Using "Big Data" For Analysis and Improvement of Public Transportation Systems in Istanbul

Ilgin Gokasar, Kevser Simsek
Bogazici University, Department of Civil Engineering.
Istanbul/Turkey
ilgin.gokasar@boun.edu.tr, kevser.simsek@belbim.com.tr

Abstract

Istanbul´s automated fare collection system generates operational big data, including ridership on the BRT-Bus Rapid Transit line. This paper discusses the automated fare collection system and pricing policies for public transportation. Operational big data concerning ridership and research avenues into transportation planning management are described. Some recommendations for improving the planning and management of the BRT-Bus Rapid Transit line are provided. Graphical and tabular presentations of a processed selection of big data make them comprehensible and meaningful.

Keywords: Smart card (Istanbulkart); automated fare collection; transportation planning; transit data; BRT-Bus Rapid Transit transportation demand; ridership statistics; micro-simulation.

1. Introduction

With the improvement of smart card technology, automated fare collection systems have become the most common collection method used by public transit authorities. Since their invention in the 1969, new uses have been added to the original purpose of smart cards. They are portable and durable, features that make them useful for many purposes, such as authorization, payment, and identification. One important use of the smart card is the collection and processing of data.

Smart cards have become the fast, contactless, wireless technology of choice. Contactless transport systems are successful because they respond to the requirements of operators and end-users alike. Because of these advantages, the smart card, appropriately named the Istanbulkart, was put into service in Istanbul's public transportation system in 2004. Use of the Istanbulkart has been increasing ever since (Figure 1).

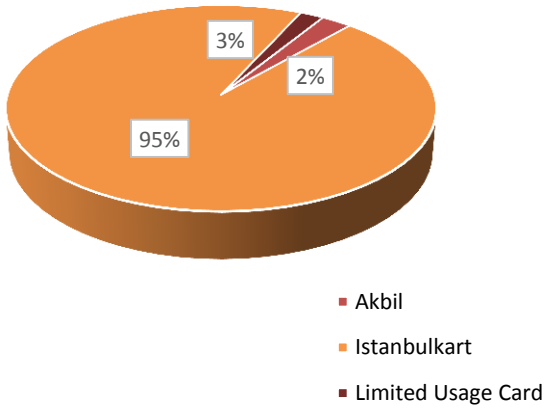

Figure 1. Istanbulkart usage percentages for BRT – line

The operators of Istanbulkart technology reduced the system's fraud and maintenance costs by replacing the inefficient paper-based fare collection system. Istanbulkart saves time by getting rid of queues for paper tickets. With use, the data collection and reporting capacity of the Istanbulkart has been improved. Istanbulkart is similar in look and size to a credit card. It is used to pay for transportation in place of more traditional methods, such as tickets and cards with magnetic stripes. Each Istanbulkart has a unique serial number. A card can be assigned to a specific individual or it can be anonymous. Also, Istanbulkarts can be specially designed for population groups such as elderly people, students, and government employees. A range of payment and fare options can be created electronically. Currently, for example, 37 different media types and 69 different fare types are provided by the Istanbulkart (Table 1). Bus fares are collected by an automated reader (Validator) next to the driver or at a turnstile before boarding a transport vehicle. The Validator is a smart device that does validity checks, collects fares in accordance with specified tariffs, and records the result of all transactions. Data collected from the stations are transferred from the stations' data transfer computers to the automated fare collection server located in the data center by means of an established external network.

Table 1. Istanbul Public Transportation System Ticket Types

Ticket Types				Definition
CR Media		SJ		Single Journey Ticket
		SV	AT	Anonymous Ticket
			RFT	Reduced Fare Ticket
			FET	Free Entry Ticket
			ST	Seasonal Ticket
		FC		Function Card
From Istanbul AFC		Akbil		Obsolete anonymous ticket
	Limited Usage Tickets	birGec		1 Journey Ticket
		ikiGec		2 Journeys Ticket
		beşGec		5 Journeys Ticket
		onGec		10 Journeys Ticket
	Stored Value Tickets Istanbul kart	Anonymous		Anonymous Ticket
		Discounted		Reduced Fare Tickets
		Free		Free Entry Tickets
		Function		Function Cards

There are refund machines at the exits from BRT- Bus stations. These machines recognize the cards of travelers who have used only a portion of the line and credits them with refunds up to 46% of the

ISBN: 978-1-62561-000-3

full fare. Hence, the trip time of each traveler can be calculated from information taken from the refund machines (Figure 2).

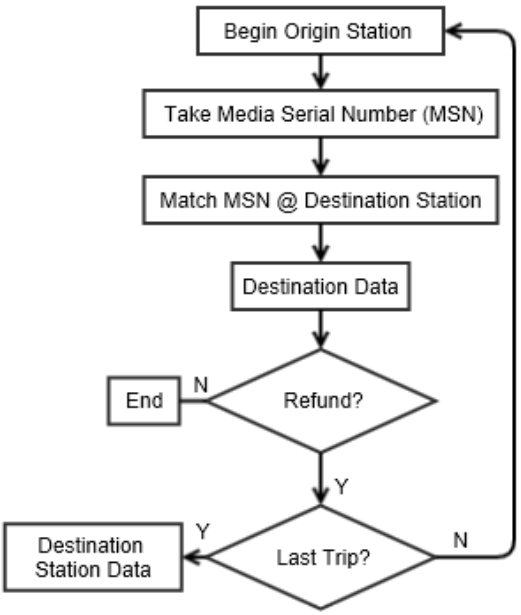

Figure 2. BRT-Destination station data process flow chart.

Typically, the date and time of the transaction, status of payment (transfer, acceptance, or refusal), card ID, fare type (student, adult, senior), route ID, and other related data are stored in the Validator and the central server [3]. These data are valuable because they can be used for transit planning, long term planning, and daily management of the transportation system. Masses of data are collected and stored, complicating the process of analyzing the data.

In brief, by using a smart card system, transportation authorities have access to:

1. Personal travel data of millions of people,
2. Information about each card and/or traveler,
3. Continuous trip data including refund information,
4. Identity of user and frequency of use.

Despite the accumulation of so much information, it is difficult to improve the Istanbulkart's usability and accessibility for the following reasons.

1. Its prevalent purpose is not to monitor performance of the transportation system; hence, additional passenger trip information such as destinations and delays cannot be directly retrieved.
2. Each passive data collection method has its disadvantages, and processing usually requires additional knowledge. Interoperating and mining heterogeneous datasets would enhance both the depth and reliability of transportation studies.
3. The amount of data obtained is increasing tremendously (approximately 6 million distinct Istanbulkart data every day) and traditional data processing methods might not be equal to the task.

Such data barriers make the development of a large-scale transportation performance monitoring system cumber-some and slow. [4]

In this study, we use SQL and SPSS software to extract passenger origin information from the Istanbulkart data. To avoid the problems of analyzing big data, we enable visualization and analysis of transportation performance measures by using SPSS and Microsoft Excel to highlight connections among heterogeneous transportation data sets, including Istanbulkart data. These programs provide a data-rich visualization platform from which to monitor transit network performance for purposes of planning and operations. It will also prove to be useful for data-driven transportation research.

2. Public Transportation in Istanbul

It is claimed by the Metropolitan Municipality of Istanbul that more than 22% of commuters drive their own cars to work (Table 2). This figure implies that traffic congestion can be decreased drastically if public transit takes a larger share of the commuting traffic. However, a commuter's choice depends on the convenience of the alternatives. To attract more riders, transit service must be improved to gain customer satisfaction and reduce operation costs. However, some of the barriers to improvement include:

1. Multi-headed management of city transportation;
2. Eighteen different bus companies;
3. More than 57,000 private operators;
4. Operator redundancy;
5. Three administration operators (IETT, IDO, ULASIM A.S.) and one government operator;
6. Prevalence of roadway transportation (88%);
7. Prevalence of private initiatives (71%);
8. Insufficient coordination and integration of administrative and legal framework;
9. Unfair competition and resource waste due to inadequate planning;
10. Difficulty of processing big data.

Table 2. Distribution of Daily Passengers According to Transportation Types

Type	Fleet	%	Trip/day	%
IETT BRT Line	334	0.02	715	5.61
IETT Bus	2,501	0.15	1,500,000	11.8
Private Buses	2,057	0.12	1,225,000	9.6
Car	1,602,730	96.02	2,800,000	22
Jitney	590	0.04	70	0.55
Minibus	5,860	0.35	2,000,000	15.7
Taxi	17,416	1.04	535	4.19
Employee Shuttle	36,902	2.21	2,419,000	19
TCDD	58	0	141	1.11
Light Metro	126	0.01	390	3.06
Metro	46	0.003	225	1.76
Street Tram	66	0.004	295	2.31
IETT Tram	4	0	2,824	0.02
Moda Tram	4	0	18	0.01
IETT	2	0.0001	108	0.08
Kabatas Funicular	2	0.0001	1,321	0.1
Ropeway	4	0.0002	700	0.01
IDO	98	0.01	325	2.55
Ferries	393	0.02	85	0.67
Total	**1,669,193**	**100**	**12,754,334**	**100**

ISBN: 978-1-62561-000-3

There are 2,503 buses working weekdays, traveling 521,533 km. on 14,000 trips, carrying 2,200,000 passengers. Approximately 700,000 of those passengers are carried by BRT- Bus rapid Transit.

2.1 BRT –Bus Rapid Transit

The Bus Rapid Transit system (BRT) of Istanbul began service in 2007 with the intent of reducing traffic congestion on arterial roads while providing quick and comfortable transportation. The line connects the Asian and European sides of the city. Thanks to this connection, the duration of the 53 kilometer journey from Avcılar on the European side to Sogutlucesme on the Asian side has been reduced to 83 minutes (Figure 3)
The advantageous features of BRT-Bus Rapid Transit are

1. fast transportation (30 second trip interval);
2. appropriate for metropolitan area with high population;
3. environmentally friendly;
4. comfortable alternative transportation.

There are several similar applications in other metropolitan cities of the world.

Figure 3. BRT-Bus Rapid Transit Line Map

3. Presentation and Analysis of BRT Data

The yearly and monthly increase of ridership on the BRT-Bus Rapid Transit is illustrated in Figures 4 and 5. The increase is due mainly to improvement of the transfer points and the addition of new lines and stations.

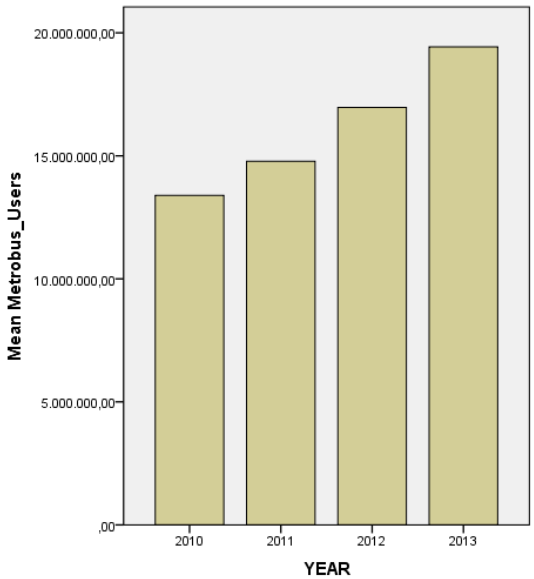

Figure 4. Number of passengers by year.

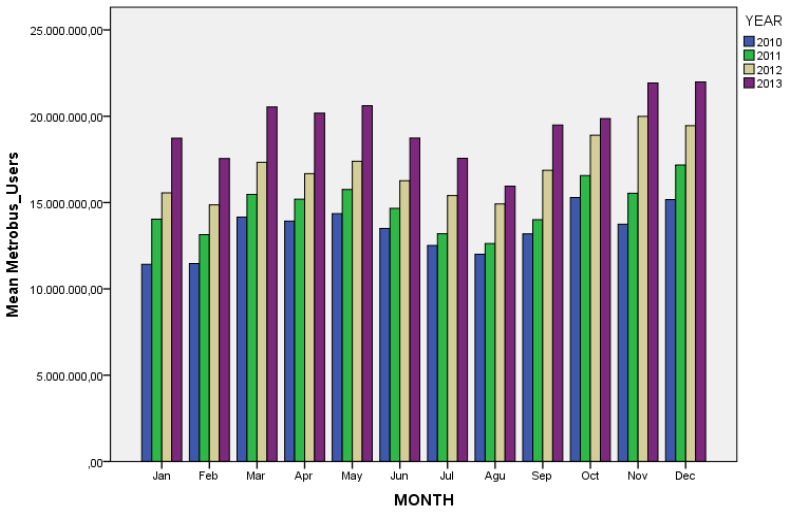

Figure 5. Number of Passengers per month

A unit of measurement is a definite magnitude of a physical quantity, defined and adopted by convention or by law, that is used as a standard for measurement of the same physical quantity.

Figure 6 shows Fare types /Stations per month (June 2013), and regular fare tickets have the highest share among the total fare types (Figure 7)

ISBN: 978-1-62561-000-3

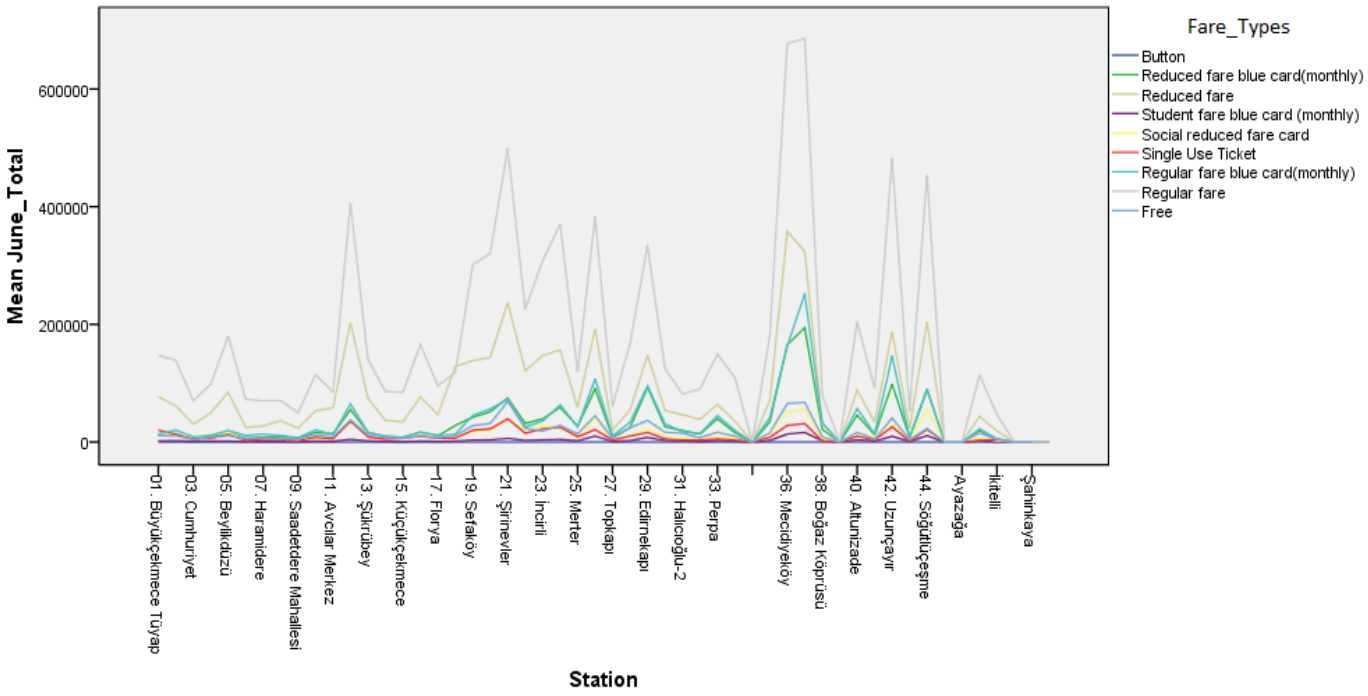

Figure 6. Fare types /Stations per month (June 2013)

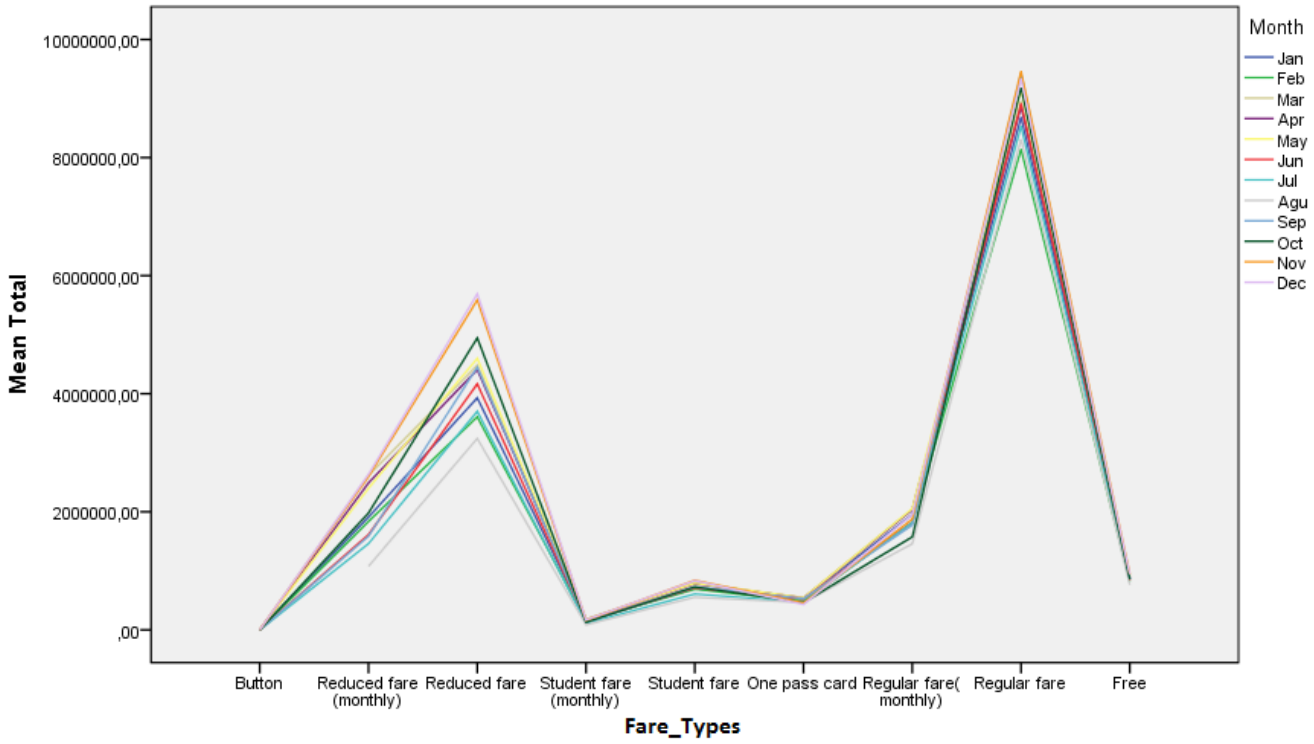

Figure 7. Fare types /Stations per year (2013)

ISBN: 978-1-62561-000-3

Besides being the central business district of Istanbul, Mecidiyekoy (station 36 in Figure 6) is a shopping center with big malls and broad streets. Universities and well-known high schools are located nearby. As a result, the area is crowded all day long. Mecidiyekoy has always been one of Istanbul's transport hubs. Shared taxis (minibuses) bring people from the suburbs, and many of the bus routes converge on Mecidiyekoy, making connections with the busy ferry ports. Moreover, of the sixteen stations along the original Metro line, Mecidiyekoy is the biggest. Not surprisingly, the number of passengers reaches a peak at this location.

There used to be a big bus station for buses that cross the bridges connecting the European and Asian sides of Istanbul. The BRT-Bus Rapid Transit was started by the city's transportation authority (IETT) to reduce traffic intensity on arterial roads and across the bridges. With the advent of the BRT, the number of bus rides between the European and Asian sides has been significantly reduced.

According to March, 2013, data, over 80,000 people are using Mecidiyekoy BRT station daily. Of these, 42,000 people are using buses after getting off the BRT. After getting off a bus, 70,000 people use the BRT.

In the light of overwhelming ridership, the station's use rose beyond its capacity, and the transportation authority of the Istanbul Metropolitan Municipality decided to build a new BRT interchange station, which also involved a pedestrian crossing to and from the underground Metro system. Now in operation, the new interchange station provides fast, safe, integrated access to all the transport modes.

The BRT-Bus Rapid Transit pricing system depends on the trip distance. When a traveler first uses his or her card, the full price is charged; then, at the destination station, he or she uses the card again, either at a refund machine or on another transit vehicle. First and last usage dates (by mining more than 6,000,000 different passenger data per day) were recorded to obtain mean travel time and origin-destination station data.

The graphics are obtained for approximately 36% of all travelers because only 255,188 of 715.000 total travelers receive refunds to their Istanbulkarts. Figure 10 shows number of passengers from origin station to destination station. Figure 11 shows mean travel time from origin station to destination station.

Istanbulkart was the source of data shown in Figures 8, 9, 10, and 11.)

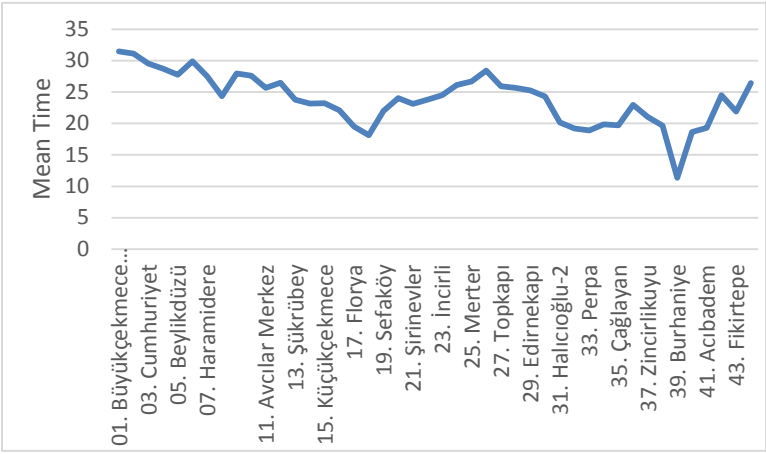

Figure 8. Origin station-Mean Time (min)

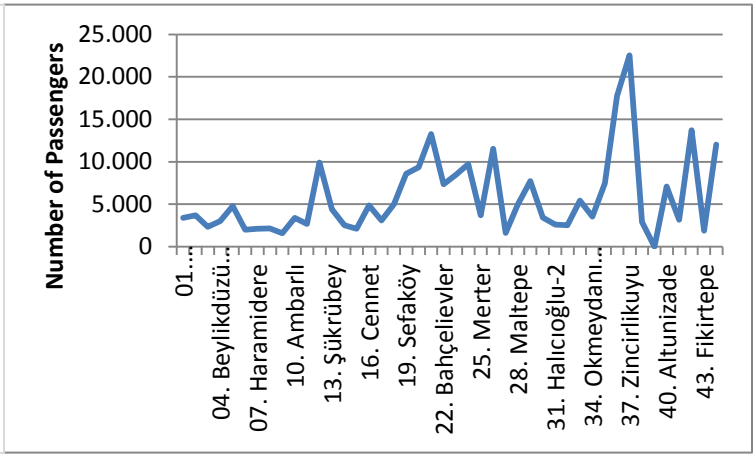

Figure 9.Station-Number of Passengers

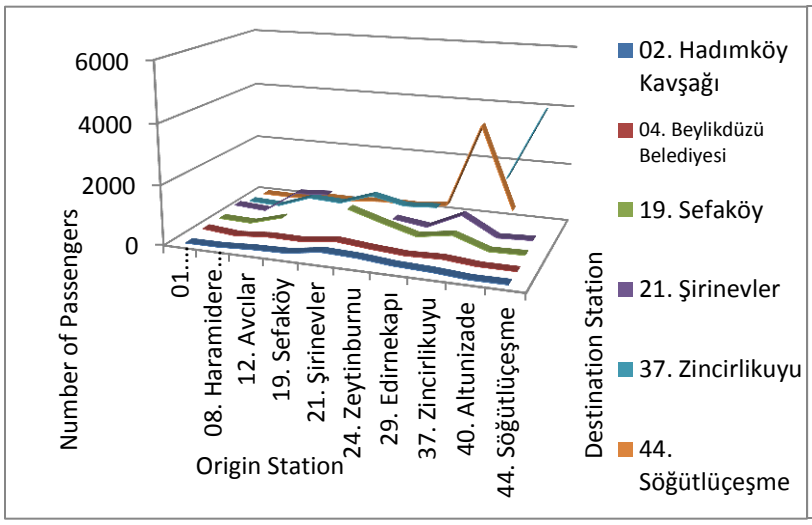

Figure 10. Origin station to destination station Number of Passengers.

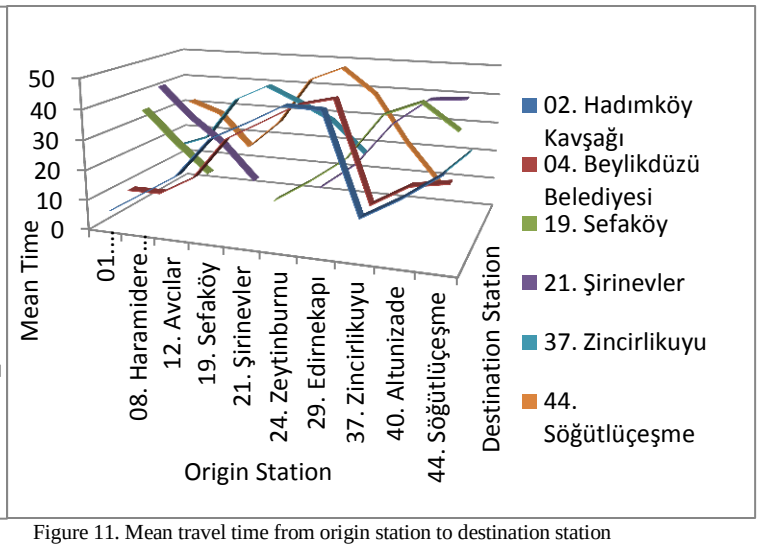

Figure 11. Mean travel time from origin station to destination station

ISBN: 978-1-62561-000-3

4. Conclusions

Improving the data-mining process while reducing the time needed for data processing has become an urgent problem. In this paper, Istanbul´s automated fare collection system and pricing policies, the operational source of big data, have been introduced. Even though examination of this data source alone is not enough to measure the full potential of Istanbulkart data for planning purposes, the processed data used in this study reveal the potential for problems that might be caused by inappropriate handling of the big data that Istanbulkart makes available. Therefore, topics for immediate research leading to the improvement of BRT-Bus Rapid Transit line planning and management should be

1. deploying models to mine large datasets for the purpose of discovering patterns of persistent problems;
2. developing super-fast visualization tools for this type of big data;
3. using the Istanbulkart data to develop OD tables by developing algorithms.

In the second stage of this study, an integrated data fusion procedure that models the travel patterns and regularities of transit riders along the BRT-Bus Rapid Transit line will be developed. This procedure will incorporate transit riders' trip chains based on their temporal and spatial characteristics and effectively capture their historical travel patterns. We will examine big data along four dimensions: “volume,” the scale of data; “variety,” the different forms of data; “velocity,” the rate of streaming data; and “veracity,” the relative certainty of data [6]. Then, through examination of travel patterns and transfer data, rider-level destinations can be estimated from multi-day observations and “agent-based micro-simulations” [7]. In these simulations, travelers and vehicles are modeled through agents that interact with the public transport system according to their individual goals. We use the MATSim simulation package which has an active user-base and has been applied to other large-scale scenarios [8]. Using MATSim, we hope to adapt demand management plans in such a way that utility is improved. We will use simulations to adapt pricing strategy and get a better match between available capacity and the demand emerging from the population [9]. We believe that an improved version of our SQL\SPSS visualization methods and use of agent-based micro-simulations can improve the design of demand management systems, including origin-destination based schemes, in the realm of public transport.

References

[1] IETT www.iett.gov.tr
[2] IBB www.ibb.gov.tr
[3] Blythe, P., 1998. Integrated Ticketing-Smart Cards in Transport. IEE Colloquium: Using ITS in Public Transport and Emergency Services, Paper No. 4, 20 p
[4] http://www-01.ibm.com/software/analytics/spss/
[5] Catherine Morency, Martin Trepanier, Bruno Agarda Measuring transit use variability with smart-card data, December 2010
[6] Data Mining Solutions: Methods and Tools for Solving Real-World Problems (27 July 1998) by Christopher Westphal, Teresa Blaxton
[7] Recognizing Demand Patterns from Smart Card Data for Agent-Based Micro-simulation of Public Transport, Paul Bouman, Milan Lovric, Ting Li, Evelien van der Hurk, Leo Kroon, Peter Vervest, ATT 2012
[8] Multi-agent transport simulation toolkit, 2012. http://www.matsim.org.
[9] E. Miller. Microsimulation and activity-based forecasting. In Activity-Based Travel Forecasting Conference, 1997.