Istanbul’s BRT Trip Destination Estimation and Validation

Fahrettin Eldemir¹, Murat Akad², Aydin Deli³

¹ Department of Industrial Engineering, Yıldız Technical University, İstanbul, Turkey
² Department of Industrial Engineering, Istanbul University, İstanbul, Turkey
³ TÜBİTAK, Gebze, Kocaeli, İstanbul, 41401, aydindeli@tubitak.gov.tr

Abstract

The Bus Rapid Transit (BRT) system of Istanbul, located in the main artery of the city, is the mode of transportation allowing significantly faster trips. Among transportation modes in Istanbul, there is no mode serving more passengers than BRT systems do. However, due to high demand in rush hours, occupancy rates in the vehicles are exceeding capacity limits which in turn cause capacity crushes. While BRT systems are designed to carry 20,000 passengers per hour in one direction, in Istanbul BRT System in one direction 45,000 are transported. Limited stop services, express services and other solutions that will reduce occupancy rate require origin and destinations (OD) estimation of the trips in advance. For this reason, both passenger’s alighting station (destination) and boarding station should be known at the same time for successful planning and mentioned solutions.

Personal smart cards are used in Istanbul. The cards IDs and trip attributes are recorded with each trip. On the average, there are 6.1 million trips realized per weekday in Istanbul's public transportation system, where BRT trips have a significant percentage that is 14%, close to one million trips. The data regarding passengers’ origin station and boarding time are available. However trip destinations are not readily available. There are few methods commonly used to determine the destination stations. The accuracy of the most used OD matrix generation method is checked by two approaches:

1. In Istanbul’s BRT system distance based pricing policy is used. In this policy, number of stations from origin to destination are counted and the fee taken at origin are returned by refund machines at destination stations accordingly. This in return helps verifying the accuracy of OD matrix.

2. The trip records of the passengers those use bus system, metro or other modes of transportation after BRT system are investigated. From these records the boarding station of the next trip is taken. The nearest BRT station to the next trip’s boarding station is assumed as the destination station of BRT trip. This helps verifying the accuracy of OD matrix.

As a result, it is observed that a significant part of OD matrices namely around 95%, are accurate and they were able to be checked and verified by far more successful methods than classical survey methods.

Keywords: Public transportation, origin destination matrix (OD), smart card data, planning management, Bus rapid transit (BRT)
1. Introduction

Bagchi and White (2005) express that smartcard data can be used to analyze travel behavior, turnover rates and trip rates.

Smart cards have become the most important part of trip data after being introduced to public transportation. To execute better strategies on public transportation services, the trips of the users should be analyzed from smart card data. Information regarding the origins and boarding times of the trips are readily available in smart card data.

In addition to the information extracted from smart card data; some other information needs to be gathered such as trip destinations, transfer points, network connection points and trip times. (Ma et al., 2013). It is crucial to know the destination stations of the trips as well as the origin stations in planning of public transport and optimization of network (Li, 2009). However the destination stations are not available on most smart card systems.

Origin Destination matrix (OD) provides basic information about the passengers that travel from one geographical area to another geographical area and vehicle flow in traffic. In this aspect, OD matrix has a very critical role in traffic and transportation management (Li, 2009).

Several methods of creating the OD matrix are as follows;

- Surveys of passenger or households
- Counting passengers at stations
- Card reading systems or sensor at vehicle

There are several drawbacks associated with these methods. Household surveys are costly and time consuming (Ma et al., 2012). Counting passengers at the stations give information for limited time range. When the sensor option is considered, many buses are not equipped with the required technology. The best option is to use smart card systems. At the end of 1990’s, smart card payment systems started being used in some cities such as Washington and Tokyo. This new technology has performed a very rapid development and it became significant part of public transportation payment systems. Efficient Planning and operating of public transportation come true by smart cards (Munizaga and Palma, 2012). Besides, some practices such as special fee for express lines or special discount for off-peak hours are put in to practice (Trepanier, 2007).

Barry (2002) examined automated payment systems as known Metro card in New York Transportation and developed a methodology that estimate OD trips from station to station. The first algorithm consists of two basic assumptions:

1. Users begin to their next trips from the end station of their previous trip.
2. Users finish their last trip of the day at the station where they begin their first trip on that day. When theses suppositions are tested at metro users with surveys, they are verified at %90 level. Zhao et al (2007) developed OD matrix estimation algorithm from smart card data’s of Chicago Transit Authority rail systems where only trips’ origin stations were available. Trepainer (2007) proposed object oriented method that estimated destination station for bus transportation in Gatineau, Canada. He assumed “The station that begin next trip was the nearest station of previous trip” and he examined first trip of next day to estimate correct destination point of the last trip of the day unlike Barry (2002). He pointed out with this assumption destination stations were identified at %66 level accuracy. Attanucci and Wilson (2011) built up an algorithm that detected destination stations from trip records in London Transportation Systems where 6 million trips took place daily. Munizaga and Palma (2012) used Smart card data along with Global Positioning System (GPS)and Geographical Information Systems (GIS). They estimated users’ alight points from trip chains of smart cards. At the same time they produced a lot of valuable information like OD matrix, user’s behavior analysis, cumulative passenger loads at the road sections. They assumed that “Every trip record begins close to the end point of the previous trip and the last trip of the day is close to the station where the start of the first trip of the day”. Gordillo (2006) studied London rail systems and estimated OD matrices. Farzin (2008), also, had used the same methodology to estimate the OD matrices for buses and BRT systems in Sao Paolo.
2. Istanbul’s BRT System

Istanbul’s BRT system consisting of 44 stations and having 52 km length began to serve in 2009 in a complete manner.

Before it has been begun the inference of OD matrix, to understand position of the Istanbul’s BRT system in public transport, it is going to be introduced and realized system analysis as used trip data in this section. The BRT system comprises 2 dedicated lines, one is for Western direction and the other is for Eastern direction. The system carries 850,000 passengers in weekdays as of May 2015. BRT is crucial transport system that accounts for 14% of daily passengers of Istanbul’s public transit. The system is preferred by citizens of Istanbul highly for the reasons such that being located in the main artery of the city, having large population in close proximity and having high frequency service. Because of the high preference, the load factor in BRT vehicles became extremely high. To provide solutions such as express or limited local services, additional lanes, bus assignments; it is necessary to analyze passenger trip behaviors. For this reason OD matrix should be generated. In order to apply the method that is proposed by Barry (2002), smart card data is used. Daily usage of Istanbul’s BRT system is 822,699 on 10 March 2015. The hourly distribution of the trips are given in Figure 1.

![Figure 1: Istanbul BRT Stations](image)

Daily frequencies of the passengers in the system are given in Table 1. There are 207,996 passengers that use the system only once at that day. The data that belongs to passengers who use the system once are not usable in OD estimation. However most of the passengers use the system twice a day. The number of passengers that use the system more than three times in a day is very low.

<table>
<thead>
<tr>
<th>Trip Frequency</th>
<th>Trips</th>
<th>Trip Frequency</th>
<th>Trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 times</td>
<td>207,996</td>
<td>5 times</td>
<td>1,650</td>
</tr>
<tr>
<td>2 times</td>
<td>242,133</td>
<td>6 times</td>
<td>809</td>
</tr>
<tr>
<td>3 times</td>
<td>25,831</td>
<td>7 times</td>
<td>163</td>
</tr>
<tr>
<td>4 times</td>
<td>9,192</td>
<td>8 times</td>
<td>104</td>
</tr>
</tbody>
</table>

![Figure 2: Hourly trips](image)
3. Methodology

Destination point of trips of Istanbul’s BRT system is estimated in this study with the methodologies and assumptions mentioned in previous section. For validation and verification purpose two different source of data used. The first data source is refund machines existing in Istanbul’s BRT stations. The second source for verification of the estimates is the travel data of other transportation modes. The nearest BRT station to the trips of origin of other modes such as bus, metro and ferry lines is assumed as destination station. OD matrix accuracy is very valuable and will provide significant benefits. OD estimations helps to analyze the system correctly and helps detection of bottlenecks in the system, the definition of express or limited stop services, determination of corridor capacities and section demands.

Validation of the method is answering the questions such that “Do the trip records in a mass transit can lead to system planners to Origin and destination data?”, “Are the assumptions made valid?” or as aforementioned, “Does next trip start on the point where the previous trip ends?” Or “Do passengers return to the point where they started their trip initially?” As using only BRT smart card data, it is not possible to do any inference about the destination point of passengers that having 1 record in the system in a day. As described earlier, according to the approach proposed by Barry (2002), the data of the passengers using the system 2 times are used. The passengers using the system more than 2 times in one day are excluded from estimation due to being used for other passengers.

242,133 passengers that use the system twice in one day cover 484,266 of the total trips. This account more than half the all trips. The first station of passenger that punch the smart card is recognized as origin station and the second station of the same passenger that punch the smart card is recognized as destination station. A passenger that uses the system 2 times on the same day has records at two different stations, for example, one record of 42nd station named “Uzuncayır” and one record at 19th station named Sefaköy. Then is assumed that this user boards on 42nd station and travels to 19th station in the first trip. It is also assumed that this user boards on 19th station and travels to 42nd station in the second trip.

In other words hypothesis of “passengers go and return between the first and second trip” was used to estimate the destination point. When it is estimated destination station of passenger, it is known boarding and alighting passenger at each station. Thus the accumulated count of passengers at stations and between stations can be calculated. The OD pairs obtained from the trip record do not include all the trips as mentioned before. Therefore the number of trips estimated between any two stations needs to be adjusted by correction factors (or weights). For example on March 10th, 2015 between at 18:00:00 and 18:59:59 time range 80,270 trips are recorded. Only 46,519 of these records are included in OD estimation. Therefore a correction factor for the time range is needed. This correction factor can be calculated by dividing 80,270 by 46,519 which is 1.72. In this case, each value in OD the matrix for the selected time range should be multiplied by 1.72. Many times planners use single correction factor for a specific time range. However we strongly advice to use different correction factor for the stations.

When the system is analyzed at 18:00 o’clock (between at 18:00:00- at 18:59:59) which is the rush hour of the BRT system, boarding and alighting activities of stations are as shown in Figure – 3 and cumulative numbers of passengers passing the stations are shown in Figure-4.

![Figure 3: Boarding and Alighting of Stations at 18:00 hours](image-url-buffer)
So how valid is this method that infers to the cumulative passenger numbers? Do passengers return to their previous trip’s destination station? Do passengers return to the first origin station of the day? The OD estimation approach is validated with 2 different methods which are reported in the following section.

4. Validation Method-1: Refund Data

As the obtained OD matrix is be able to validate, Istanbul’s BRT system presents an invaluable application area. As in Istanbul’s BRT system distance based pricing is used, there are refund machine on the station. The refund machines provide flexible payment options with adjustable structure. In this way, number of stations from origin to destination are counted and the fee taken at origin are returned by refund machines at destination accordingly. This in return helps verifying the accuracy of OD matrices, as it is reserved to destination station of trips which known the origin station in refund data. However this data is not available for all passengers because this policy is not applicable to students and social card users. They are already benefiting the discounted rates or they are using the system free. There are only flexible payment systems for regular card owners. These regular card owners who travel more than 28 stations are not eligible for repayment. So their destination data are not also recorded in anywhere. For these reasons, the refund machines usage rate is not sufficient to represent the whole population in sampling. They do not infer the passenger OD matrix adequately (27% using rates of total passenger in a day). Although OD matrices cannot be inferred from refund machines, the accuracy of obtained OD matrix can be checked by utilizing repayment records at the refund machines. In this situation “trips which are from 2 transactions per day group” and “at the end of these trips having 2 refund transaction” data can be used for the validation. On 10 Mart 2015 match, 112,988 trips were seen in both OD estimation process and repayment records. 94,368 records of 112,988 refund records (84%) were exact match when destination stations and repayment station were compared. Namely 84% of passenger using refund machine get repay the first refund on the station where having second boarding record and the second refund on the station where having first boarding record.

The station that having refund records is different from the station that having trip records to 18,620 of the resting trips (16%). Namely passenger don’t repay the first refund station which having second boarding. But being of estimation of more than 1-2-3 station deviation from the true destination will not affect the general approach of OD matrixes. There are 7,757 of records (7%) only 1 station deviation from the true destination. There are 3,372 of records (3%) only 2 station deviation from the true station and 2,274 of records (2%) only 3 station deviation from the true station. Consequently estimating of OD matrix are validated with refund data and is corrected accuracy of its 95% rates 1-2-3 station from the true destination.

5. Validation Method-2: Next trip point location

While generating the OD estimates, only BRT records of the passengers’ data are used. However passengers often do not use a single mode of transportation while moving from one place to another. There are frequently intermodal transfers in Istanbul's public transportation system when realizing 6.1 million trips per weekday. In this case successive trip records of passengers, who use more than one mode of the transportation system, can contribute to the discussion of whether the correct destination is determined. Therefore in the validation process, we analyzed the BRT stations’ locations and all systems’ stations (BRT, Bus, metro, light rail etc.) locations that
are close to the Istanbul’s BRT system (the stations that have distance less than 500 meters). We analyzed other trips (BRT, Bus, metro, light rail etc.) of users who use BRT at the same time.

683,628 trips of 822,699 BRT system trips on 10 March 2015 were useful in estimating the destinations. (It provides very reliable estimation of OD matrix). 421,546 trips that come from the passengers who use the system twice in one day are examined to validate the accuracy. When the results of the two approaches are compared to test the hypothesis of passengers travel between their first trips’ recorded stations and their second trips’ recorded stations, it has been observed that 362,839 trips overlap completely. So, 86% of these passengers’ destinations estimations are found to be true destinations. This rate is going to increase if the estimation of the passengers’ destinations includes 1-2-3 stations deviations as above. There are 21,067 of records (5%) only 1 station deviation from the true destination. There are 11,798 of records (2.8%) that are only 2 station deviated from the true station. And 4,279 of records (1%) are only 3 station deviated from the true station. In summary, OD estimation matrix is validated with next trip’s starting point and it has been observed that the accuracy of the method can reach 95% rates when 1-2-3 station deviations from the true destination are acceptable.

6. Conclusion

The accuracy of OD estimation methods are usually checked by conventional household and field surveys. Sometimes these surveys are used also to estimate OD matrices. But to reach sufficient number of survey respondent is not possible in Istanbul where BRT system has 820,000 average trips on a regular weekday. Although with this method it is possible to obtain OD matrix, it causes loss of time and cost.

Two novel approaches are used to check the validity of OD estimation method that presented Barry (2002). When 1-2 or 3 stations deviation from true destination assumed to be acceptable, the accuracy of the method is found at level of 95%. When boarding and alighting stations of trips are known, we can calculate cumulative passenger numbers at the stations, section loads, average trip times and average number of stations travelled by the passengers. With this knowledge, practical solutions can be put forward at the services operations, resolving congestions in short time and vehicle assignments. This method can be used in identifying the bottlenecks of the system.

There are the limitations of this study. It is not known the destination points of passengers using the system once in a day. There might be some solutions available which are needed to be researched. The vehicle passenger count systems and sensors can be placed on the vehicles to obtain OD matrices one hundred percent. This system can provide section load profile and cumulative passenger counts unlike passenger based destination estimation methods. Counting devices at the stations for leaving passenger can be placed as well.
References


Bagchi, M. and White, P.R., 2005, the potential of public transport smartcard data, Transport Policy, 12, 464-474


Farzin, J. M. (2008), constructing an automated bus origin destination matrix using fare card and global Positioning system data in Sao Paolo, Brazil, Transportation Research Record: Journal of the Transportation Research Board 2072, 30-37

Gordillo, F. (2006), The value of automated fare collection data for transit planning: An example of rail transit OD matrix estimation (Doctoral dissertation, Massachusetts Institute of Technology)

Li, B. (2009), Markov models for Bayesian analysis about transit route origin-destination matrices, Transportation Research Part B43, 301-310

Munizaga, M. A., Palma, C. (2012), Estimation of disaggregate multimodal public transport origin-destination matrix from passive smartcard data from Santiago Chile, Transportation Research Part C


Biography

Fahrettin Eldemir is an Assistant Professor in the Department of Industrial Engineering at Yıldız Technical University, Istanbul, Turkey. He has a Ph.D. in Decision Sciences and Engineering Systems, an M.E. in Operations Research and Statistics, B.S in Industrial and Management Engineering from Rensselaer Polytechnic Institute.

Murat Akad is an Assistant Professor in the Department of Industrial Engineering at Istanbul University. He has a B.S and Ph. D. in the Department of Industrial Engineering at Istanbul Technical University. His research interests are focused on transportations, operation planning and management.

Aydın Deli works as a researcher at TÜBİTAK (The Scientific and Technological Research Council of Turkey). He graduated a BS from the department of Industrial Engineering at Yıldız Technical University in 2009, and a MS from the department of Industrial Engineering at Istanbul University in 2015. Also he continue on a MS Computer Engineering at Gebze Technical University.