LINE-BASED APPROACH FOR ESTIMATING LANDING STOPS USING SMART CARD DATA.

Sh. Xidirov

Namangan state technical universityAssoc. Prof.,

E-mail: shuhrat1228@gmail.com

Annotation: Today, due to the increasing population and widespread use of private cars, urban transportation is becoming an increasingly big problem and the importance of public transportation for livable cities is increasing day by day. Especially in terms of minimizing the traffic problem in metropolitan cities as well as the damage to the environment, public transportation has become a problem that needs to be solved all over the world. In order for local governments and public transportation service providers to provide a successful transportation service, they need to know when and between which parts of the city the passengers are traveling and make plans based on this information. Due to developing technology, the use of electronic fare collection systems in public transportation has increased and thanks to the smart cards used, it has become easier to reach the objective numerical information required for successful planning. In almost all smart card systems, there is no data on the passengers' boardings, despite the detailed data kept on their disembarkations. Due to this deficiency, estimating the landing information is becoming quite important in terms of both the creation of OD matrices showing the travel movement in cities and other statistics, especially in terms of correctly determining the passenger density in the bus. Within the scope of the study, a method has been studied that estimates the landing stops based on the basic logic that passengers will get off at the closest line stop to their next boarding, using the stop and line information they get on, and according to certain simplifying assumptions for passengers who board only once during the day. In this study, the results of the method have been evaluated on a current data set belonging to Tashkent, where bus, metro, ferry and light rail systems have been used integrated since 2021.

Keywords: Smart Card data, Transportation Planning, Passenger statistics.

Аннотация: Сегодня, в связи с ростом населения и широким использованием личных автомобилей, городской транспорт становится все более серьезной проблемой, и важность общественного транспорта для городов, пригодных для жизни, возрастает день ото дня. Особенно с точки зрения минимизации транспортных проблем в крупных городах, а также ущерба окружающей среде, общественный транспорт стал проблемой, которую необходимо решать во всем мире. Для того чтобы местные органы власти и поставщики услуг общественного транспорта могли успешно предоставлять транспортные услуги, им необходимо знать, когда и между какими районами

города путешествуют пассажиры, и строить планы на основе этой информации. Благодаря развитию технологий использование электронных систем оплаты проезда в общественном транспорте расширилось, а благодаря используемым смарт-картам стало проще получать объективную цифровую информацию, необходимую для успешного планирования. Почти во всех системах смарт-карт отсутствуют данные о посадке пассажиров на борт, несмотря на то, что сохраняются подробные данные об их высадке. Из-за этого недостатка оценка информации о посадке становится весьма важной как с точки зрения создания матриц OD, показывающих движение пассажиров в городах, так и с точки зрения другой статистики, особенно с точки зрения правильного определения плотности пассажиров в автобусе. В рамках исследования был изучен метод, который оценивает количество посадочных остановок, основываясь на базовой логике, согласно которой пассажиры будут выходить на ближайшей остановке линии к своей следующей посадке, используя информацию об остановке и линии, которую они получают, и в соответствии с некоторыми упрощающими допущениями для пассажиров, которые садятся только один раз в течение тот самый день. В этом исследовании результаты метода были оценены на основе текущего набора данных, относящегося к Ташкенту, где с 2021 года интегрированы системы автобусного сообщения, метро, паромной переправы и легкорельсового транспорта.

Ключевые слова: Данные смарт-карт, Планирование перевозок, Статистика пассажиров.

Annotatsiya: bugungi kunda aholi sonining ko'payishi va xususiy avtomobillardan keng foydalanish tufayli shahar transporti tobora katta muammoga aylanib bormoqda va yashash uchun qulay shaharlar uchun jamoat transportining ahamiyati kundan-kunga ortib bormoqda. Ayniqsa, metropoliten shaharlaridagi transport muammosini va atrof-muhitga etkazilgan zararni minimallashtirish nuqtai nazaridan jamoat transporti butun dunyoda hal qilinishi kerak bo'lgan muammoga aylandi. Mahalliy hokimiyat organlari va jamoat transporti xizmatlarini etkazib beruvchilar muvaffaqiyatli transport xizmatini taqdim etishlari uchun ular yo'lovchilar shaharning qachon va qaysi qismlari o'rtasida sayohat qilishlarini bilishlari va ushbu ma'lumotlarga asoslanib rejalar tuzishlari kerak. Rivojlanayotgan texnologiyalar tufayli jamoat transportida elektron tariflarni yig'ish tizimlaridan foydalanish ko'paydi va ishlatilgan smart-kartalar tufayli muvaffaqiyatli rejalashtirish uchun zarur bo'lgan ob'ektiv raqamli ma'lumotlarga erishish osonlashdi. Deyarli barcha smart-karta tizimlarida yo'lovchilarning tushishi haqida batafsil ma'lumotlarga qaramay, ularning taxtalari haqida ma'lumot yo'q. Ushbu etishmovchilik tufayli qo'nish ma'lumotlarini baholash shaharlarda

ISSN 3030-3907

sayohat harakatini ko'rsatadigan OD matritsalarini yaratish va boshqa statistik ma'lumotlar, ayniqsa avtobusdagi yo'lovchilar zichligini to'g'ri aniqlash nuqtai nazaridan juda muhim ahamiyat kasb etmoqda. Tadqiqot doirasida qo'nish bekatlarini yo'lovchilar keyingi bortga chiqish uchun eng yaqin yo'l bekatida, ular to'xtaydigan to'xtash va chiziq ma'lumotlaridan foydalangan holda tushadigan asosiy mantiq asosida taxmin qilish usuli o'rganildi. faqat bir marta bortga chiqadigan yo'lovchilar uchun ba'zi soddalashtirilgan taxminlarga ko'ra. kun. Ushbu tadqiqotda usul natijalari 2021 yildan beri avtobus, metro, parom va yengil temir yo'l tizimlaridan foydalanilgan Toshkentga tegishli joriy ma'lumotlar to'plami bo'yicha baholandi.

Kalit so'zlar: Smart-karta ma'lumotlari, transportni rejalashtirish, yo'lovchilar statistikasi.

Introduction

Today, the spread of cities to larger areas and the rapid increase in population have made transportation planning one of the most important issues for local governments in order to carry out urban public transportation services more effectively and efficiently. In addition, traffic problems and environmental problems in metropolitan cities are increasing the importance of urban public transportation services. In addition to all these, detailed analyses of objective data obtained from smart cards in particular have a very important place in public transportation planning, and it is a fact that the use of these analysis results in transportation planning increases success.

Smart card technology can contain very important information such as photographs, fingerprints, DNA results, various biometric data, medical data, banking data, and is therefore

It has found its place in sectors. Parallel to this development in other sectors, public transport service providers have also started to use smart card technology instead of traditional tickets due to its advantages such as security, modern appearance and high data quality, especially in terms of collecting fares [8]. Smart cards are used in many countries of the world and continue to become widespread in the public transport sector [7]. Despite these advantages, the high research and development costs of smart card systems [5] and the difficulties experienced in analysis due to the highly complex structure of the data [3] constitute the negative aspects of smart cards.

Although smart card data includes all the detailed data on passengers' boarding, there is no data on passengers' disembarkation. On the other hand, landing information is required for analyses on when and between which regions passengers travel. Since landing data is not included in smart card systems, passenger landing information needs to be estimated with various approaches. The

ISSN 3030-3907

success of these predictions is very important in terms of creating the OD matrix showing the travel movement in the cities, and also in terms of determining the passenger density for the bus correctly and various statistics such as passenger-km, passenger-time, etc.

In this study, a method was worked on to estimate the stops where the passengers get off using the stop and line information they get on. Within the scope of the study, the accepted assumptions were examined, the effects of the assumptions on the results were investigated, and the results obtained using the smart card data between September 22-26, 2014 were shared.

LITERATURE

Smart cards are plastic cards that have the ability to store and process various information thanks to the microchips placed inside them. The most well-known and used smart cards are bank and credit cards, and today, the use of these cards for different purposes in many different sectors is becoming increasingly widespread. Although there has been an increase in the use of smart cards and the number of studies on this subject in recent years, smart card technology is not a new concept. The smart card patent was first obtained in 1968 by German inventors Dethloff and Groutrupp [9]. Although there were many studies on smart cards in the 2010, there was a significant increase in the use of smart cards in the 2018 in parallel with the rapid growth in internet and mobile communication technologies [4].

The success of smart cards in data storage has revealed some advantages over traditional methods.

Olmazor and White stated in their study that survey methods in data collection regarding processes have decreased to a minimum level as a result of the use of smart cards [2]. The same people, in their later studies, have suggested that the importance of statistical studies has increased as a result of the combination of personal data and travel data of smart cards [3]. In parallel with the increasing use of smart cards in public transportation, various studies have been conducted with the help of the obtained data, especially in transportation planning. Some of these studies are various analyses aimed at understanding passenger behavior such as boarding time, boarding stop, and travel duration [10]. In 2005, Olmazor and White analyzed the boarding and alighting places of the passengers and created an origin-destination matrix, aiming to improve the service provided with the help of this matrix [3]. Trepanier and Vassiviere used various statistical methods to find the alighting points of the passengers [12]. Lianfu et al. stated that the data on smart cards played a very important role in urban public transportation planning, and showed that many of the data obtained from complex travel surveys could be obtained by analyzing the data on smart cards [6]. Studies conducted using smart cards in the public transportation sector are grouped under three headings: strategic,

ISSN 3030-3907

tactical and operational [8]. In the study where long-term studies are called strategic, it is emphasized that long-term big data used in strategic studies yields considerably more results compared to other data collection methods [1]. In addition, it is stated that quite successful results have been obtained especially in marketing and customer loyalty issues, together with the acquisition of personal information of smart card owners as a result of other studies [14]. At the tactical level, there are mostly studies related to improving the service. These studies are generally aimed at determining bus times and routes depending on passenger movements [15]. The critical point in studies at the tactical level is that there is no data regarding the stop where the passengers get off. For this reason, some algorithms prepared for estimating the passenger's get off stop are frequently used in the literature [7, 11]. Operational level studies generally consist of research on performance indicators such as daily vehicle-km, passenger-km, line, and personal travel times [13].

METHODOLOGY

Using smart card data, determining the density of the routes followed by passengers in public transportation systems and creating OD matrices are quite difficult due to the lack of any data on passengers' disembarkations. The lack of data on passengers' disembarkations in smart card systems requires estimating their disembarkations using other passenger data and accepting some assumptions for these estimates. The estimated values that change as a result of the assumptions made are quite important since they will directly affect transportation planning.

In this study, 5 basic assumptions are accepted and each assumption is summarized below with its requirements.

Assumption 1: The stop where the passenger gets off is the closest stop to the next boarding stop. The next boarding stop can be at most 3 km away from the disembarkation stop, otherwise Assumption 5 is valid.

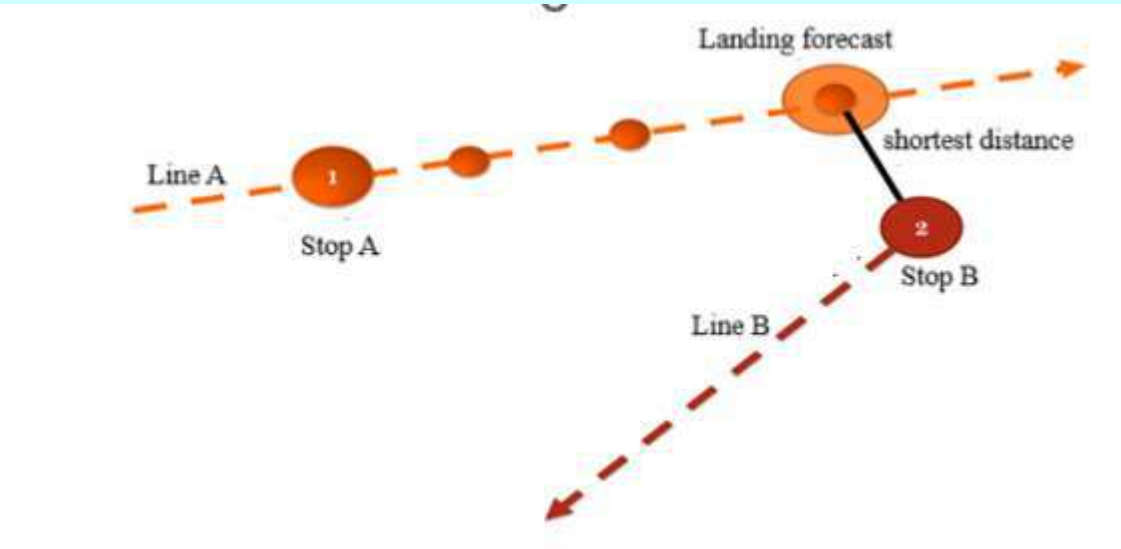


Figure 1. Example of Assumption 1

ISSN 3030-3907

The basic logic in estimating passengers' getting off stops is to use information about the line they boarded and the stop where the passenger boarded next. Since passengers will tend to get off at the stop closest to their next boarding stop, it is assumed that passengers will get off at the stop closest to the stop they boarded next on the line they used. However, since the assumption will lose its validity if the distance difference between two stops is too large, a distance limit of 3 km has been introduced and the reliability of the assumption has been increased in this way. The distance between two stops has been calculated using the Haversine formula.

Assumption 2: Passengers return to the place where they first boarded at the end of the day.

The vast majority of passengers have a regular travel habit. Especially in home-work and home-school trips, the person's last journey of the day ends where the first journey of the day began. As a result of this situation, which is more reliable especially in weekday trips, it is assumed that the passenger's last boarding ends at the stop where he/she first boarded during the day according to the criterion specified in Assumption 1 (max. 3 km).

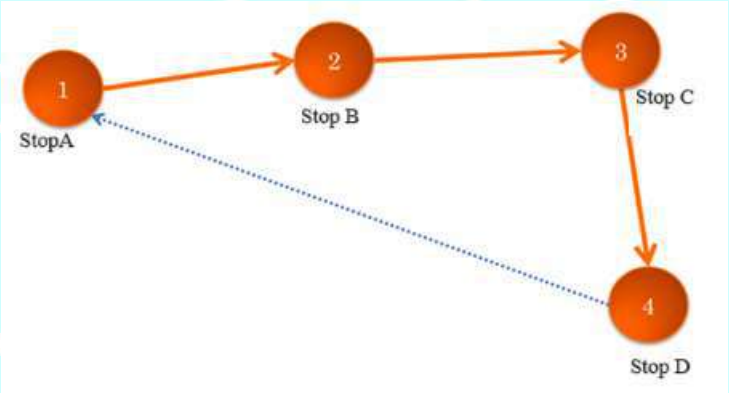


Figure 2. Assumption 2 Example

Assumption 3: Passengers who board only once during the day get off at the last stop of the line they boarded.

Approximately 8%-9% of the passengers board only once during the day. Since these passengers do not have a subsequent boarding, the first two assumptions become invalid. Although it is seen in the literature that passengers who board once a day are generally excluded from the analysis, instead of ignoring the boardings of these passengers, assuming that these passengers get off at the last stop of the line they boarded would be a pessimistic but more realistic approach in terms of vehicle occupancy.

ISSN 3030-3907

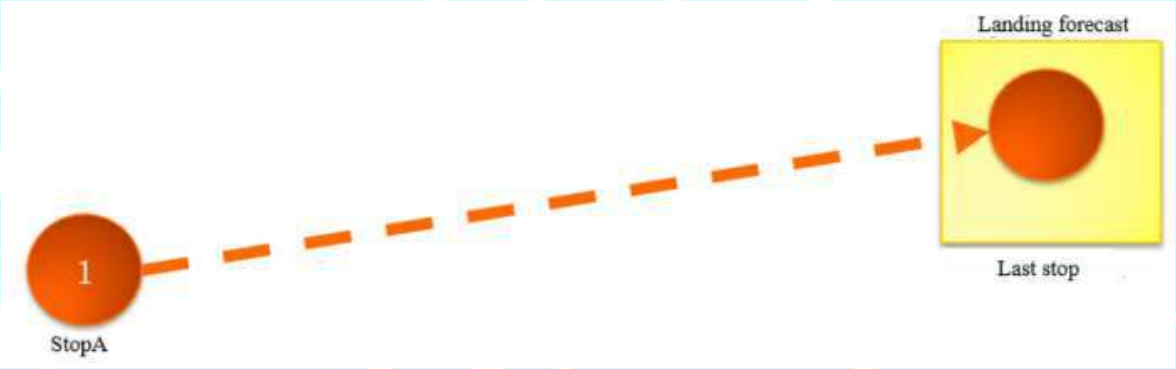


Figure 3. Example of Assumption 3

Assumption 4: If there is more than one boarding for a card in the same trip, the first boarding is included in the analysis, and the other boardings are terminated at the last stop.

Due to reasons such as some of the passengers not having a smart card or not having sufficient balance, etc., there may be more than one boarding data for a card in the same trip. Based on the assumption that these passengers do not move together, Assumptions 1 and 2 are applied for the first boarding data, while it is assumed that the other boardings are disembarked at the last stop.

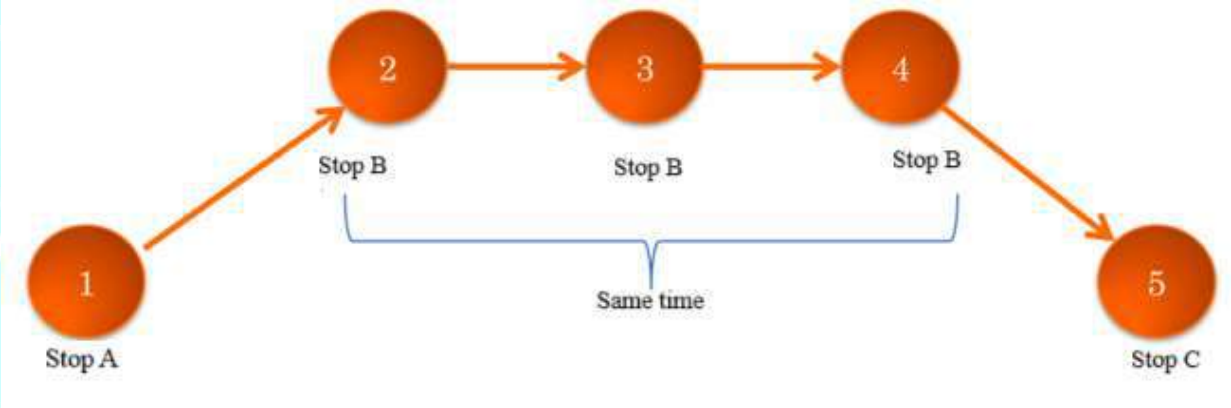


Figure 4-a. Example of Assumption 4 (Current Situation)

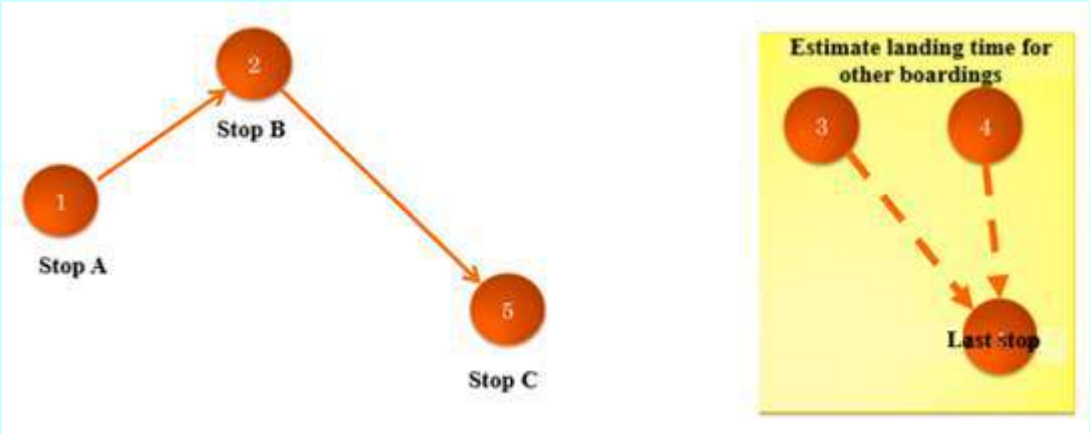


Figure 4-b. Example of Assumption 4 (Assumption Result)

Assumption 5: Passengers whose landing stop cannot be predicted with Assumptions 1, 2 and 4 get off at the last stop of the line they boarded.

ISSN 3030-3907

Due to the distance criterion in Assumption 1 and Assumption 2, for boardings whose landing stops cannot be predicted, and for boardings where the same card has more than one boarding at the same time, it is assumed that these journeys end at the last stop of the line where boarding is done with a pessimistic approach. The reason for the pessimistic approach in all assumptions is to be able to determine both the passenger density in the bus and the OD matrix values with realistic upper limits.

In line with the assumptions summarized above, all procedures for compiling the data within the scope of the study were developed on SQL Server 2012 and Oracle 11.2 databases, and the analysis results regarding vehicle occupancy on a line basis were shared with the planning units through Oracle Business Intelligence system interfaces.

APPLICATION

As an example, within the scope of the application, the passenger density values obtained from the actual boarding numbers and the boarding estimations regarding the boardings on a stop basis for a journey were examined. In the study, the effects of the accepted assumptions on the results were also investigated and the obtained results were compared with the trip-chaining approach used in the literature for boarding estimation. Within the scope of the study, the current boarding data of 8,144,690 boardings made with 1,032,046 different cards in public transportation vehicles in Tashkent between September 22-26, 2021 were analyzed.

With a population of approximately 3,500,000 in the city center, Tashkent has a public transportation system where metro, ferry, bus and light rail systems are fully integrated, and an average of 1,700,000 boardings are made per day.

In the city where the smart card system has been valid in all public transportation vehicles for over 10 years,

almost all passengers have their own smart cards. Thanks to these cards, data regarding many variables such as card type (full, student, disabled etc.), boarding time, boarding stop, type of vehicle boarded (metro, ferry etc.), line number, boarding order, boarding type (paid/transfer) are recorded for each boarding.

In the application, firstly, the boarding data of the return journey starting at 07:35 on the line numbered 169 on 23.09.2014 was examined as an example and the boarding and disembarkation numbers obtained on a stop basis are summarized.

When evaluated in general, it is assumed that approximately 15% of all boardings made during the day are outside the distance criterion (3 km) specified in Assumption 1 and get off at the last stop of the line they boarded. In addition, the total number of boardings made by passengers who board only once a day constitutes 7.9% of the total number of boardings and, as required by Assumption 3, the assignment of the disembarkation stop for these passengers is made to the

ISSN 3030-3907

last stop of the line they boarded. Similarly, as required by Assumption 4, the share of those who are assigned to the last stop and board with the same card on the same journey is 2.5% of the total. In summary, the ratio of boardings whose disembarkation stop could not be predicted and were assigned to the last stop to all boardings is 25.4%, and the minimum success rate for estimating the disembarkation stop is 74.6%.

CONCLUSION AND EVALUATION

In urban public transportation, information about passengers' boarding and disembarkation is of great importance in terms of transportation planning. While smart cards, which have become increasingly widespread in public transportation in recent years, provide detailed information about passengers' boarding, since there is no direct data about disembarkations, it has become necessary to estimate the passengers' disembarkation locations. Although there are various algorithms in the literature to estimate passengers' disembarkation stops, in this study, a method was worked on to estimate the disembarkation stop for bus passengers in Tashkent and the results obtained were shared. Accordingly, while the disembarkation stop was estimated for approximately 75% of all boardings, no definite estimate could be made for the rest, and the last stop of the line where these passengers boarded was determined as the disembarkation stop. However, when it is taken into consideration that 10% of the 25% of the people whose landing stop cannot be predicted are people who board only once a day and who board more than once with the same card at the same time, it can be evaluated that the prediction success of the method used is more than 75%.

It is clear that the success rate will change depending on the assumptions used to predict the landing stop, and rearranging the assumptions closer to reality will increase the success rate. In this direction, various extensions such as using the boarding data of the passengers who board only once a day, the previous day or the next day, and/or using the direction statistics of the stop they boarded for passengers whose landing stop cannot be predicted, will increase the success rate in predicting the landing stop.

In this study, the landing stop prediction for passengers boarding only on bus lines was studied, and it is also possible to predict the landing stops for passengers using other modes of transportation such as metro, light rail, and ferry with similar assumptions. Although studies to be conducted for other modes have some advantages and disadvantages compared to studies conducted on a bus basis, an integrated study covering all modes will provide significant convenience in obtaining city-wide passenger trend statistics and OD matrix studies.

ISSN 3030-3907

RESOURCES

1. Agard, B., Morency, C. ve Trepanier, M., 2006, "Mining public transport user behaviour from smart card data", The 12th IFAC Symposium on Information Control Problems in Manufacturing (INCOM), Saint-Etienne, France.
2. Bagchi, M., White, P.R., 2004, "What role for smart-card data from bus system?", Municipal Engineer, 157, pp. 39–46.
3. Bagchi, M., White, P.R., 2005, "The potential of public transport smart card data" Transport Policy, Vol. 12, pp.464–474.
4. Blythe, P., 2004, "Improving public transport ticketing through smart cards", Proceedings of the Institute of Civil Engineers, Municipal Engineer, 157, pp. 47-54.
5. Deakin, E., Kim, S., 2001, "Transportation technologies: implications for planning", University of California Transportation Center, 536, pp. 27.
6. Lıanfu, Z., Shuzhı, Z., Yonggang, Z., Zıyın, Z., 2007, "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.
7. Munizaga, M., Palma, C., Mora, P., 2010, "Public transport OD matrix estimation from smart card payment system data", 12th World Conference on Transport Research, Lisbon, Paper No. 2988.
8. [8] Pelletier, M.P., Trépanier, M., Morency C., 2011, "Smart card data use in public transit: a literature review", Transportation Research Part C, Vol. 19, pp. 557-568.
9. Shelfer, M., Procaccino, J.D., 2002, "Smart card evolution", Communications of the ACM, Vol. 45, No:7, pp. 83-88.
10. Trépanier, M., Chapleau, R., 2006, "Destination estimation from public transport smartcard data", 12th IFAC Symposium on Information Control Problems in Manufacturing (INCOM), Saint-Étienne, France.
11. Trépanier, M., Tranchant, N., Chapleau, R., 2007, "Individual trip destination estimation in a transit smart card automated fare collection system", Journal of Intelligent Transportation Systems, Vol. 11, No:1, pp.1-14.
12. Trépanier, M., Vassiviere, F., 2008, "Democratized smart card data for transit operators", 15th World Congress on Intelligent Transport Systems, Washington D.C.
13. Trépanier, M., Morency, C., Blanchette, C., 2009, "Enhancing household travel surveys using smart card data", 88th Annual Meeting of the Transportation Research Board, Washington, 15.
14. Trépanier, M., Morency, C., 2010, "Assessing transit loyalty with smart card data", 12th World Conference on Transport Research, Lisbon, Paper No. 2341.

ISSN 3030-3907

15. Utsunomiya, M., Attanucci, J., Wilson, N., 2006, “Potential uses of transit smart card registration and transaction data to improve transit planning”, Transportation Research Record: Journal of the Transportation Research Board 1971, pp. 119-126.

ISSN 3030-3907