1hr, 1day)</td>
</tr>
<tr>
<td>The risk that an individual is identified</td>
<td>Large</td>
<td>Small</td>
</tr>
<tr>
<td>Social resilience risk</td>
<td>Large</td>
<td>Relatively restrained</td>
</tr>
<tr>
<td>Publicity</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Data Publish</td>
<td>Hard (Due to contract between private sector)</td>
<td>Easy (Due to data acquire public groups such as governments)</td>
</tr>
</tbody>
</table>
Estimated People Movement based on VinaPhone CDRs

※movie

No photographs

(C) Mr. MUROI Toshiaki, Japan Transport and Tourism Research Institute, 2020

(C) Mr. MUROI Toshiaki, Japan Transport and Tourism Research Institute, 2020
Future MBD use in transportation plans

• Growing expectations for PPP/PFI
  ➢ ASEAN governments are having financial difficulties.

• Current demand forecast has issues with accuracy.

• Risk aversion when demand risk is high
  ➢ Investors call for unfailing recovery of their investment
  ➢ Decision makers consider the balance between improvement and investment

• Too much focus on demand risk (price) will result in

Highly accurate demand forecasts and statistics development contribute to truly addressing traffic problems.
Thank you for your attention

This survey was conducted with support from the Nippon Foundation.