Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 52 (2015) 178 – 185

Procedia
Computer Science

The 6th International Conference on Ambient Systems, Networks and Technologies
(ANT 2015)

Public Transportation Service Evaluations Utilizing Seoul Transportation Card Data

Ji-Young Song[a], Jin Ki Eom*[a], Kwang Sub Lee[a], Jae Hong Min[a], Keun Yul Yang[a]

a Korea Railroad Research Institute, Ui-Wang, Kyung-Ki, 437-050, South Korea

---

Abstract

This study evaluated transit service performance in Seoul using data collected from the Automatic Fare Collection (AFC) system in Seoul. The distance-based fare system in Seoul allows a maximum of four transfers with no additional charges to encourage transit ridership. In order to analyze the transit transfers, this study developed quantitative indicators for public transportation evaluations differentiated from those of previous studies by the fact that it utilizes data mining techniques which incorporate massive amounts of data (over 10 million transits per day) derived from the smart card system. This study not only carried out an evaluation to improve public transportation quality but provided comparative analysis of the mobility handicapped and an evaluation of public transportation users' regional equity. This evaluative analysis of Level of Services (LOS) for various items is expected to be adopted for analyzing LOS status and generating improvement priorities and to be utilized as an objective database for public transportation policy decisions.

Peer-review under responsibility of the Conference Program Chairs

Keywords: Evaluarion indicators, Level of service, Transit smart card data, K-means clustering

---

1. Introduction

The transportation system is being reorganized to prioritize the public transportation and walking in order to improve public transportation travelers' convenience as well as to reduce energy consumption. The government,

---

* Corresponding author. Tel.: +82-31-460-5467; fax: +82-31-460-5022.
E-mail address: jkom00@krri.re.kr

1877-0509
Peer-review under responsibility of the Conference Program Chairs
doi:10.1016/j.procs.2015.05.053

especially, established a master plan focusing on building an urban-railway-oriented public traffic network in order to promote public transportation utilization and to systematically support public transportation. Owing to these efforts by government, the urban traffic system is expected to be reorganized with public transportation as the focus. Despite the introduction of a variety of public transit priority policies, the utilization of individual transportation has been increasing. This is because the problem areas of the service have not been resolved.

It is required to closely review the local connections by analyzing the accessibility and transferability of the public transportation service in order to increase public transportation utilization. However, previous research has neither properly considered the service evaluation items and indicators nor diversified the evaluation items and data in accordance with the changes in the transportation environment. Additionally, the previously used data was not easy to obtain so the studies were mainly based on sample surveys and focused on the individual line networks with partial traveler data which limited their accuracy.

In 2007, an integrated payment system was introduced into the Seoul capital area transportation system, accompanied with Smart Cards, which store an individual's transfer data, allow for the collection of observational complete enumeration data on travellers and makes it possible to overcome the limitation to a single line in previous studies, allowing for integrated analysis of the entire transportation network. Thus, it is possible and necessary to analyse travel patterns using this information. The analysis also needs to be extended to the travel of senior citizens and the disabled, who are under movement constraints, as well as to normal travellers.

This study was done to take advantage of the newly available data through data mining techniques to develop new evaluation criteria and use a service quality evaluation model considering the characteristics of service satisfaction to generate improvement items. The public transportation service evaluation model developed via this study analyses the data of individual travellers so that the regional service can be evaluated. Furthermore, it will be able to be utilized as an objective database for the practical considerations of policies such as service status analysis and improvement elicitation.

2. Data

The data used in this research includes the transportation card data from Seoul, Gyeonggi province, and Incheon in 2010, provided by the Korea Smart card system. This research used data on the passengers who used 482 stations on 12 lines of metropolitan subway and over 15,000 stops on 405 bus lines located in the city of Seoul for an entire day. The total transportation card data consists of 22 columns of attributes and 153 digits of width including transit data on times, spaces, and methods.

The integrated distance proportion fare system of the public transportation system in the Seoul capital area calculates the basic fare when boarding and charges additional fares when departing depending upon the distance traveled. The fare is calculated in proportion to the distance regardless of the number of transfers (max 4) between buses and trains. This calculation method is available only when boarding with a transportation card, and the integrated fare is applied if transfers are completed within 30 minutes (between 07:00 and 21:00) or 1 hour (between 21:00 and 07:00).

According to the ratio of whole-day traffic volumes of public transportation systems based on transportation modes and the number of transfers derived from the smart card data, the integrated proportion fare system allows a maximum of 4 transfers. The transfer combinations of bus and subway were analyzed into a total of 54 types including bus only, subway only, and bus-subway-bus-subway-bus. The cases of using bus or subway only and 1 or 2 transfers account for over 99% of the total trips, and trips with more than 3 transfers account for the remaining 1%. The trips with more than 3 transfers are a minority of the total and only a few trips with certain combinations were found, so the characteristic of the indicators determined whether or not to include them. Interline transfers are only recorded for 2 of the 12 metropolitan subway lines. This research does not consider transfers occurring between subways since estimating the transfers on the rest of lines may weaken the integrity of the observational complete enumeration data.

3. Service evaluation indicator

Various studies have been conducted in regard to the quality of public transportation services. The city of Seoul has continued surveys of passenger satisfaction concerning its public transportation system since 2006. However, most of them are focused on the assessment of suppliers or facilities. There was no research for new indicators that reflect the changes since the reformation of the public transportation system. Indicators for transfer system functionality, transfer convenience, mobility, etc. that passengers perceive has not been developed.

There are currently no criteria for service evaluation on transfer times, transfer convenience, mobility, and equity between the two modes (subway-bus), so this research classified the levels of service into 6 levels from A to F through clustering. The evaluation of transfer times was calculated using data from individual trips and the levels of service level were categorized based on the individual passenger data collected, while the transfer convenience and mobility data were collected by Origin-Destination (O/D) in borough units and LOSs were categorized based on the O/D in borough units. To calculate the equity indicator, LOSs were divided collecting the travel data by borough as the income levels in borough unit was used. Figure 1 shows traffic analysis zone (TAZ) ID, names and locations of borough in Seoul for presenting results of analysis.

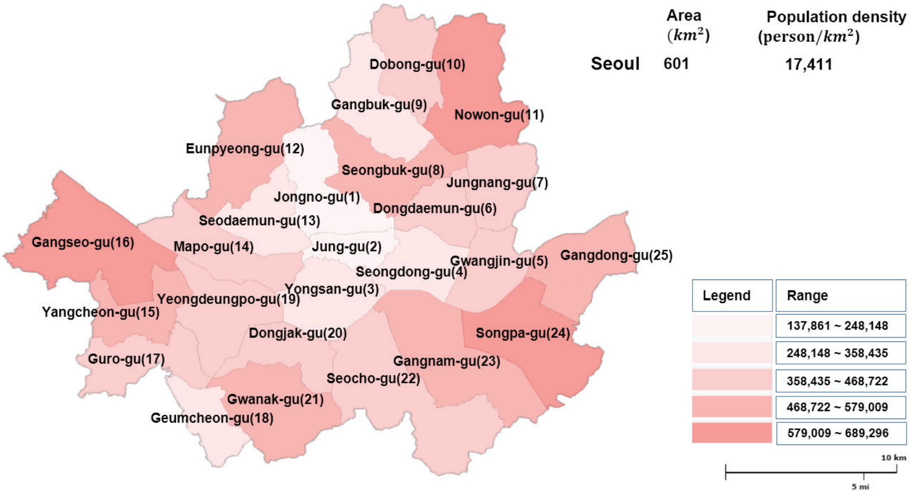


Fig.1 Map of Borough 'Gu' in Seoul

The evaluation data was classified into levels by employing the clustering method to establish the criteria for each indicator. A clustering method is one key technique and uses an algorithm to group large amounts of data for each cluster to include similar data units. It finds the best cluster model through an iterative refining process. In particular, the k-means algorithm uses the Euclidean distance calculation method for calculating the distance between two data units, and ensures the cluster totals k in number in order to minimize the value calculated by the square error function. Thus this method is relatively efficient for dealing with large sets of data, of which the algorithm complexity is O (nkt), the data units being n in number, the clusters k, and iteration counts t. This study takes advantage of this efficiency and extensibility. The use of the public transportation card data has enabled a wider analysis including a time series technique. This study presents an evaluation fit for the new public transport system that overcomes current limitations by using smart card data as source data.

3.1. Transit Attribute Evaluation Indicators

The total transit time of passengers using public transportation includes access time, in-vehicle transit time, transfer time, transfer transit time, dwell time etc. For this study, the access information was not considered because it is unobtainable from the transit smart card data. Transfer access time, transfer transit time, dwell time data are obtained from transaction information of transit smart card data. To calculate total transit time summarized in-vehicle transit time and transfer time.

$$T^{T} = T^{IVT} + T^{TF} \tag{1}$$

Where,
T^{T}-Total transit time
T^{IVT} – In vehicle transit time
T^{TF} – Time for transfer between transit modes

This analyzed the transit information from individual passengers to calculate the total transit, and aggregated by borough 'Gu'.

In the case of the subway tags time of the traffic card is a point of entry to the platform or station. The waiting time until boarding the subway cars is included in the tag time. Thus tagging time include access walking time, in vehicle transit time, egress walking time.

In contrast, for the bus tag, it is during and after boarding of the vehicle. There are some cases that the passengers' tag before getting off the vehicle, while some cases, the running time is excluded, because the tag time is included not only in-vehicle transit time but the dwell time at the arrival station recurrently.

Although errors may exist of the tag time, in this analysis travel time is calculated based on the tag time (in minutes), as shows:
T^{IVT} – In vehicle transit time, minutes

Travel distance is the transit distance between the subway stations between the bus stops. The in-between of subway stations were analyzed by the standard measured in accordance with the locations of each station. The distance between the bus stops analyzed on the base of measurement of geographic information each bus stops of bus routes. The distance passes between the same stations are also analyzed by various routes. Unit of travel distance m and symbol is as follows.
D_{ij}^{T}- The average of travel distance of all passengers from origin i to destination j

Transit speed is measured by dividing the distance by the transit time. Transfer time is not suitable for measuring the speed that presents in the idea of space, rather than the concept of distance and includes moving and waiting time. Thus the speed measurement is by dividing the distance by the in-vehicle transit time. Unit is m/min., equation is as follows:

$$S_{ij}^{T} = \frac{D_{ij}^{T}}{T_{ij}^{IVT}} \tag{2}$$

Where,
S_{ij}^{T}- The speed of transit of all passengers from origin i to destination j
T_{ij}^{IVT}- In vehicle transit time of all passengers from origin i to destination j
D_{ij}^{T}- The average of travel distance of all passengers from origin i to destination j

In addition, the difference of traffic speed was significant hourly and by traffic conditions. The difference was found significant between average speed Of O/D during the morning peaks and the entire day in Seoul.

In addition, the introduction of new routes and facilities are confirmed with an improved rate of transit speed. Thus transit speed is a useful indicator for service analysis of public transport passenger.

The number count of passengers using the transportation card gives the information aggregated by O/D of borough 'Gu'. Also, it obtained and utilized the card information the numbers of passengers boarding together by one transit card. The unit is thousand people, and symbol is as follows:
n_{ij}-the number of passengers using public transportation from origin i to destination j

3.2. Transfer Attribute Evaluation Indicators

There are two definitions of linked transfer. First, the transfer route from the station to the bus stop (facility size). Second, the waiting time for the bus at the bus stop (linked transfer).

Previous studies measured the transfer time by actually investigating the number of steps, elevators, and the transfer distance, etc, but this research easily deducted the transfer time by extracting the time from the drop-off at station i to the boarding time at stop i' from the smart card data. It is calculated as follows.

$$T_i^{TF} = T_{ii'}^{move} + T_{ii'}^{wait} \tag{3}$$

Where,
T_i^{TF} – Time for transfer between transit modes at station i, minutes
$T_{ii'}^{move}$ – Moving time between station i and stop i'
$T_{ii'}^{wait}$ – Waiting time at stop i'

The total transfer time including the moving time required to transfer by the facility size of the station and waiting time which can be considered of the interoperability between the two modes is referred to as linked transfer time.

Transfer time is defined as the time difference between the transportation card tag for getting off the first vehicle and the tag for getting on the second vehicle when traveling from an origin to a destination by means of public transportation. In a smart card, travel time data is stored when it is touched to a card reader, in which the transfer time can be obtained. For this part, the analyzed transit information from individual passengers to transfer time were calculated between the railway station and bus stops, and aggregated by borough ‘Gu’. Average number of transfers were obtained from dividing sum of total transit by number of passengers. As shown below:

N^{TF}-number of transfers in one travel

Time taken per transfer is found by calculating the ratio of travel time to a variable of the transfer volume of all the travelers who transferred once or more within each travel section and averaging them for each borough. For the calculation, see the following equation:

$$\propto_{ij} = \frac{(\Sigma_n {T^{TF}}_n / {N^{TF}}_n)_{ij}}{n_{ij}} \tag{4}$$

Where,
$\propto_{ij}$ - Average time taken per transfer of all passengers from an origin, i to a destination, j
${T^{TF}}_n$- Transfer times of all the n passenger from an origin, i to a destination, j
${N^{TF}}_n$- Number of transfers of n passenger from an origin, i to a destination, j
n_{ij}- Number of passengers from an origin, i to a destination, j

By multiplying the number of passengers from an origin to a destination using their smart cards by the number of transfers of each of the passengers and then dividing the transfer time used in the first equation by the result, the value of a transit time taken per transfer can be obtained. It is an indicator for classifying transit problems in transfer trip, regardless of the numbers of transfers and travelers.

The former is considered to have a lot of subway transfer stations and an efficient transfer system was developed, although long-distance travels are frequent in the sector, while the latter has traffic congestion due to poor transfer facilities and downtown is located in the area. Because this indicates regardless of transfer frequency and the number of transferring people, that it is particularly useful to evaluate the convenience of long transit sectors where multiple transfers are frequent.

4. Application of the Evaluation Model of The City of Seoul

In the previous sub-chapter presented various evaluation methods with the analysis of the transit attribute, transfer property and equity. The Level of Services (LOS) for the City of Seoul was set by selecting the indicators that are meaningful to set the standard and evaluate transit performance. The evaluation model for public transportation service using smart card data, developed in this paper, may extend its analysis area by changing the counting method such as evaluations by unit sections and types using various items. This evaluation model was applied to the travelers who used 482 stations on 12 lines of metropolitan subway and 15,000 stops on 405 bus lines in the city of Seoul. To apply this model, relevant data was provided from Korea Smart Card Co., Ltd. for the use of research. This data had been collected for fare calculation between operators. The transportation card data has the same structure as the data used for the development of the evaluation items. The evaluation model was applied to normal travelers and the transportation vulnerable respectively. The total number of cases used for the total transit time evaluation indicator analysis was 1,845,860 and they were counted by each of the 25 boroughs of Seoul. For counting, the travel data of 600 O/Ds between boroughs were used, and intra-borough travel data was excluded. Using an SAS K-means clustering method based on the data of each borough, the LOS are defined from A to F. Total transit time indicator in the table 1, the largest amount of data belongs to LOS B and A. From LOS B, the lower the LOS is, the less data is available.

Table 1 LOS Classification of Total Transit Time Indicator, Transit Speed Evaluation Indicator, Number of Transfers Evaluation Indicator by Clustering

LOS	Total Transit Time Evaluation Indicator		Transit Speed Evaluation Indicator		Evaluation Indicator of Time Taken per Transfer	
	number of clustering	Indicator	number of clustering	Indicator	number of clustering	Indicator
A	143	0 ~ 32.31	38	411.95 ~ ∞	2	0 ~ 2.4026
B	156	32.31 ~ 44.46	92	377.42 ~ 411.94	441	2.4027 ~ 4.4855
C	140	44.46 ~ 55.43	139	345.51 ~ 377.41	144	4.4856 ~ 6.2556
D	108	55.43 ~ 68.39	158	311.23 ~ 345.5	10	6.2557 ~ 8.2357
E	52	68.39 ~ 92.51	124	274.25 ~ 311.22	1	8.2358 ~ 10.8428
F	1	91.51 ~ ∞	49	0 ~ 274.24	2	10.8429 ~ ∞

5. Validation of Evaluation Indicators

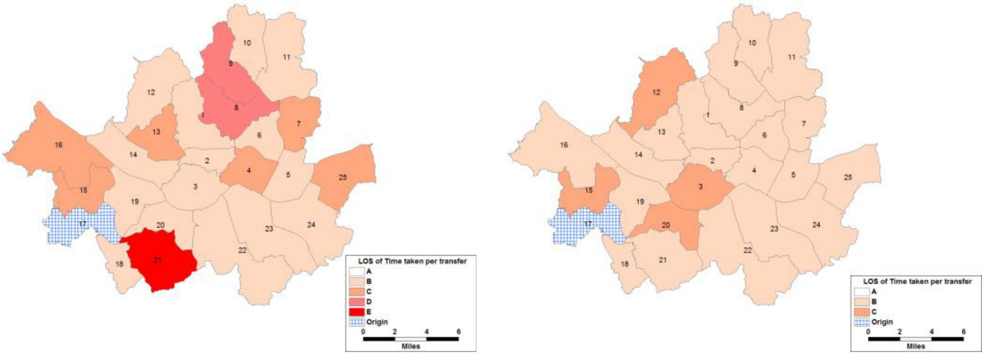
This chapter utilizes them directly in the evaluation of public traffic systems, and overviews the changes in the evaluation values caused by changes in public traffic systems. By the use of indicators, it intends to evaluate the changes in public traffic systems found in the analysis of LOS with facility openings, analysis of LOS with meteorological changes, and analysis of LOS with an aging population.

At Guro Digital Complex Station's transit center, the transit movement distance was reduced from 60m to 10m by the removal of fences in it carried out in order to reduce inconvenience of transit between the subway and buses. In the center and the station, fences had been installed and operated inevitably in consideration of the safety of the users due to the operation of the transit parking lot.

Service improvement was numerically verified before and after repair of the facility through analysis of traffic evaluation indicators on April 18, 2011 (before fence demolition at the Transit Center of Station Guro Digital Complex), and after November 24, 2011.

Changes in LOS are found in the figure 2 LOS of transit time taken per transfer from Guro-Gu (ID17) to destination j before and after repairing transfer center. The transfer time from Guro-Gu to boroughs in Seoul was

4.56 min per transfer before repairing the transfer center; this time after repair was 3.97 min, which is a difference of approximately 0.6 min. The transfer time from boroughs in Seoul to Guro-Gu was approximately 0.3 min. From Guro-Gu to Gwanak-Gu (ID21), the time per transfer improved more than 5 min, from 8.90 min to 3.57 min. The LOS of transit speed changed from Level B 31 to 40, Level C 14 to 8, and Levels E to F after repair.



(a) (b)

Fig. 2 LOS of Transit time taken per Transfer from Guro-Gu (ID17) to Destination j before (a), after(b) Repairing Transfer Center

Service improvements before and after opening Line 9 were analyzed numerically by evaluation indicators using transit card data from June 18, 2009 (before the opening of Line 9), and October 15, 2009 (after the opening of Line 9). Evaluation indicators of transit time and number of transfers were utilized for this analysis.

There were 732,434 transit card data values during the morning peak on June 18 2009, before opening Line 9; on October 15, 2009, there were 941,509. Thus, the number of trips increased. Figure 3 provides a map of metro lines in Seoul, which highlights Metro Line 9. This map shows the area through which Line 9 passes.

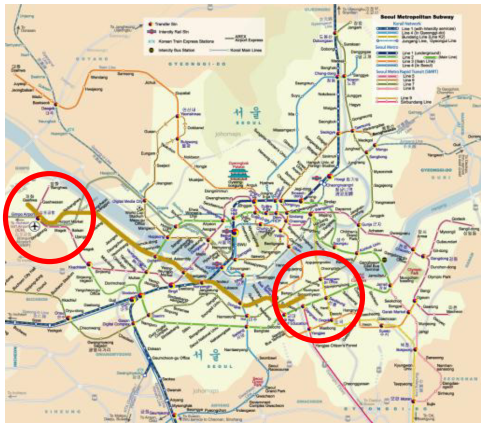
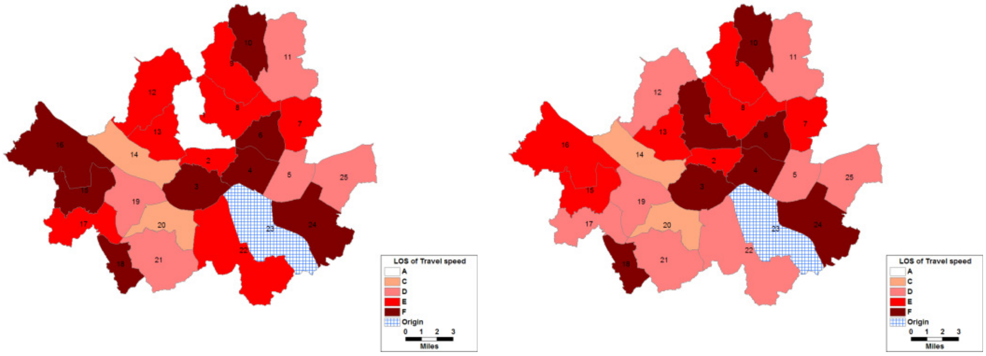


Fig.3 Map of Metro Lines in Seoul with the Highlights of Line Number 9

One objective behind opening Line 9 was improving the windiness of road at the start and end points; Gangnam-Gu (ID23) and Yangcheon-Gu (ID15) confirm the effect of speed improvement after the opening of Metro Line 9. By comparing the transit patterns of Gangnam-Gu (ID23) and Yangcheon-Gu (ID15) as the start and end points, and the north parts of Seoul as the routes connected by transfer with the other areas, the beneficial effects from opening Metro Line 9 are obvious.



(a) (b)

Fig. 4 LOS of Transit Speed from Gangnam-Gu (ID23) to Destination j before (a), after (b) Opening Metro Line Number 9

The LOS of O/D from Gangnam-Gu (ID23) and Yangcheon-Gu (ID15) was improved as a result of the increased traffic speed by Line 9. Unexpectedly, the greatest change interval was not the O/D above Line 9, but the O/D between Eunpyung-Gu (ID12) and Gangnam-Gu (ID23) because it enables seamless connection with other lines. The result was a speed boost outside the route. The regions that affected by the metro line 9, became more convenient to transit, but more regions tended to decrease the LOS after metro line 9 was opened. This is not to show the service deterioration after opening of metro line 9, but to reflect the system changing. Transit condition of the big system has lower level than that of small system. This system of LOS has this kind of limitation.

6. Conclusions and Implications

What makes this study distinguished from previous studies is its use of the public transportation card data for analyzing service levels, considering transfer for the first time in the area of public transportation. Previous studies mostly depend on qualitative analysis such as measuring satisfaction levels based on a questionnaire survey, or objective data collection and definition. Many changes made in public transportation, which include the introduction of an integrated fare system and the improvement of the transfer system, have required conducting the transportation service level analysis from users' view. This study used objective data in analyzing service levels, focused around transfer, and has obtained meaningful evaluation results.

This study presented a passenger service level analysis model, with regard to complex public transportation systems that have not been clearly analyzed in the previous service level analyses, and performed an evaluation on the service levels of Seoul.

The results of the service level evaluation analysis of various factors conducted using individual trip data suggest that this analysis approach can be used in the development of policies for public transport expansion, facility expansion to enhance convenience for the disabled and elderly, service route expansion plans, and current situation analysis of the service.

References

1. S. p. office, "Quality/Level of service handbook", 2002.
2. J. M. Samuels, "Performance Measures to Improve Transportation Systems and Agency Operations", Transportation Research Board, National Transit Institute, October 29–November 1, 2000 2000.
3. U. S. F. T. Administration, T. D. Corporation, M. International, T. C. R. Program, and C. Systematics*, A Handbook for Measuring Customer Satisfaction and Service Quality* vol. 47: Transportation Research Board, 1999.
4. H. P. Benn, *Bus route evaluation standards*, 1995.
5. Kittelson, Associates, U. S. F. T. Administration, T. C. R. Program, and T. D. Corporation, *Transit Capacity and Quality of Service Manual* vol. 100: Transportation Research Board, 2003.
6. H. S. Levinson, *Supervision strategies for improved reliability of bus routes*, 1991.