Origin-Destination Matrix Generation Using Smart Card Data: Case Study for Izmir

Efendi Nasiboglu[1], Umit Kuvvetli[2], Mefharet Ozkilcik[2], Ugur Eliiyi[3]
[1]Dokuz Eylul University, Izmir, Turkey
[2]ESHOT General Directorate, Izmir Metropolitan Municipality, Izmir, Turkey
[3]Izmir University of Economics, Izmir, Turkey
[1]*efendi.nasibov@deu.edu.tr*, [2]*{ukuvvetli, mozkilcik}@eshot.gov.tr*, [3]*ugur.eliiyi@ieu.edu.tr*

***Abstract*— In parallel with growing populations and the development of the cities, the traffic densities and respectively the need for public transportation increase. While the local authorities transport the passengers to most of the locations they want, they must also maintain a high-quality public transport system to keep passengers' satisfaction levels high. The most reliable way to conduct such a good transport planning study is to benefit from actual numerical data. The number of the studies related to public transport planning by using smart card data has increased along with the developing technology and widespread usage of electronic fare collection systems. In Izmir city, which has employed an electronic fare collection system since 1999, a fully integrated public transport system has been developed that includes bus, light rail system, metro and ferry modes. In this study, the stages of constructing an origin-destination (OD) matrix that shows the regions where passengers board and alight are evaluated, and an example OD matrix generated by using actual data is interpreted.**

***Keywords*— transportation system; smart card data; OD matrix; electronic fare collection system**

I. INTRODUCTION

One of the most important services that are presented to citizens by local authorities is public transport. The significance of public transport gradually increases due to the environment, traffic and various transport problems in the developed cities that are overpopulated and the population continuously increases. Local authorities need to carry out several works to orientate the citizens to public transport, to make the public transport easier, faster and more efficient for these and suchlike reasons. Public transport is an area that the local authorities have a lot of problems and the performances of the local authorities on public transport affect their future. Because of these reason and reasons like this, successful planning of the public transport system, especially in the developed cities, is of capital importance.

Planning public transport by conducting studies based on data rather than subjective opinions will affect the reliability of the study and will provide obtaining more realistic results. Nowadays, developing technology opportunities and recording and storing data easily make the data more accessible. Conversely, the factor that provides the success is to use these data accurately and appropriately. In recent years, almost all sectors has benefitted from the magnetic cards that are plastic cards including microchip to give more personalized and qualified service to their customers. Firms can obtain personal information as how often, when and how much their customers use their product/service through magnetic cards. With these information firms can organize more personal campaigns, in other words they increase their customers' loyalty and satisfaction by considering their behaviors and habits. Accordingly in recent years the usage of the cards that are named as smart cards and store data as boarding, fare, etc. has increased significantly.

II. SMART CARD DATA AND OD MATRIX IN TRANSPORT PLANNING

A. *Smart Card Data*

Smart cards are cards that are similar to credit cards from the point of look and size and can store information within. The passenger fare types (adult, student, etc.) and remaining information can be defined in smart cards. Although the smart cards are frequently used nowadays, smart card technology is not new. The first patent was published in 1968 by two German inventors Dethloff and Grotrupp [1]. However, a lot of studies were realized about smart cards in the seventies, the use of the smart card has become significant since 1990 with the exponential growth of the internet and the mobile communication technologies [2]. The smart cards have many usages as access control and storing information (biometrics, photos, fingerprints, etc.). Public transport firms has preferred to use smart cards due to facts as security of payment, modern system look and the quality of data, too [3].

Public transport firms have always needed to know their passengers' trip demands. They have determined the routes that met these demands by benefitting from the information that passengers travelled from where to where by using which routes. Before the smart card technologies became common, public transport firms had benefitted from the various surveys to be able to specify their passengers' demands. But, the number of the studies that use the smart card data rather than the surveys to determine the travel demands has increased due to the cost overrun of the surveys, the difficulties in the application stages of these surveys and more reliable and realistic results of the smart card studies.

In the scientific literature, the number of studies that are related to the usage of the smart card data by the public transport firms is increasing rapidly. Bagchi, Gleave and White firstly investigated the nature of the smart card, and then they analyzed the travels and bus to bus transfers by using the pilot smart card data of England Merseyside Bus Firm and Bradford First Bus Firm for their study in 2003 [4].

In 2007, Morency, Trepanier and Agard worked on measuring the variability in public transport network usage by using smart card data. They identified and estimated the spatial and temporal variability of transport usage for Canada transport network [5]. In their two other studies in the same years, they illustrated smart card usage for estimating the performance of transport network supply, different network performance measures as statistics based on passenger service and how the passengers' behaviors are obtained by using smart card data and data mining methods [6,7].

Lianfu, Shuzhi, Yonggang and Ziyin indicated that smart card data play a crucial role in the planning stage of the urban public transport systems and they showed that many of the information obtained from the complex travel surveys could be attained by analyzing smart card data [8]. Pelletier, Trepanier and Morency dealt with various aspects of smart card usage for transport planning in their studies in 2009 [3].

B. OD Matrix

Origin-Destination (OD) matrix is a matrix that shows the total numbers of trips of OD pairs in the transportation network. It provides traveling information of passengers where their journeys start and where they end. So OD matrix is very important in planning of public transportation system [9].

Gaudry focused on the four approaches to estimation of OD matrix, dealt the role of this matrix, classifies the matrix determination procedures and provided some simple examples [10]. An algorithm to estimate a Bus Passenger Trip OD Matrix based on Automated Fare Collection (AFC), Automatic Passenger Count (APC) and Automatic Vehicle Location (AVL) data is studied by Cui in 2006. He applied the algorithm to estimate the OD matrix for a selected corridor of the Chicago Transit Authority (CTA) bus network [11].

Wang studied about the usage of Automatic Data Collection Systems (ADCS) data for transportation planning and obtaining bus passenger Origin-Destination (OD) inferences and investigated the travel behaviors in London in 2010. In this research the trip-chaining method was used to obtain bus passengers' boarding and alighting places and the results were validated by comparing with Bus Passenger Origin and Destination survey data [12]. Bert developed a method to determine dynamic Origin-Destination (OD) matrices for urban networks in 2009 [13]. Wan proposed a combined trip generation, trip distribution and traffic assignment model that made combined use of the Conditional Demand Analysis (CDA) and Traffic Assignment (TA) technique basing on the concepts of the conventional 4-step transportation model and statistical approaches using traffic counts in 2004 [9].

By using smart card data, to achieve the information like what time and from which bus stop each passenger boards is possible. While preparing OD matrix by these data, there are especially two important points for special attention. These subjects are respectively districting the city and recording or estimating (in the cases that obtaining the places passengers alight is not possible) the places passengers alight.

The first stage of creating the OD matrix for a city is districting the city. However, the bus stops where passengers board are available in the smart card data, creating the OD matrix between the bus stops increase the dimensions of the matrix many times more, making difficult to interpreting the matrix and accordingly the benefits obtained from the matrix reduce. Therefore, combining the bus stops considering specific criteria is needed. Two different ways can be chosen for linking the bus stops to the regions. First one is to gather the bus stops that are close to each other under a region by using the geographic coordinate information. This study should be conducted in company with an expert that knows the region well; the passenger densities of the bus stops are not considered in this option. The other method is to district the bus stops by basing on bus stop densities and using techniques like clustering analysis. In this matching, the bus stops that are similar for passengers and very close to each other might be assigned to different regions and this situation can be a disadvantage.

When the regions for the OD matrix are determined, if the number of the regions is too few or too many, the benefit obtained from the matrix might reduce. If the number of the regions is more than enough, it will increase the dimension of the matrix and can be hardly interpreted. On the other hand if the number of regions is too few, the interpretation will be easier but the results will be too general and not helpful. While the number of regions is determined, the criteria as the topography of the city, the size of the regions and the number of the bus stops must be considered as well.

After the necessary regions and each region's bus stops for OD matrix are determined, boarding and alighting regions of the passengers must be determined. For OD matrices, especially that are created by smart card data, the only option to record the alighting regions of the passengers is to make passengers to swipe their cards to validators while they are alighting from the bus. If the passengers swipe their cards while they are both boarding and alighting, the bus stops/regions the passengers travel between are known precisely. But this option is not frequently preferred because of the waste of time, dissatisfaction of the passengers and the cost of the system. Due to these reasons, generally the boarding regions of the passengers are obtained and knowing the alighting regions of the passengers is not possible. The alighting regions of the passengers are estimated under specific assumptions.

While the alighting regions of the passengers are inferred, the alighting bus stops of the passengers are ordered according to boarding sequence and the passenger travel between these bus stops is accepted. As for inferring the last alighting bus stop of the passenger, the first boarding bus stop is generally used. This inference method that construct boarding chains is used frequently in the literature. Another approach to infer the last alighting bus stop can be to assume the next day boarding bus stop as today last alighting bus stop especially if long time data is available.

III. APPLICATION

Izmir is the third largest city of Turkey. Situated on Aegean Sea coast, Izmir has the most developed public transport system of Turkey that integrates bus, metro, light rail system and ferry transport vehicles. The city accommodates approximately 4,500,000 people, 500,000 of which achieve 1,500,000 passes. Only one type of smart card is in use for all of the public transport vehicles in Izmir and these vehicles are fully integrated. The city electronic fare collection system, which has received numerous awards, allows the passengers to use all types of the transport modes/vehicles unlimited times by a single fare charge within 90 minutes after the first pass.

Since 1999, various studies have been carried out to restructure Izmir public transport system in parallel with the development of city and the population by the data obtained from smart card system. To improve satisfaction levels of the citizens in the public transport, travel habits and demands of passengers have great importance is clear.

In the first stage of the study, 6,597 bus stops/stations were linked with the regions according to their geographical positions in such a way that each bus stop/station would belong to only one zone. In accordance with the results of linking process, Izmir was divided to 563 subregions, 396 regions and 39 top regions. These top regions (orange), metro (red) and light rail system (blue) stations and ferry ports (light blue) were shown in Figure I. The dimension of the travel origin destination matrix would be obtained was induced and interpreting the matrix was eased. Then, the data including card type (adult, student, disabled, etc.), boarding date, boarding time, boarding bus stop/station, line used, boarding type (paid, transferred), boarding subregion, region and top region variables for a weekday in Izmir were analyzed. As a result, the OD matrix that shows the number of people who travel between the regions was created. Analyzing approximately 1.5 million boardings, regions and passengers were matched according to specific criteria. Special transportation solutions are also proposed for passengers who travel persistently. In addition to these boardings, the regions that have the maximum demand are examined for every region, and special price, special route and special frequencies are proposed for each region to meet their demands and provide higher passengers satisfaction. As a final analysis regarding the boardings, natural transfer points (the points that passengers frequently board/transfer) were also identified and some suggestions have been made to develop physical and social opportunities in these points to provide the social aspects of passenger satisfaction in addition to a successful planning.

The OD matrix was obtained as a result of the analysis of all daily boardings. A sample OD matrix showing the number of boardings and alightings for 12 regions that have the highest number of passengers are presented in Table I. The number of passengers (boardings) between 12 regions constitutes 78.1 percent of all boardings that occur during a week day.

When Table I is examined, it is observed that each region has its maximum demand toward itself except for Uchkuyular. Uchkuyular, has its highest demand as 27.1 percent toward Hatay region. In addition, passengers who prefer light rail system (Izban), metro and ferry (Izdeniz) prefer these vehicles for their next boardings again.

When the number of passengers is examined, the maximum demand is from Bornova to Bornova. Bornova has a very wide geographical area and its passengers' demand is in its own regions. This demand is 57 percent of all the passenger demand in this area, and has a rate more than the demand for all other regions in Bornova. The next popular region for Bornova departures is observed as the Metro system.

Another important point is that the percentage of transportation by ferry (Izdeniz) for Bayrakli, Alsancak and Karshiyaka, which are seaside regions, is not high as expected. The percentages for these regions (0.8 percent, 1.9 percent and 9.0 percent respectively) show that transportation by ferry must be developed in Izmir. Although the option of transportation by ferry is available between Alsancak and Karshiyaka, the demand from Alsancak to Karshiyaka by the buses is 10.7 percent whereas the demand of Alsancak to ferry (Izdeniz) is only 1.9 percent. These values show that most passengers who want to go from Alsancak to Karshiyaka region prefer buses over ferry.

ACKNOWLEDGMENT

We would like to thank Izmir Metropolitan Municipality ESHOT General Directorate for their invaluable support by sharing their database and know-how for public transport operations.

REFERENCES

[1] M. Shelfer and J.D. Procaccino, "Smart card evolution," Communications of the ACM, vol. 45(7), pp. 83-88, 2002.

[2] P. Blythe, "Improving public transport ticketing through smart cards," Proceedings of the Institute of Civil Engineers, Municipal Engineer, vol. 157, pp. 47-54, 2004.

[3] M.P. Pelletier, M. Trepanier and C. Morency, "Smart card data in public transit planning: a review," Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), November 2009.

[4] M. Bagchi, S.D. Gleave and P. White, "Use of public transport smart card data for understanding travel behaviour," European Transport Conference, 2003.

[5] C. Morency, M. Trepanier and B. Agard, "Measuring transit use variability with smart-card data," Transport Policy, vol. 14, pp. 193-203, 2007.

[6] C. Morency, M. Trepanier and B. Agard, "Measuring transit performance using smart card data," World Conference on Transport Research, San Francisco, USA, June 24-28, 2007.

[7] B. Agard, C. Morency and M. Trepanier, "Mining public transport user behaviour from smart card data," Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT), October 2007.

[8] Z. Lianfu, Z. Shuzhi, Z. Yonggang and Z. Ziyin, "Study on the method of constructing bus stops OD matrix based on IC card data,", Wireless Communications, Networking and Mobile Computing (WICOM), pp. 3147-3150, 2007.

[9] T.Y. Wan, "Application of conditional demand analysis in origin-destination (OD) matrix estimation using traffic counts and zonal characteristics," Department of Management Science Master of Philosophy, City University of Hong Kong, 2004.

[10] M. Gaudry, "The four approaches to origin-destination matrix estimation for consideration by the MYSTIC Research Consortium," Project Funded by the European Commission Under the Transport RTD Programme of the 4th Framework Programme, 2002.

[11] A. Cui, "Bus passenger origin-destination matrix estimation using automated data collection systems," Master of Science in Transportation, Massachusetts Institute of Technology, 2006.

[12] W. Wang, "Bus passenger origin-destination estimation and travel behavior using automated data collection systems in London, UK,"

[13] Master of Science in Transportation Massachusetts Institute of Technology, 2010.

[14] E. Bert, "Dynamic urban origin-destination matrix estimation methodology," Ecole Polytechnique Federale de Lausanne, 2009.

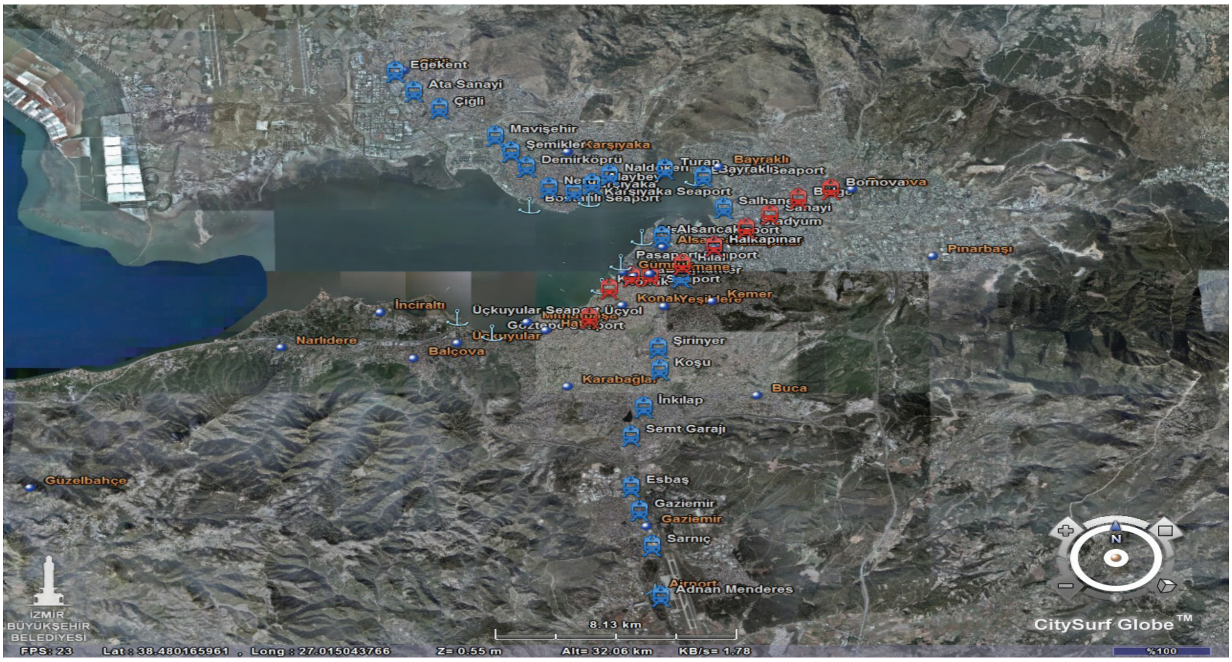

Figure 1. Top regions, metro and light rail system stations and ferry ports on Izmir map (Izmir Metropolitan Municipality- 3D City Guide-2012)

TABLE I. EXAMPLE OF OD MATRIX FOR SELECTED 12 REGIONS

			ALIGHTING REGION												
			Alsancak	Bayrakli	Bornova	Buca	Hatay	Izban	Izdeniz	Karabaglar	Karshiyaka	Konak	Metro	Uchkuyular	TOTAL
BOARDING REGION	**Alsancak**	f	5,379	3,816	5,024	2,725	2,524	3,607	767	2,922	4,436	4,616	3,812	1,737	**41,365**
		%	%13.0	%9.2	%12.1	%6.6	%6.1	%8.7	%1.9	%7.1	%10.7	%11.2	%9.2	%4.2	
	Bayrakli	f	4,133	11,893	4,341	1,236	28	2,880	279	63	5,171	454	2,985	36	**33,499**
		%	%12.3	%35.5	%13.0	%3.7	%0.1	%8.6	%0.8	%0.2	%15.4	%1.4	%8.9	%0.1	
	Bornova	f	4,922	5,857	60,112	3,258	135	2,858	309	614	1,802	1,146	24,066	112	**105,191**
		%	%4.7	%5.6	%57.1	%3.1	%0.1	%2.7	%0.3	%0.6	%1.7	%1.1	%22.9	%0.1	
	Buca	f	2,561	1,384	2,542	53,467	3,216	9,635	507	2,111	1,154	10,470	4,417	1,579	**93,043**
		%	%2.8	%1.5	%2.7	%57.5	%3.5	%10.4	%0.5	%2.3	%1.2	%11.3	%4.7	%1.7	
	Hatay	f	2,562	41	211	3,265	12,244	521	273	9,264	39	3,321	6,873	5,859	**44,473**
		%	%5.8	%0.1	%0.5	%7.3	%27.5	%1.2	%0.6	%20.8	%0.1	%7.5	%15.5	%13.2	
	Izban	f	5,045	3,363	4,163	9,323	879	50,358	1,857	674	3,966	1,235	11,085	72	**92,020**
		%	%5.5	%3.7	%4.5	%10.1	%1.0	%54.7	%2.0	%0.7	%4.3	%1.3	%12.0	%0.1	
	Izdeniz	f	757	79	112	123	94	1,896	12,872	119	5,427	2,740	1,830	973	**27,022**
		%	%2.8	%0.3	%0.4	%0.5	%0.3	%7.0	%47.6	%0.4	%20.1	%10.1	%6.8	%3.6	
	Karabaglar	f	2,822	45	319	2,650	5,116	736	555	18,956	34	12,077	9,392	1,333	**54,035**
		%	%5.2	%0.1	%0.6	%4.9	%9.5	%1.4	%1.0	%35.1	%0.1	%22.4	%17.4	%2.5	
	Karshiyaka	f	4,514	5,088	1,980	1,480	57	4,512	4,240	101	22,786	1,051	1,050	111	**46,970**
		%	%9.6	%10.8	%4.2	%3.2	%0.1	%9.6	%9.0	%0.2	%48.5	%2.2	%2.2	%0.2	
	Konak	f	4,436	497	1,542	10,337	4,386	1,354	1,277	14,756	1,329	21,022	6,453	3,968	**71,357**
		%	%6.2	%0.7	%2.2	%14.5	%6.1	%1.9	%1.8	%20.7	%1.9	%29.5	%9.0	%5.6	
	Metro	f	1,654	1,722	23,932	2,181	11,118	11,809	2,075	2,336	1,097	14,587	52,238	477	**125,226**
		%	%1.3	%1.4	%19.1	%1.7	%8.9	%9.4	%1.7	%1.9	%0.9	%11.6	%41.7	%0.4	
	Uchkuyular	f	1,301	84	113	1,720	5,437	291	666	1,435	174	3,105	2,491	3,213	**20,030**
		%	%6.5	%0.4	%0.6	%8.6	%27.1	%1.5	%3.3	%7.2	%0.9	%15.5	%12.4	%16.0	
TOTAL			**40,086**	**33,869**	**104,391**	**91,765**	**45,234**	**90,457**	**25,677**	**53,351**	**47,415**	**75,824**	**126,692**	**19,470**	**754,231**