

PRELIMINARY INVESTIGATION OF BOARDING BEHAVIORS OF TRANSIT USERS: A CASE OF IZMIR'S INTEGRATED FARE SYSTEM

Hassan Shuaibu Abdulrahman¹, Mustafa Ozuysal²

¹The Graduate School of Natural and Applied Sciences, Dokuz Eylül University, Izmir, Turkey

²Department of Civil Engineering, Dokuz Eylül University, Izmir, Turkey

Received 7 October 2020; accepted 6 April 2021

Abstract: This paper contributes to an emerging literature on the application of smart card fare payment data in public transportation planning. While this data is not mainly aimed at planning purposes, it can provide very useful evidences on travel behavior variability over space and time. Consequently, the analysis of various groups' behaviors in using a multi modal transit system a case Izmir city is evaluated. It argues a case for the potential of smartcard data in revealing a daily pattern especially using boarding information. In this study, 1,967,955 boarding data of 7 November 2018 in the smart card system of İzmir Metropolitan Municipality were analyzed. Analysis based on fare categories, type of boarding, number of daily boarding, and daily cost inquired by the various groups were estimated. The distribution of most users across all public transportation modes fits well with generalized extreme value distribution. These groups (adults, teachers and students) seem to be commuters and mostly active during the peak periods while the non-commuters tend to have a uniform boarding across time. The application of boarding information data can reveal important patterns, which is relevant to planners and policy makers.

Keywords: boarding behaviors, smart card data, integrated public transportation system.

1. Introduction

Sustainable urban mobility often has to do with the efficient utilization of the different mass transportation systems in order to deal with ever glaring problems such as invasive problems associated with car ownership whose impacts are in the form of traffic congestion, air pollution, accidents, etc. Different social groups especially those that cannot own or drive a car are catered for by mass transportation systems, which in turn contribute to the overall wellbeing of its

people and environment. The main demand for urban travel stems from the need to partake in the different activity centers by the various social groups. To derive a meaningful benefit from the transit system there is need to understand the behavior of the transit users. Methods for of analyzing transit users' behavior is carried out by the use of data extracted from on-board transit passenger surveys or household travel diaries that often demands a huge data user. Conventional techniques are used in evaluation of existing systems and policies, enacting new policies,

¹ Corresponding author: hassan.abdulrahman@ogr.deu.edu.tr

planning and management of the system by comparing alternatives. A strong case was made by Dieleman *et al.* (2002), they used this type surveys in their analysis and one of their conclusions was that personal characteristics have strong impacts on travel behavior. It is expected that for most of the trips will largely be regular trips (to and from work or study) or periodic. Therefore, understanding the ridership behavior should be of interest to the regulators and transit service providers. This will reveal salient insights into ridership patterns and allow for future planning in order to make mass transit attractive and lucrative for both users and operators alike Boyle *et al.* (2000). However, they have an inherent disadvantage of being capital intensive, time consuming and very difficult to conduct. This necessitates an alternative data source that can be used to discern information regarding travel patterns of various social groups. Advancement in technology has provided a potential rich source of data, which is mainly used for other purposes such as fare collection but contains important information that can be used in the field of transportation engineering. These cards can find utility in surveys regarding origins and destinations matrix estimation, longitudinal nature of trips, route choices as well as mode choices especially when dealing active transit users. Understanding transit users' behaviors represents one of the main challenges in smart card data analysis. Many studies have shown this ability in a wide spectrum of transit evaluation, planning, and operations etc.

Bagchi and White (2005) highlighted the importance of having access to larger sets of individual data over long periods for the better understanding of transit users. They studied the passenger behaviors of

the users using smart card data even though the absence of alighting information limits their analysis.

Utsunomiya *et al.* (2006) engages in a pre-analysis of smart card data from Chicago Transit Authority, USA to show the potential of using smart card data. Their work laid emphasis on daily traveling patterns based on access distances to transit, frequency and consistency, and the different users of smart card by residential area.

Trepanier *et al.* (2007) proposed a data mining method which can aid in grouping transit users into important market segments. This helps in classifying the users of transit services into groups based on their behaviors of daily and regular trip patterns.

The work by Pelletier *et al.* (2011) highlighted three levels of planning (long, medium, short) which information from smart data can be used to improve. They suggested that, the areas where smart card data may be used to improve includes; planning horizons such as designing a new policy of the system, which is strategic, or service adjustment and network definition and operational statistics of the system.

Ma *et al.* (2013) used spatio temporal characteristics of individual passengers from smart card to links trips, DBscan algorithm to deduct patterns and then cluster and classify travel pattern regularities.

In a similar work by Mahrsi *et al.* (2014), came up with a methodology which was applied to transit systems (bus and subway) in a city in France. They cluster similar temporal behavior based on boarding time which can be used to understand the demand and

investigated the effect of distribution various group characteristics.

In behavior analysis Ali *et al.*, (2016), used MATsim and smart card data in analyzing the travel behavior of transit users. They used the demand generated from smart card data as an input along with waiting time in the optimization of transit vehicle schedules and postulated the generate the microsimulation travel demand models using smart cards

Zhao *et al.* (2017), tried to answer questions regarding the individual spatio-temporal travel patterns of metro passenger in order to improve the services of the metro system in Shenzhen, China. They proposed a data-mining methodology which investigates travel patterns at individual levels and subsequently used clustering techniques to identify hidden attributes which was used to classify passengers based on similarities of their travel patterns (Zhou & Murphy, 2019). In a more recent study, the author developed a methodology for the utilization of smart card data in terms of processing the data so as to identify and visualize low demand routes (ride sharing) which was applied in a case study, Brisbane, Australia.

Izmir a city in Turkey which operates one of such automatic fare collection system and it is one of the most advanced because the fare collection system is an integrated one and making it possible for the user to use a single card for all modes of mass transportation available. In a single day over a million transactions are recorded across all modes. This forms a huge potential for behavior analysis.

This study attempts to fill this void by identifying the behavior of the different

users across different modes as captured in smartcard data and seeks to answer the basic question of which transit riders can be classified as commuters or non-commuters through an empirical analysis. Therefore, useful information on classification of commuters and non-commuters travel behaviors with emphasis on boarding data across different modes of mass transportation in an integrated fare collection system in Izmir City are presented.

2. Smart Card Data Izmir Case Study

Izmir a city in Turkey has an urban population of about 4,061,074 (Turk Stat, 2013). It has one of the most developed and well integrated public transit systems in the country because it can boast of quite a number of different and well integrated modes. In Izmir, the urban mass transportation network is comprised of the public bus system bus (ESHOT AND IZULAS), three rail systems (METRO-rapid transit, the IZBAN-commuter rail Service, TRAMVAY- light rail service), and a ferry service (IZDENIZ) which operate in the inner bay. In the mass transportation network, trips are divided into five different categories: Student, free (disabled, policemen, press card-holder professionals, the unpaid), 3-5 journey boarding pass (for passengers without a smart card), discounted (teachers, municipal workers, the elderly over 65: price differs according to the type of pass), and full fare boarding (all passengers not mentioned above). The fare system operated has various classes for example (ESHOT/METRO/IZDENIZ/TRAMWAY) has a time based fare system. The first boarding is charged fully, and subsequent transfer boarding are charged additionally within transfer time of 120

minutes. While for IZBAN has a distance based fare. For example, a full fare card type, the first 25km the fare is charged fully and next 25km is charged additional token. In this study, 1,967,955 boarding data of (7

November 2018) in the smart card system of the İzmir Metropolitan Municipality were analyzed (ESHOT, 2020). This data is quite large and needs some data mining techniques to extract meaningful data.

Table 1
Types of Boarding Based on Mode of Transport

Mode of Transport	Type of Boarding						Total	Percentage (%)
	3 5 Card	60 Years	Student	Teachers	Free	Full Fare		
ESHOT (Bus)	522	25033	254972	7955	121621	401151	811254	41.2
IZBAN (Commuter Train)	989	12269	150050	6776	29888	288977	488949	24.8
IZDENIZ (Ferry)	634	1761	8676	494	6179	33581	51325	2.6
IZULAS (Bus)	64	5820	55632	1673	27867	96720	187776	9.5
METRO (Rapid Transit)	918	8324	109783	3583	32118	172411	327137	16.6
TRAMVAY (Light Rail)	136	4120	21635	1312	16062	58183	101448	5.1
TOTAL	3263	57327	600748	21793	233735	1051023	1967889	100

Note: Data description 7th November, 2018

The paper aims to identify changes of travel behavior in time and space and therefore employs the variables to reflect spatial and temporal characteristics of travel behavior. Frequency is the most representative characteristics of travel behavior and can be said to represent both spatial and temporal aspects of travel. Temporal characteristics

are represented by timing of a trip along the day. The characteristics here suggested is the longitudinal nature of time and considering all social groups boarding from morning till evening. The figure below shows the diagram representing the travel behaviors investigated. This was done using simple Microsoft Access and Excel to sort the data and analysis.

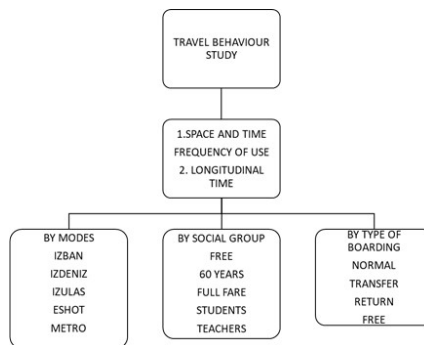


Fig. 2.
Methodology for Preliminary Investigation

2.1. Aggregate Pattern Recognition

Figure 3 below shows the fare type as it relates to the different modes of transportation. By far the most dominant group is the full fare and the student groups forming about 60% in IZBAN and about

80% for other modes of transportation. However, unlike the distribution of students in other modes, the portion of students who use IZBAN seems to be very low which may suggest that that most students live close to their schools and will be constrained by living in suburbs.

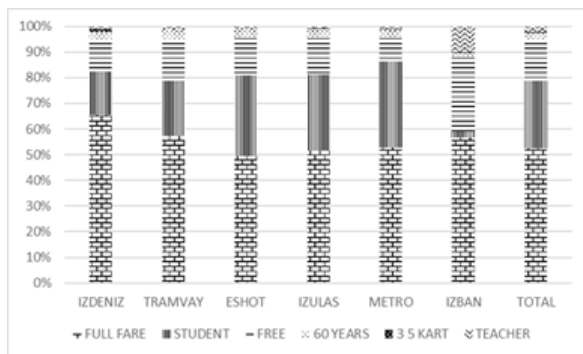


Fig. 3.
Classification by Users

A look at Figure 4 shows the boarding types which include three categories; Normal boarding, Transfer and the free groups. In all modes the normal boarding forms more than 50% of all boarding across all modes while transfers make about 30% and free boarding about 10%. This may suggest the number of direct trips being about 50% even

though this may be misleading since the user may use his card more than once in a day. However, this extraction may be useful when estimating the accrued revenue. Also, in a time based fare type this will appear to be detrimental as there is about 40% of free or discounted transaction when considering the revenue generated.

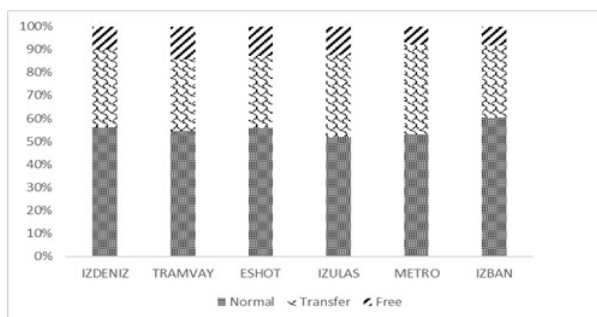


Fig. 4.
Classification by Boarding Type

Figure 5 illustrates that, for most of the cases, there are identifiable peaks between (7:00-8.00 am) and relatively flatter peak in the evening (16:00-18:00 pm). A uniform curve is noticeable for modes such as for TRAMWAY and IZDENIZ. This shows a general nature of all boarding which may be

used to identify peak periods across all modes of transportation. However, this does not show dominant social groups even though it can be inferred that most of the users engage in regular trips in the morning and evening for all modes with exception of the aforementioned modes which tends to be used for sightseeing.

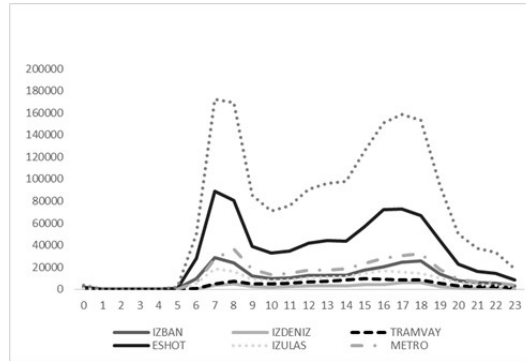


Fig. 5.
Temporal Analysis According to the Mode

The number of usage of cards can be measured by frequency which is depicted in Figure 6. The frequency was classified into 5 groups 1 to >5. The largest frequency is 2 across all

modes Again, this shows that most users are commuters and in some cases like IZDENIZ, where the frequency is about 60%, may imply trapped users who must engage on two-way trip.

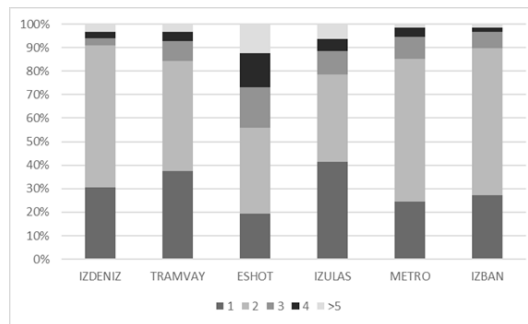


Fig. 6.
Number of Boarding per User per Mode

Investigating further since the transit system is well integrated and most users use more than one mode for their trips. The combined data shows that even number of uses (2, 4,)

are peak more than an odd number of uses, showing the dominance of round trips. The distribution of the frequency was evaluated by fitting different distribution to see which

best describes the distribution. From the Figure 7 below it can be seen that the generalized extreme value distribution fits best among the distribution chosen. This distribution is skewed to the left implying

that most people predominantly use their cards twice on this day. This may also signify that that most travelers engage in direct trips and transfer trips may be minimal or that the travelers hardly engage in other activities.

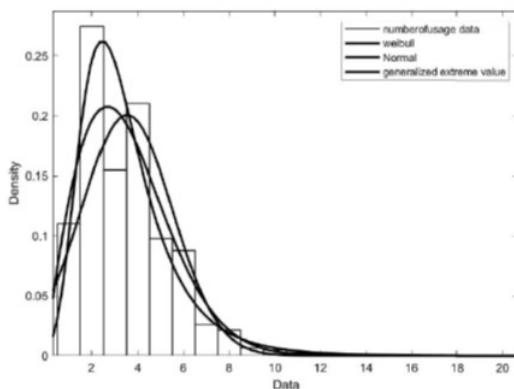


Fig. 7.
Fitting Distributions to Number of card Usage

The Table 2 below presents the statistics of the usage of various modes. The mean of individual modes usage is tending towards 2 even though the median value for TRAMWAY and IZULAS is 1 that is to

say that, they may not necessarily be used for commuting trips. However, when the data was combined the mean was increased to well above two showing increase in trip rates which signifies that multiple modes are used.

Table 2
Descriptive Statistics of Card Usage per Mode

	IZDENIZ	IZBAN	TRAMVAY	METRO	ESHOT	IZULAS	ALL
Mean	1.551001	1.607993	1.493815	1.6804	1.985828	1.461864	2.525187
Standard Error	0.003292	0.001506	0.002555	0.001538	0.001793	0.002046	0.002403
Median	2	2	1	2	2	1	2
Mode	1	2	1	2	1	1	2
Standard Deviation	0.589179	0.602381	0.654678	0.676162	1.124748	0.713811	1.534967
Sample Variance	0.347132	0.362863	0.428604	0.457194	1.265059	0.509527	2.356122
Kurtosis	1.547966	0.695279	1.858749	2.174577	3.730695	3.421434	1.721387
Skewness	0.873285	0.664346	1.311399	1.033819	1.601637	1.747897	1.28643

In the overall transit system, the revenue generated from boarding has ESHOT the highest earnings because it is by far the dominant mode, far more accessible, and only a few number of the buses operates

the distance based fare system. Also fare type used in IZBAN is based on distance while others are based time, therefore, some percentages of the boarding fees are returned to alighting passengers from IZBAN.

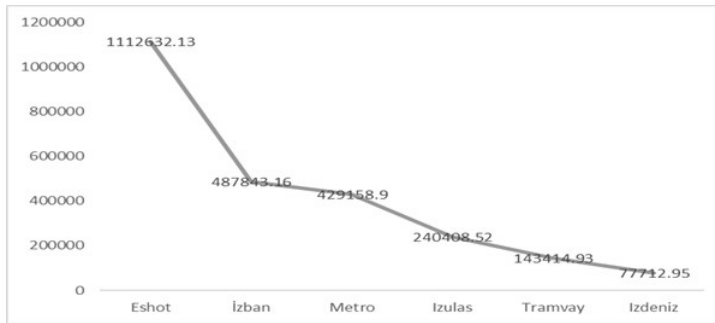


Fig. 8.
Cost of Boarding per Mode

2.2. Individual Group Behavioral Patterns

Investigating further the individual social groups across all modes reveals in all the groups can be further classified as commuters and non-commuters. The commuter group include full fare, student and teachers as

depicted in figure below all has sharp peaks in the morning and evening periods. This may imply that all these groups have a round trip across all modes of transportation. The non-commuter groups include; free, 60 years with the 35card users having a noisy distribution across all modes of transportation.

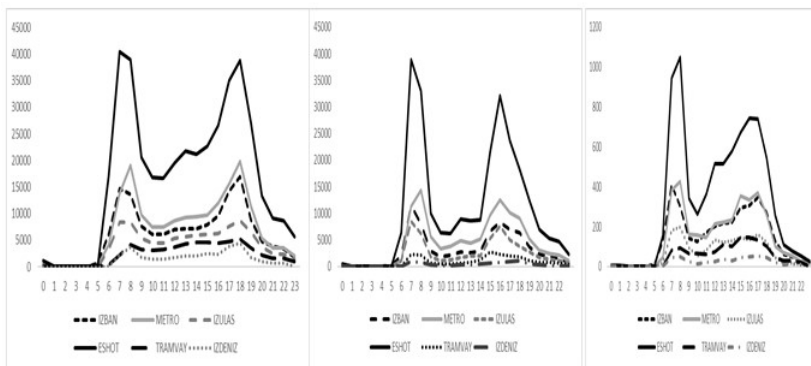


Fig. 9.
Commuter Temporal Patterns (Full Fare, Students and Teachers Respectively)

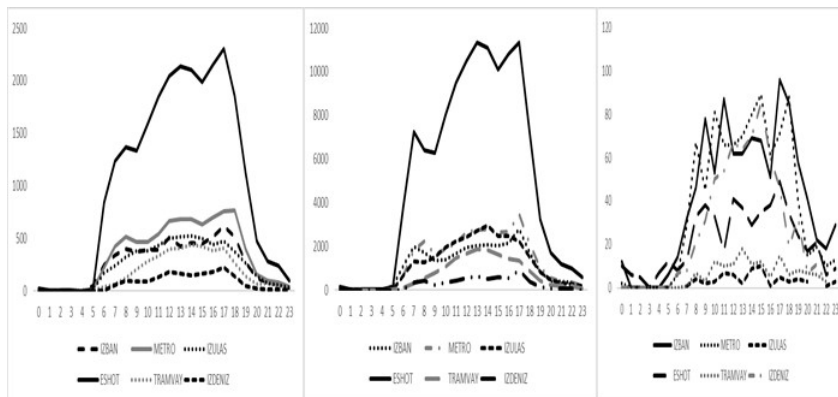


Fig 10.
Non-commuter Temporal Patterns (60 Years, Free and 35 Card Respectively)

The figure below shows probability mass function of all groups. All social groups hardly use their card once but a multiple of twice or thrice for this particular day. But by far most

of the users use their smart cards twice. This result should not be surprising since mass transportation system is predominantly used by low to middle income earners and students also.

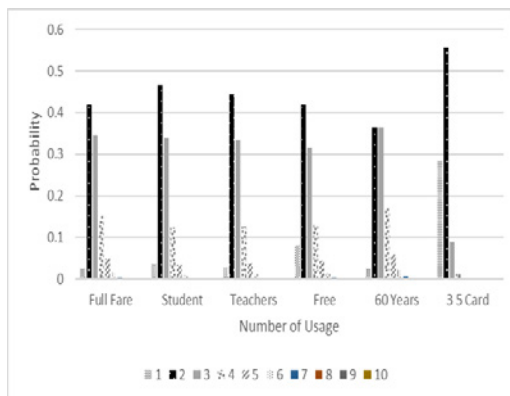


Fig. 11.
PMF of Number of Card Usage for Different Social Groups

3. Conclusion

This study is a preliminary assessment on boarding behavior of social groups by extracting data from an integrated Fare Collection System in Izmir, Turkey. Smart

cards can reveal various aspects regarding routine and changes in travel behaviors of social groups but cannot replace traditional surveys because there are limitations regarding alighting information. And they also offer substantial benefits for transit

agency to support transit market analysis and improve their transit service. In this study, various groups were further classified as commuters and non-commuters, peak periods were identified, modes with most users and transfer, boarding cost across all modes, and frequency of card users. From the results of the study the social groups were broadly categorized into commuters and non-commuters based on spatio-temporal behaviors of the smart card data. Thus, a better understanding of the variability of the transit demand could help the transit operators to move towards a more customized supply for types of users that would match the observed variability. Also, for further work on spatio-temporal variability behavior of different card users may be evaluated based on weekly, monthly or yearly basis which has a huge potential, it will require more data, and consequently more powerful data mining techniques.

Acknowledgements

The authors are grateful to ESHOT for the provision on their database.

References

- Ali, A.; Kim, J.; Lee, S. 2016. Travel behavior analysis using smart card data, *KSCE Journal of Civil Engineering* 20(4): 1532-1539.
- Bagchi, M.; White, P. R. 2005. The potential of public transport smart card data, *Transport Policy* 12(5): 464-474.
- Boyle, D.K.; Foote, P.J.; Karash, K.H. 2000. Public transportation marketing and fare policy, *Transportation in the New Millennium*. 6p.
- Dieleman, F. M.; Dijst, M.; Burghouwt, G. 2002. Urban form and travel behaviour: micro-level household attributes and residential context, *Urban Studies* 39(3): 507-527.
- ESHOT. 2020. Eshot Genel Müdürlüğü Resmi Web Sitesi. Available from Internet: <<https://www.eshot.gov.tr>>.
- Ma, X.; Wu, Y. J.; Wang, Y.; Chen, F.; Liu, J. 2013. Mining smart card data for transit riders' travel patterns, *Transportation Research Part C: Emerging Technologies* 36: 1-12.
- Pelletier, M. P.; Trepanier, M.; Morency, C. 2011. Smart card data use in public transit: A literature review, *Transportation Research Part C: Emerging Technologies* 19(4): 557-568.
- Trepanier, M.; Tranchant, N.; Chapleau, R. 2007. Individual trip destination estimation in a transit smart card automated fare collection system, *Journal of Intelligent Transportation Systems* 11(1): 1-14.
- Turk Stat. 2013. Population of provinces by years, 2007-2014. Turkish Statistical Institute. Available from Internet: <<http://www.tuik.gov.tr/UstMenu.do?metod=temelist>>.
- Utsunomiya, M.; Attanucci, J.; Wilson, N. 2006. Potential uses of transit smart card registration and transaction data to improve transit planning, *Transportation Research Record* 1971(1): 118-126.
- Zhao, J.; Qu, Q.; Zhang, F.; Xu, C.; Liu, S. 2017. Spatio-temporal analysis of passenger travel patterns in massive smart card data, *IEEE Transactions on Intelligent Transportation Systems* 18(11): 3135-3146.
- Zhou, J.; Murphy, E. 2019. Day-to-day variation in excess commuting: An exploratory study of Brisbane, Australia, *Journal of Transport Geography* 74: 223-232.