An Effective Use of Tokyo Metro Passengers Flow by Visualization of Smart Card Ticket 'PASMO' Origin-Destination Data for Public Transport Network to be Sustainable

Satoshi Kawamura[1], Yoshimitsu Tomita[2], Masahiko Itoh[3], Daisaku Yokoyama[4], Masashi Toyoda[5] and Masaru Kitsuregawa[6]

[1,2] Tokyo Metro Co.,Ltd, Japan
[3,4,5] the University of Tokyo, Japan
[6] National Institute of Informatics / the University of Tokyo, Japan

Corresponding e-mail: s.kawamura@tokyometro.jp

Abstract

This paper explores the possibility of improving operation and service of the Tokyo Metro system based on large scale trip records of smart card PASMO. Understanding accurate passengers' flow is important for Tokyo Metro to provide sustainable services. Though we already have several statistics such as transportation census, they do not have recent and temporal traffic information. The lack of such information makes it difficult to verify long-term demand forecast, to understand gaps between passengers' behavior and operators' plan, and to dynamically predict demand in case of big events. Trip records of PASMO, which is originally used for fare calculation, enable us to grasp detailed and temporal passengers' flow. We have analyzed four years' worth (410 million) records of PASMO, and have developed a visualization system for examining passengers' flow every ten minutes. We show several findings obtained from our analysis and visualization. First, we describe verification of estimated transportation statistics in the case of opening a new line. Second, we depicts unexpected phenomena and their reasons that lead operators to construct solutions. Finally, we show the analysis of passengers' flow in the events such as train suspension by typhoon, which reveals problems of train operations that brought heavy congestion. These lessons would help to construct countermeasures for better transfer operations and services.

Keywords: PASMO, visualization, OD Data

1. Introduction

The Basic Act on Transport Policy *[1] was enacted on December 2013. It is aimed at "improving the lives of citizens through comprehensive and systematic implementation of policies on transport." In particular, it defines basic plans for measures and policies from the perspective of building cities and responding to major disasters. We believe, however, that there is a need to comprehend and analyze the status of use of railways and conduct investigative research needed in formulating policies in order to promote the implementation of this law.

Thus far, the Ministry of Land, Infrastructure, Transport and Tourism has been conducting an Urban Transportation Census as part of large-scale surveys on transport every five to ten years; with the census to be conducted in 2015. The Urban Transportation Census began in 1960, and is carried out by handing questionnaires and collecting responses from passengers in three major metropolitan areas, namely, Tokyo, Nagoya, and Osaka. The data obtained from this survey serve as important reference in the formulation of urban plans by municipal governments and of route and transport plans by transport operators.

The census, however, surveys information only for specific days. Moreover, since collation of results takes more than one year, it is difficult or impossible to get a picture of the latest situation.

To address these problems, transport operators have conducted their own studies based on magnetic tickets actually used by passengers. This method, however, did not provide information on what routes the passenger took, and only covered a period of a few days in a year, making it difficult to accurately comprehend the actual situation.

To date, transport operators have therefore depended only on experience and intuition to carry out improvements in operations and services. However, in response to the shift from quantity to quality in public transportation demands by society, they need to be able to make prompt and detailed decisions pertaining to route and transportation plans based on user needs. This calls for the use of digital data and ICT to be able to clearly describe actual passenger flow, hasten management decision-making, and improve their operations and services.

The introduction of smart cards for transport in the Tokyo metropolitan area in March 2007, however, has drastically changed the situation in the transportation industry. Since then, mechanisms have been

*[1] http://www.mlit.go.jp/common/001096409.pdf

developed to collect passenger records from smart cards, understand changes and trends in passenger movement, and promptly reflect them in urban and transportation plans.

In this paper, we discuss the possibilities for improving Tokyo Metro operations and passenger services using large-scale passenger records from PASMO, a common passenger card system for railways and buses in the Tokyo metropolitan area, to address the problems discussed above. We have developed a system for analyzing passenger movements using PASMO data since four years ago. The system we developed collects and analyzes PASMO data to visualize passenger movements between stations at ten-minute intervals. We analyzed several case studies and presented measures that are useful for improving operations and services.

First, we quantitatively determined and analyzed passenger flow at the start of direct through-services following the opening of new lines. Next, we analyzed the gaps between the actual passenger volume at certain transfer stations from that forecasted by the operator and determined the causes for the gaps. Finally, we analyzed passenger movements during big events and similar phenomena using a case wherein train services were partially suspended due to a typhoon and a case wherein a public parade was held, to discuss issues that arise upon occurrence of similar phenomena and present possibilities for service improvement.

The paper is organized as follows. Chapter 2 provides a review of related research. In Chapter 3, we explain the significance of this research based on PASMO data. In Chapter 4, we present an overview of PASMO and the data generated during its use. In Chapter 5, we explain the outline of the analysis platform used in this study. In Chapter 6, we present the case studies we looked into, and in Chapter 7, we state our conclusions.

2. Related Work

Pelletier et al. [1], one of the first to focus on the use of smart card data in the formulation of public transportation policies, systematically presented case studies for utilizing smart card data by object-oriented modeling of transportation systems. Their study is considered as the first step towards utilizing smart card data for improving transport operations. Sun et al. [2] conducted trials to extract passenger spatio-temporal density and train trajectory from smart card data for subway in Singapore and succeeded in predicting the time passengers spend in stations and the movement trajectory of trains. Caepa et al. [3] conducted a similar study for London subway. Their study, however, was limited to a single train line and cannot be applied to the entire subway network, unlike our study, wherein we succeeded in visualizing the nine complexly intersecting Tokyo Metro lines. Ushida et al. [4] on the other hand, used Tokyo Metro traffic records rather than smart card data to visualize status of delays and develop a method for improvement. Their study suggests that the combination of smart card data and traffic records can be used to further improve operations. Meanwhile, Shimizu [5] used PASMO bus boarding history to obtain information on road travel times in Tokyo and showed the possibility of its use for improving bus operation schedules. An issue that must be looked into in preparation for the 2020 Tokyo Olympics and Paralympics is the creation of a seamless transportation system that integrates trains and buses, which points to the need for linking bus and train data. It was from this perspective that Zeng et al. [6] carried out a study to predict travel times for Singapore subway that include the bus routes for accessing the train stations. Tsunoda et al. [7] used smart card data to quantify the effects of disturbances in train operations to users and studied how they can be used to improve services.

The above studies carried out limited aspects of analysis in terms of operations during normal or abnormal times respectively. Our goal, however, is to overcome the effects of such limitations.

3. Significance of Analysis using PASMO Data

Transport operators have predicted passenger demands based on surveys such as the Urban Transportation Census in formulating train routes and transportation plans. These surveys, however, took more than one year to consolidate data, making it difficult or impossible to comprehend the latest situation.

We therefore analyzed passenger records from PASMO smart cards in an attempt to address problems that could not be solved using conventional methods. PASMO data are mainly used for fare calculation and settlement, and there have been no previous studies aimed at utilizing the data to improve operations and services from the operator’s perspective. Also, no previous attempts have been made so far to apply actual data from complex subway networks to the improvement of services and operations.

4. Fare Collection System using Smart Cards

Adding e-money functions to smart cards has automated fare collection and settlement operations previously done using magnetic tickets, leading to increased convenience for passengers. And the use of train-use data generated through this fare collection system enables comprehending passenger flow. Further, analysis and visualization of passenger flow provide a means to infer passenger behavior and the factors causing the behavior.

4.1 PASMO

PASMO was introduced in March 2007 as a common passenger card system for railways and buses in the Tokyo metropolitan area, including the Tokyo Metro lines. On March 2013, mutual use of smart cards from ten transport operators in major cities around the country became possible, enabling seamless operations that allow passengers to use a single PASMO card in any major city throughout Japan. Currently, there are 30 million PASMO cards in circulation, totally more than 80 million smart cards are interoperable in Japan including JR East and other operators' cards.

Tokyo Metro manages nine lines covering a total of 195.1 km and transports 6.84 million passengers a day (as of March 2015), serving as an important means of transportation in the Tokyo metropolitan area. *[2] In this research, we conducted case studies using PASMO data (410 million transactions) for Tokyo Metro and its connecting lines covering a four-year period.

4.2 PASMO Transactions

A PASMO card is assigned a unique ID. Although PASMO cards can be used as either named (registered to user's name) or unnamed cards, they are managed based on their ID number. When users pass through automatic ticket gates or charge their cards at ticket vending machines, information such as train station, time, and amount paid are recorded as part of card use history. These data are stored both in the PASMO card and at the PASMO center servers.

For example, for a passenger that used a card with a charge of 2,000 yen to get on a train at Shibuya at 6 pm on November 30, 2015, then got off at Ginza station, charged his or her card with 1,000 yen, got on again, and finally got off at Shibuya station, the following data items are generated for the five transactions:

① 20151130,180000,Shibuya,get on
② 20151130,181700,Ginza,get off,-195,1805
③ 20151130,181800,Ginza,charge,1000,2805
④ 20151130,182000,Ginza,get on
⑤ 20151130,183700,Shibuya,get off,-195,2610

Among the data stored in the PASMO center servers, those pertaining to Tokyo Metro lines are transmitted to Tokyo Metro servers.

5. PASMO Data Analysis Platform

5.1 Overview of Analysis Platform

We have built a system for collecting and analyzing PASMO data spanning more than four years (starting 2011).

First, transaction data were linked using the card ID to create an Origin-Destination data (OD data). Due to privacy considerations, we used data from unnamed cards, and randomly changed ID numbers as necessary. Use of OD data enabled a detailed understanding of transportation demands at much higher accuracies than information obtained from traditional censuses and surveys. It also enabled analyzing gaps between forecasted and actual passenger flows. Moreover, the availability of information on usage status also makes it possible to analyze the status of card refunds and other transactions, which is beyond the scope of this study. Next, OD data were then interpreted as flows, since there are a variety of possible routes between the Origin and Destination. The route used by the passengers, however, cannot be determined if they transferred lines without going through ticket gates. To address this problem, we inferred the routes based on the expected time to travel from origin to destination. We create a database for the time needed to travel between adjacent stations and the time to make transfers at each station, and compute shortest path for arbitrary origin-destination stations. We assume that all passengers will travel with this shortest path. This enabled us to determine the number of persons using a particular link – a section between adjacent stations – at ten-minute intervals based on the flow inferred from OD data. This system serves as a useful tool for analyzing passenger flow when there is a concentration of passengers during events and other extraordinary situations [9], [10].

5.2 Visualization

To facilitate analysis, we developed two methods for visualization, namely the heat map view (Figure 1) and the animated ribbon view (Figure 2). [9]

In the heat map view, it is possible to identify the line and time in which an unusual passenger flow occurs. This method is suitable for obtaining an overview of movements over a period of one or a few days. It expresses the data as a 2D map, with the train station on the vertical and time on the horizontal axis. To more clearly show the volume of passengers, we calculated the average and the distribution of persons moving between stations and denoted time intervals with volume of passengers higher than the average as red points and those with volumes lower than the average as blue points, using color variations to depict deviations from the average.

On the other hand, the animated ribbon view expresses the flow of users on a map for each station or line at a certain point in time, and is therefore suitable for looking at trends spanning a few minutes or hours. It shows a 3D map to express passenger flow over the train route map, and expresses deviations from normal through time using variations in color. Red represents high, while blue represents low volume of passengers.

*[2] http://www.tokyometro.jp/en/index.html

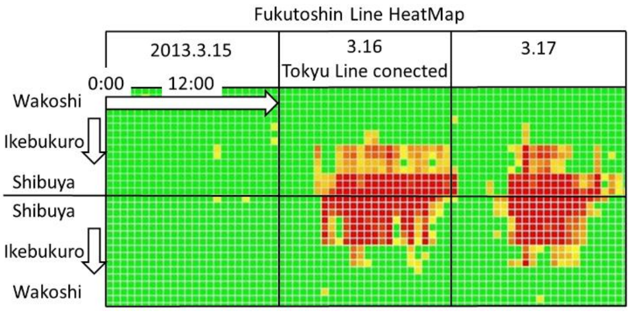

Figure-1 example of Heat Map View

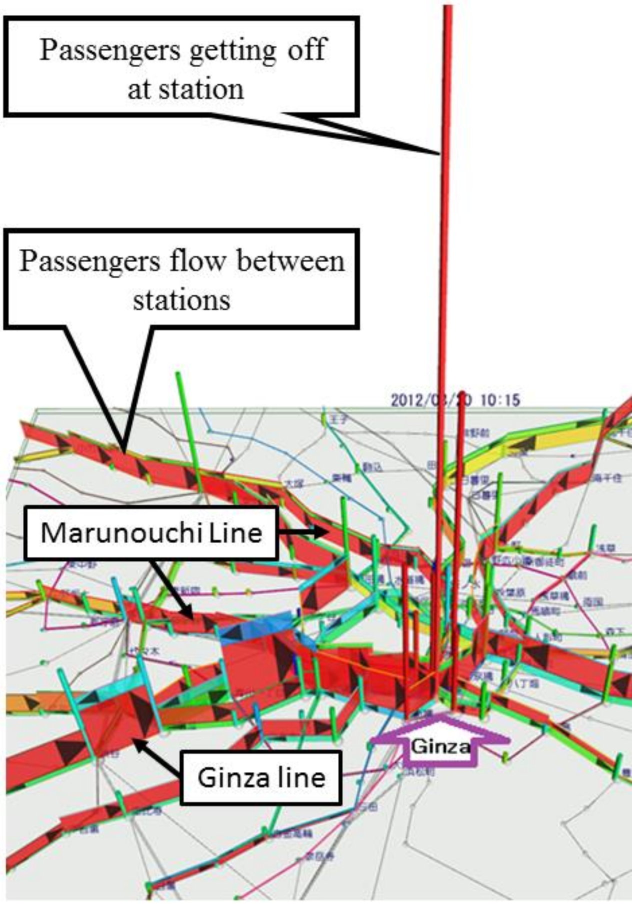

Figure-2 example of Animated Ribbon View

We looked into the following case studies using the heat map and animated ribbon view.

6. Case Studies

6.1 Increase in users as a Result of Direct Through-Services between Fukutoshin Line and Tokyu Toyoko Line on March 16, 2013

We determined the effect of direct through-services between Fukutoshin Line and Tokyu Toyoko Line at Shibuya station by visualizing the extent of increase in users before and after the start of the direct through operations. Figure 1 shows an increase in users especially between Shibuya and Ikebukuro (red areas). We were able to quantitatively determine that users previously changing trains for Yamanote or other lines at Shibuya Station instead took a train without changing at Shibuya station all the way to Shinjuku or Ikebukuro.

We were thus able to quantitatively measure the benefit to users in terms of reducing the inconvenience of having to change trains. We also demonstrated that demands arise from improving convenience by providing direct through-services between lines without a new line construction.

6.2 Analysis of Passenger Behavior during Transfers at Shinjuku-Sanchome Station

There are 2 transfer routes between Fukutoshin line and Marunouchi line at Shinjuku-sanchome station. (a) One is direct through from Fukutoshin line platform to Marunouchi line platform. (b) The other is that once he/she gets off Fukutoshin line and again gets on Marunouchi line through concourse. Former is 1 trip but latter is 2 trip so that fare is different (Figure 3). Complaints, however, have been received from passengers who mistakenly get out of the gates when changing lines.

We therefore tried to analyze passenger behavior during transfers using PASMO data. We extracted passengers' records getting off at Shinjuku-sanchome station and getting on the same station of the other line within 30 minutes from passengers records selecting route (b). We regarded the numbers of extracted passengers' records as passengers getting off by mistake

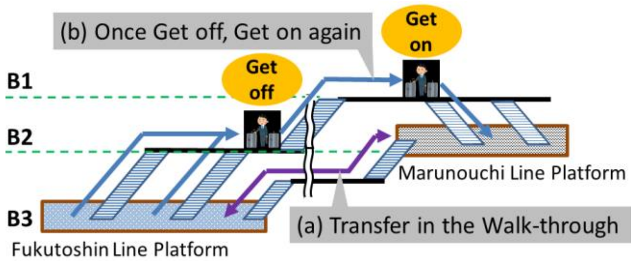

Figure- 3 Transfer at Shinjuku-sanchome station

Before PASMO cards were introduced, passengers who mistakenly got off had their tickets collected at the gates and had to buy a new ticket to get on. It was therefore not possible to monitor their movements. Use of PASMO data, however, allowed us to determine the actual number of passengers getting off by mistake and getting on again.

Our results showed that although there was a large number of passengers who mistakenly get off the station in changing lines from Fukutoshin Line to Marunouchi Line, there were few passengers making mistakes in changing lines going the opposite direction. This disparity is attributed to the confusing structure of the Fukutoshin Line platform, because there are three ascending routes from Fukutoshin Line platform to Marunouchi Line platform. On the other hand path from Marunouchi Line to Fukutoshin Line is a single descending walk-through.

This shows the importance of designing more user-friendly facilities by considering the location of stairs during the renovation or construction of new stations.

Last year, a new walk-through for changing lines was built in Shinjuku-sanchome Station, for which the above data were used in the design to investigate the number of fare gates to install.

6.3 Passengers Flow in case of Disaster or Big Event

6.3.1 Analysis of User Behavior during Suspension of Tozai Line Services due to Typhoon No. 17 on September 30, 2012

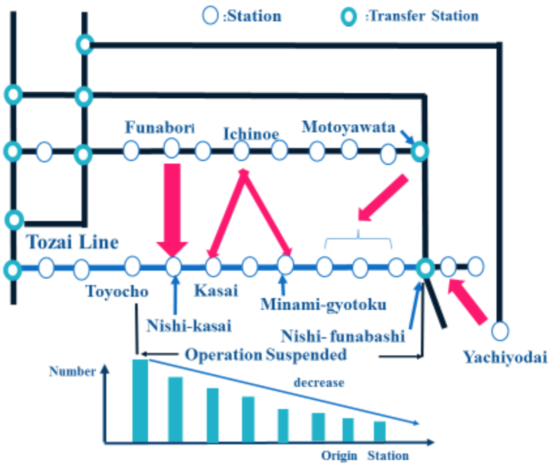

Figure -4 Passenger flow at train suspension by typhoon

The Tokyo Metro Tozai Line (blue line in Figure 4) is a long subway line that includes aboveground tracks and bridges going over several rivers. For this reason, services are usually suspended when risks of derailment arise due to strong winds or tornados brought about by approaching typhoons or low-pressure. When this happens, passengers either have to wait until services resume, or use other means to reach their destination, such as using alternate train routes and stations and then riding buses or taxis. To date, however, no studies have been made on the choices that passengers make on such occasions. We therefore used PASMO data to identify passengers that appear to have not reached their destination and investigated the circumstances behind the choices they make during such incidents.

First, we assumed that the station where passengers got on in the morning is their point of origin. Next, we identified and analyzed data for passengers who made a final stop at night at a station different from their station of origin, regarding them as having done so as a result of train suspension or other unavoidable circumstances.

Figure 4 shows results of analysis of data during an almost-four-hour suspension of train services starting 8 pm between Toyocho and Nishi-Funabashi stations due to an approaching typhoon in 2012. Pink arrows point to the stations of origin on the Tozai Line in the morning for passengers who got off at a station on another train line (black line) at night, with the thickness of the arrows reflecting the number of the passengers. Passengers getting off at another station were assumed to have taken buses from that station to get home. The bar graph at the bottom shows the number of passengers departing from each station between Toyocho and Nishi-Funabashi who got off at Toyocho and did not transfer to another line or get off at another station. These passengers were assumed to have either walked or taken the taxi from Toyocho Station, in the absence of other means of public transportation, as shown by the decreasing number of passengers coming from stations farther from Toyocho Station.

Our analysis revealed important results pointing to the importance of providing information on alternative means of transportation during suspension of services, in consideration of the expected time for resumption of operations. They also point to the need for collaborating with bus operators during such incidents. Since smart cards can be used also in buses, a future issue to investigate is the use of bus boarding history to analyze passenger behavior and demands.

6.3.2 Analysis of the Massive Concentration of Passengers related to the Victory Parade for the London Olympics on August 20, 2012

An unusual flow of passengers was observed around Ginza Station during the London Olympic parade held at Ginza. As shown in Figure 2, there was a concentration of passengers using the Ginza and Marunouchi Lines and getting off at Ginza Station starting from around 9 am, two hours before the start of the parade. On the other hand, as shown in Figure 5, an increased flow in the opposite direction was observed during the one-hour period from the start to the end of the parade. This indicates that there were marked increases in movement of passengers around Ginza at particular time intervals. These results were obtained by visualizing data at small intervals of 10 minutes.

Passengers getting on from Ginza travelled towards Shibuya, Shinjuku, Ikebukuro, and Ueno as their final destinations. Since Ginza Station is located at the center of Tokyo and functions as an important hub that connects to Shibuya, Shinjuku, Ikebukuro, and Ueno, which serve as hubs for suburban railways. In particular, lines serving many passengers on regular days showed a much higher increase in passenger volume. Presently, the Ginza Line is undergoing renovations to improve the branding of the line. Like the Ginza Line, Marunouchi Line, which includes Ikebukuro, Shinjuku, and other major stations, also has a high brand value, and both lines are considered to be high priorities for future investments. In particular, these two lines are foreseen to cater to a large volume of local and foreign tourists during the 2020 Tokyo Olympics.

We therefore believe that increasing the name value of these two lines by providing multilingual support and promoting them as convenient train networks for tourists would be an effective approach to cater to future user needs.

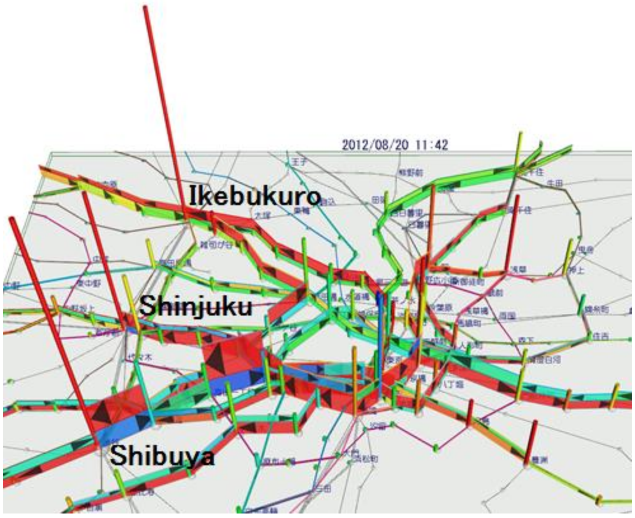

Figure -5 Animated Ribbon View after the parade

7. Conclusion

It is essential for transport operators to first correctly comprehend passenger demands before they can offer smooth means of transportation, which is the goal of the Basic Act on Transport Policy.

We used a system for collecting, storing, analyzing, and visualizing PASMO smart card data to obtain knowledge that could not be elucidated using conventional methods.

First, using our system, we were able to understand the detailed dynamics of passenger flow during the opening of new lines and stations, providing us with information useful in developing future services and in improving their convenience. Second, we were able to analyze detailed data on the differences between operator's forecasts and actual passenger behavior, and devise methods to resolve those differences. Finally, we were able to examine passenger behavior and demands during likely future events and phenomena, such as typhoons and public events, and present effective measures in dealing with similar occurrences in the future.

Our method is useful also in formulating plans in preparation for the 2020 Tokyo Olympics and Paralympics and in devising recovery plans for the near-field earthquake predicted to occur in the Tokyo metropolitan area. We believe that this method will serve as a useful tool in achieving the goals of the Basic Act on Transport Policy, namely, building of safe and convenient cities and formulating measures against disasters, as well as contribute to the improvement of operations and sustainable services.

8. References

[1] M.-P. Pelletier, M. Trepatier, C. Morency, "Smart Card Data Use in Public Transit: A Literature Review," transportation research Part C: Emerging Technologies, vol. 19, no. 4, pp. 557-568, 2011

[2] L. Sun, D.-H. Lee, A. Erath, and X. Huang, "Using Smart Card Data to Extract Passenger's Spatio-Temporal Density and Train's Trajectory of MRT System," in Proc. UrbComp'12, 2012, pp.142-148

[3] I. Ceapa, C. Smith, and L. Capra, "Avoiding the Crowds: Understanding Tube Station Congestion Patterns from trip Data," in Proc. UrbComp'12, 2012, pp. 134-141

[4] K. Ushida, N. Tomii, "Increasing Robustness of Dense Timetables by Visualization of Traffic Record Data and Monte Carlo Simulation, WCRR, 2011

[5]E. Shimizu "Visualization of duration by time-map in city road transport," 2011 Institute for Transport Policy Studies (in Japanese)

[6] W. Zeng, C. Fu, S. M. Arisona, A. Erath, and H. Qu, "Visualizing Mobility of Public Transportation System" IEEE Trans. Vis. Comput. Graph, vol.20, no.12, pp. 1833-1842, 2014

[7] F. Tsunoda, M.Yamashiro, R. Otsuka, M. Kato,H. Sukeda, K. Ozeki, "Customer-Oriented Evaluation Method of Railway Performance" 6th International Conference on Railway Operations Modeling and Analysis-RailTokyo 2015, Narashino, Japan, 2015

[8] I. Ceapa, C. Smith, and L. Capra, "Mining Mobility Data to Minimise Travellers' spending on Public Transport," in Proc. of KDD'11, 2011, pp1181-1189

[9] M. Itoh, D. Yokoyama, M. Toyoda, Y. Tomita, S. Kawamura, M. Kitsuregawa, "Visual Fusion of Mega-City Big Data: An application to Traffic and Tweets Data Analysis of Metro Passengers" in Proc. IEEE Big Data 2014, pp431-440

[10] D. Yokoyama, M. Itoh, M. Toyoda, Y. Tomita, S. Kawamura, M. Kitsuregawa, "A Framework for Large-Scale Train Trip Record Analysis and Its Application to Passengers Flow Prediction after Train Accidents", PAKDD 2014, pp533-544