ORIGIN AND DESTINATION INFERENCE OF BUS PASSENGERS: ISTANBUL CASE STUDY

by

Akif Fidanoğlu

B.S., Civil Engineering, Boğaziçi University, 2011

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Civil Engineering
Boğaziçi University
2015

ACKNOWLEDGEMENTS

I would like to thank first my thesis advisor Assist. Prof. Ilgın Gökaşar for her guidance and attention. Without her support and enthusiasm, this study could have never been completed. She provided me great educational experience which I could not be able to get from anyone else.

I would like to thank also to the member of thesis committee, Assoc. Prof. Nilüfer Özyurt Zihnioğlu and Assoc. Prof. Hakan Yaşarcan for their valuable guidance.

This study is generously supported by the Istanbul Metropolitan Municipality (IBB), Istanbul Municipality Information Technologies (BELBIM) and Istanbul Electric Tramway and Tunnel Establishments (IETT).

I am so grateful to my parents especially my mother. Without their support and courage they gave me unconditionally, I would not be able to accomplish anything.

ABSTRACT

ORIGIN AND DESTINATION INFERENCE OF BUS PASSENGERS: ISTANBUL CASE STUDY

With the advance technology used in the public transportation systems it has become much easier to monitor the trips of the passengers. Most of the public transportation systems record the passengers when they start their trip with the help of Automated Data Collection (ADC) systems installed to the vehicles. By this way, transit agencies records the origins of the passengers' trips. However, it is a bit challenging issue to detect the passengers' destination especially for the bus passengers since there is no data recorded when the passengers exit from the system. This study explores the methods to infer the destination of the passengers and to generate the Origin-Destination (OD) matrices for the bus passengers by conducting a case study for a single bus route in İstanbul. To estimate the destinations of the passengers, several assumptions have been set in the previous studies. In this study, these assumptions were used to infer the alighting location of the bus passengers with further assumptions. New methods are established to infer the destination location of the trips which couldn't be estimated with the present methods. The results are compared with the outputs of the surveys conducted on the studied route.

ÖZET

OTOBÜS YOLCULARININ BAŞLANGIÇ VE VARIŞLARININ ÇIKARIMI: İSTANBUL DURUM ANALİZİ

Toplu ulaşım sistemlerinde kullanılan yüksek teknolojiler ile birlikte, yolculukların takip edilmesi çok daha kolay bir hal almıştır. Birçok toplu ulaşım sistemlerinde yolcular yolculuklarına başladıklarında araçlara kurulan Otomatik Veri Toplama (ADC) sistemleri sayesinde kaydedilmektedirler. Bu şekilde ulaşım daireleri yolcuların seyahatlerinin başlangıçlarını kaydetmektedir. Fakat yolcular sistemden çıkarken herhangi bir verinin kaydedilmemesi sebebiyle özellikle otobüs yolcularının varış noktalarının tespiti zorlu bir konudur. Bu çalışma İstanbul'daki tek bir hatta gerçekleştirilen durum analizi ile yolcuların varış yerlerinin tespitini sağlayan metotları incelemiştir. Önceki çalışmalarda yolcuların varış yerlerinin tahmini için bazı varsayımlar yapılmıştır. Bu varsayımlar ilave kabuller ile birlikte otobüs yolcularının indiği yerleri bulmak için bu çalışmada da kullanılmıştır. Mevcut yöntemler ile varış noktası bulunamayan yolculuklar için yeni metotlar geliştirilmiştir. Sonuçlar çalışılan hatta yapılan anket verileri ile karşılaştırılmıştır.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

ABSTRACT iv

ÖZET v

LIST OF FIGURES ix

LIST OF TABLES xi

LIST OF ACRONYMS/ABBREVIATIONS xiv

1. INTRODUCTION 1

  1.1. General 1

  1.2. Literature Review 1

  1.3. Automated Data Collection (ADC) Systems 3

    1.3.1. Automated Fare Collection (AFC) Systems 4

    1.3.2. Automated Vehicle Location (AVL) Systems 5

    1.3.3. Automated Passenger Counting (APC) Systems 7

  1.4. Advantages of ADC Data in OD Matrix Estimation 8

  1.5. Disadvantages of ADC Data 10

  1.6. Type of ADC Systems 10

    1.6.1. Systems Which Record Only Boarding Location 10

    1.6.2. Systems Recording Both Boarding and Alighting Location 11

2. METHODOLOGY 12

  2.1. Origin Inference 12

  2.2. Destination Inference 14

3. ISTANBUL CASE STUDY 18

  3.1. Characteristics of 11 L (Bulgurlu-Uskudar) Route 18

  3.2. ADC Data Analysis 24

3.2.1. Properties of ADC Data 24

3.2.2. Ticket Types 25

3.2.3. The Difference of Ridership on Weekdays and Weekends 28

3.3. Inference of Alighting Locations 29

4. ALIGHTING INFERENCE FOR INTERCHANGES 30

4.1. Interchange in 2 Hours 30

4.1.1. Bus to Metro Interchanges 32

4.1.2. Bus to Bus Interchanges 34

4.2. Interchange After 2 Hours 39

4.3. Interchange to the Same Bus 40

5. ALIGHTING INFERENCE FOR LAST TRIPS 44

5.1. First Method 44

5.2. Second Method 45

5.3. Last Trip and No Trip on the Other Days 46

6. PROPOSED ALGORITHM 48

7. INFERRED ORIGIN & DESTINATIONS 53

7.1. Inference Rates of Methods 55

7.2. Inference Rate in Each Direction 56

7.3. Origin and Destination Matrices 57

7.4. Validation 63

7.4.1. Validation of the Proposed Algorithm 63

7.4.2. Validation of the Inferred Results with the Surveys 63

8. CONCLUSIONS AND RECOMMENDATIONS 68

8.1. Comments on the Key Findings 68

8.1.1. Times of the Last and First Trips of Commuters 68

8.1.2. Times of Repetitive Use of the Same Smartcard 69

8.1.3. Interchange Time 71

8.1.4. Passengers of Single Trip on All Day 72

8.2. Conclusions 72

APPENDIX A: ORIGIN AND DESTINATION MATRICES 75

REFERENCES 80

LIST OF FIGURES

Figure 1.1. Card reader in AFC system of Istanbul. (İETT)..5

Figure 1.2. AVL and Passenger Information System Scheme in Istanbul (BELBİM, 2014)..6

Figure 1.3. Screenshot of MOBIETT Application of IETT (IETT, 2014). 7

Figure 1.4. APC System Scheme (infodev.ca). 8

Figure 2.1. Location of the Coordinates Taken from AVL Data.....................................14

Figure 2.2. Process for Destination Inference (Wang 2010). 15

Figure 3.1. Route Schematic of 11 L in Bulgurlu Direction..19

Figure 3.2. Route Schematic of 11 L in Üsküdar Direction. 19

Figure 3.3. Location of Kültür Merkezi Bus Stop and Its Corrected Route. 21

Figure 3.4. Istanbul BRT Line Map 2014 (IETT). 22

Figure 3.5. Routes of Subways in Istanbul. 23

Figure 3.6. Routes of Ferries in Bousphorus. 23

Figure 3.7. Percentages of Different Ticket Types in All Trips. 26

Figure 3.8. Percentages of Different Ticket Types in 11 L Trips. 27

Figure 3.9. Number of Trips of the Passengers Who Use Discount Ticket vs. Time........27

Figure 3.10. Number of Trips of the Passengers over 65 Age vs. Time............................28

Figure 3.11. Number of 11 L Trips on Each Day vs Time. 28

Figure 3.12. Type of 11 L Trips and Their Percentages. 29

Figure 4.1. Time Interval between 11 L Trip and Next Trip. 31

Figure 6.1. The Proposed Algorithm for Inference of Alighting Location. 49

Figure 6.2. The Proposed Algorithm Steps for Determining the Repetitive Use of the Same Smartcard. 50

Figure 6.3. The Proposed Algorithm Steps for the Last Trips of the Day. 51

Figure 6.4. The Proposed Algorithm Steps for the Trips Having an Interchange Trip on the Same Day. 52

Figure 7.1. Inference Rate of Each Cluster. 56

Figure 7.2. The Number of Boarding and Alighting at Each Stop of 11 L Route in Bulgurlu Direction. 61

Figure 7.3. The Number of Boarding and Alighting at Each Stop of 11 L Route in Üsküdar Direction. 62

Figure 7.4. Rates of Inferred vs Surveyed Boarding and Alighting at Each Stop of 11 L Route in Üsküdar Direction. 66

Figure 7.5. Rates of Inferred vs Surveyed Boardings and Alightings at Each Stop of 11 L Route in Bulgurlu Direction. 67

Figure 8.1. Times of the Last and First Trips of Commuters. 68

Figure 8.2. Time Intervals Between the Last and First Trip of the Day. 69

Figure 8.3. Times of Repetitive Use of the Same Smartcard. 70

Figure 8.4. The Total Ridership in Hours for All Trips. 71

Figure 8.5. Consecutive Boarding Times Difference. 71

LIST OF TABLES

Table 2.1. Recorded Boarding Location in AFC Data and Calculated Boarding Location from AVL Data. 13

Table 2.2. Coordinates in AVL Data and Closest Bus Stop in 11 L Route. 13

Table 3.1. Bus Stops of 11 L Route in Both Directions. 20

Table 3.2. An Example ADC Data. 24

Table 3.3. Number of Trips by Ticket Types in All Trips and 11 L Trips. 26

Table 3.4. Type of 11 L Trips and Their Percentages. 29

Table 4.1. Type of Interchanges and Number of Trips in Each Group. 30

Table 4.2. Fare of Interchanges in 2 Hours. (İETT, 2014). 30

Table 4.3. Examples of Inaccurate Direction Record in ADC Data. 33

Table 4.4. Inference Results of Interchanges to Metro within 2 Hours. 34

Table 4.5. Examples of Missing Transfer Trips. 35

Table 4.6. Destination of Ferries and Marmaray & Assigned Bus Stops on 11 L route. 36

Table 4.7. Inaccurate Record of Boarding in AFC Data and Assigned Boarding Location from AVL Data. 37

Table 4.8. Inference Results of Interchanges to 11 L Route or Other Routes with Common Stops within 2 Hours. 38

Table 4.9. Inference Results of Interchanges to Other Buses within 2 Hours. 38

Table 4.10. Inference Results of Interchanges to 11 L Route or Other Routes with Common Stops after 2 Hours....39

Table 4.11. Inference Results of Interchanges to Other Routes after 2 Hours. 40

Table 4.12. Examples of Repetitive Use of the Same Smartcard. 42

Table 4.13. Inference Results of the Interchanges to the Same Bus. 43

Table 5.1. Inference Results of the Last Trips Using the First Method. 45

Table 5.2. Limits for the Time Differences between the Day of 11 L Trip and the Day of the Next trip. 46

Table 5.3. Inference Results of the Last Trips from Second Method. 46

Table 5.4. Inference Results of Last Trips (No Other Trips on Other Days). 47

Table 7.1. Overall Inference Results. 53

Table 7.2. Inference Results of Each Group 55

Table 7.3. Inference Results of Each Direction. 57

Table 7.4. Origin and Destination Matrix of 11 L Route in Bulgurlu Direction. 59

Table 7.5. Origin and Destination Matrix of 11 L Route in Üsküdar Direction. 60

Table 8.1. Type of Ticket Used by the Passengers Recorded Only Once on All Days. 72

Table A.1. Origin and Destination Matrix of 11 L Route in Bulgurlu Direction for the Weekdays. 76

Table A.2. Origin and Destination Matrix of 11 L Route in Bulgurlu Direction for the Weekend. 77

Table A.3. Origin and Destination Matrix of 11 L Route in Üsküdar Direction for the Weekdays. 78

Table A.4. Origin and Destination Matrix of 11 L Route in Üsküdar Direction for the Weekend. 79

LIST OF ACRONYMS/ABBREVIATIONS

ADC	Automated Data Collection
AFC	Automated Fare Collection
APC	Automated Passenger Counting
AVL	Automated Vehicle Location
BELBİM	Istanbul Municipality Information Technologies
CTA	Chicago Transit Authority
IETT	Istanbul Electric Tramway and Tunnel Establishments
IBB	Istanbul Metropolitan Municipality
IDO	Istanbul Sea Buses
LBSN	Location Based Social Network
OD	Origin-Destination

1. INTRODUCTION

1.1. General

With the help of advance transportation technologies installed to the public transportation systems it has become much easier for the transportation agencies to take necessary actions and implement dynamic measures. Especially Automated Data Collection (ADC) systems have become very helpful to store the data of the passengers and the trips. These systems enable the transit planners to make several analysis about the trips of the public transportation passengers and by the huge amount of data recorded in the systems. One of the main purposes of these analysis is to estimate the origin and destination of the passengers.

Origin and destination matrices are used as key inputs in transportation planning and operations. They give useful information about the interchange points in the transportation system, passenger flow during the day and the location of the residences and working areas. OD matrices can also be considerably helpful for many other decision makers who deal with the topics about city planning.

There have been several studies conducted about this issue. Most of them, like this study, mainly focused on the case studies and analyzed the ADC data gathered from the studied transportation systems.

1.2. Literature Review

Barry *et al.* (2002) proposed a methodology that is used to estimate origin and destination of the passengers by using MetroCard information in New York City. He applied a set of straightforward methods to each set of MetroCard to assign a destination for every origin station. He validated his assumptions at very high rate with the travel diary information stored by the New York Metropolitan Transportation Council.

Cui (2006) aimed in his study to create a model to estimate a network level bus passenger origin OD matrix. He used the ADC data from Chicago Transit Authority (CTA) to make OD estimation for the bus network in Chicago. For the inferences of origins and destinations, he applied the methodology at single route level and network level. His study for transit rides in public transportation of Chicago was based on the trip chaining OD estimation method. He achieved to infer the high portion of the origins and destinations in his study.

Zhao *et al.* (2007), developed a method to estimate the origin and destination locations of the rail passenger trips with the automated fare collection (AFC) data supplied by the Chicago Transit Authority (CTA). During his study, he also generated a software to assist the application of his proposed algorithms. He suggested the integration of the automated fare collection data of CTA which stores the trip transactions and the automated vehicle location data of CTA which records the vehicle location to infer the boarding station ID of the passenger. In his study, both the rail to rail trip sequence and rail to bus cases are studied with the help of integration of AVL and AFC data.

Trépanier *et al.*, (2007), used the smartcard data of Gatinueau, Canada to estimate the destination of the passengers in his study. In his study, applied model gave a considerably high rate of successful inference for destinations of the passengers.

Wang (2010), made OD estimations with the case studies for several routes in London. In the inference of boarding location, similar to Zhao (2007), he combined the iBus data of the buses and the Oyster data of the passengers. These two data consist of the AVL and AFC data, respectively. He also used the similar algorithm with previous studies to estimate the alighting locations in London. After the inference of alighting and boarding location of the passengers using the studied routes, he further analyzed the interchange times. He questioned the appliance of a fixed temporal threshold for the consecutive trips to be identified as linked trips. He stated importance of in-vehicle travel time and route headways in determining of interchange time and study the interchange times of the trips in his London case study.

Ma *et al.* (2013) suggested a very helpful data mining method to analyze the temporal travel patterns and regularity of the passengers in their public transportation use in Beijing, China. He analyzed the multi-day smartcard ADC data of the passengers, determined the trip chains of the passengers with the consideration of the spatial and temporal relationships and inferred the travel pattern and the travel regularity of the passengers by different methods. The algorithm he proposed for travel pattern and travel regularity mining is claimed to be useful to improve the accuracy of the origin and destination inference methods.

Jun *et al.* (2013) proposed a new method different from the methods used by Barry *et al.* (2002) and Cui *et al.* (2007), but similar to Ma *et al.* (2013) in the sense of regularity clustering, to infer the origin and destination matrices of the commuters. He used in his model the ADC data of the routes in Nanning City, China. Different from the previous studies, he analyzed and made an OD estimation for the passengers who make only one ride during the day by estimating the residences and work places of commuters and obtaining statistics on OD of the commuters during morning and evening peak hours.

Yang *et al.* (2014) proposed a model to infer OD matrix for non-commuting trips by the use of Foursquare (Most commonly used location based social network application) user check-in data in the Chicago urban area. In his study, advantages of the location based social network data over the traditional OD inference methods in terms of sample size, cost and real-time updating are stated. Validation for the usefulness of the LBSN data in the inference of origin and destination matrices is made by the comparison of trip length frequency distribution. At the end of study, very satisfying results are found for LBSN data to be used in long run travel demand changes.

1.3. Automated Data Collection (ADC) Systems

Automated Data Collection systems become widely used in public transportation systems with the implementation of the technological innovation over the past decades (Cui 2006). ADC systems have become popular as they provide effective and cheaper alternative to the conventional data collection methods. Commonly used examples of ADC systems are namely Automated Fare Collection (AFC), Automated Vehicle Location (AVL) and Automated Passenger Counting (APC) systems.

1.3.1. Automated Fare Collection (AFC) Systems

AFC systems become widely used in almost every metropolitans because of the advantages they provide. The main advantage is the reduction in costs resulting from the tickets, tokens and clerks employed to sell the tickets and tokens. It also serves very beneficial data for statistical analysis. It stores passengers' unique smart card information and the time of the travel.

Depending on the fare collection modes, AFC systems provide different information about the trips of passengers. In distance based fare collection systems, AFC data has information not only about the entry of passenger to the line but also the exit of him. In İstanbul, Metrobüs line is an example for entry-exit control system.

However, in most of the cases all around the world and in İstanbul also, AFC systems record only the entry to the system. In this systems, information regarding the exit of passengers is unknown and can be estimated only from the other trips of passengers.

In both entry-exit and entry only systems, information about the location of vehicle is taken from the AVL systems. In some applications, AFC systems are integrated with AVL systems and they record the location information. In the cases that AFC systems are not integrated with AVL systems, to derive the locations of related AFC data, matching of these two data set must be done manually by the transportation planners.

Municipal Data Processing Corporation of İstanbul, which is known as BELBİM A.Ş. was established in 1987 for data processing, project design, mapping and planning and other services for the municipal administration of İstanbul. In 1994, BELBİM introduced a smart ticket (Akbil) as AFC system to eliminate the problems in incoordination of different transportation agencies and difficulties in rendering statistical data. Full-scale application began in 1995 at İstanbul Sea Busses (IDO) Corp. and after that city bus lines, private mass transport lines, metro lines and all other transportation systems are also equipped with Akbil system. In 2007, İstanbulkart which is a smartcard stores passenger’s personal info also, is issued by BELBİM in almost all public transportation systems in İstanbul (Figure 1.1). (IBB, 2014)

Figure 1.1. Card reader in AFC system of Istanbul. (İETT)

1.3.2. Automated Vehicle Location (AVL) Systems

Automated Vehicle Location (AVL) systems through Global Positioning System (GPS) provide information about the position of the vehicle which AVL systems are installed. Since most of the AVL systems records the location of the vehicle in frequent intervals, it is easy for transit planners to detect the exact location of a vehicle and the closeness of the vehicles to the stations efficiently.

AVL records can be stored in the equipment on board or be connected to central computers by the help of wireless connection and allows transit agencies to access real-time vehicle positions.

In public transportation system of İstanbul, most of the vehicles are equipped with AVL systems. With the help of dynamic data recording of AVL, as seen in Figure 1.2, several passenger information services are applied to the public transportation system by the transit agency such as:

- Passengers on board are informed of the location of the bus, stations on the route of the bus and the next station of the bus by the help of screens installed in the buses.
- Passengers at the stations are informed by the screens installed in some of the major bus stops in İstanbul that show the estimated arrival time of each bus route to the station which the screen is installed to.
- Passengers who use smart phone are informed about the station names and scheduled departure time of every single route in İstanbul by the “MOBİETT” application serviced by the transportation agency (Figure 1.3). Estimated arrival time of buses to the stations are also available to the passengers with this application.

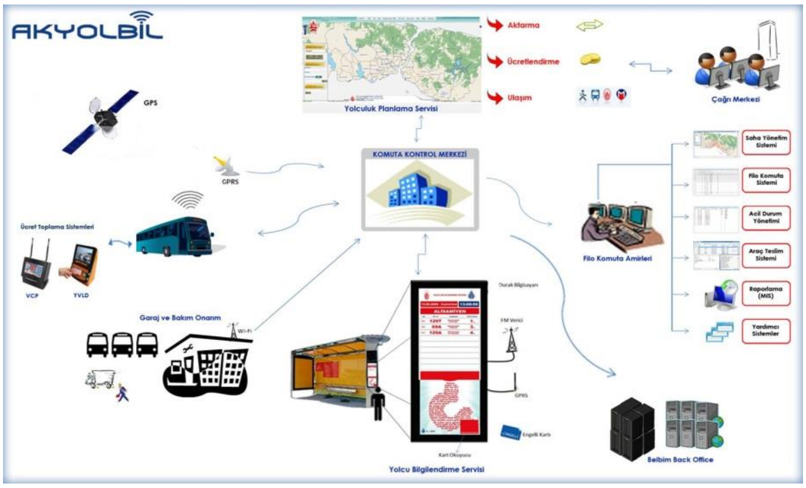


Figure 1.2. AVL and Passenger Information System Scheme in Istanbul (BELBİM, 2014).

Figure 1.3. Screenshot of MOBIETT Application of IETT (IETT, 2014).

1.3.3. Automated Passenger Counting (APC) Systems

Automated Passenger Counting system mainly consists of electronic machines that count passengers board and alight at bus stops. With the coordination of AVL systems, APC systems give the number of passengers who board and alight in each station.

APC systems operate with the help of sensors installed at the each door of the bus. When the passenger boards or alights, he breaks the infrared beam and the computer records the passenger.

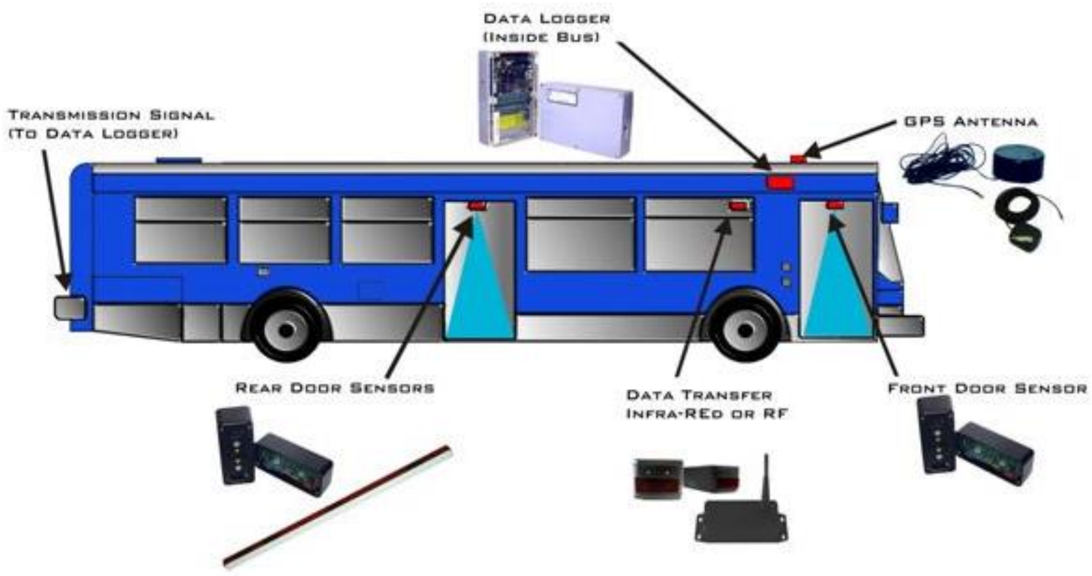


Figure 1.4. APC System Scheme (infodev.ca).

It is not necessary to install the system at every bus in public transit. With the implementation of APC system to particular portion of bus fleet, ridership information for every route can be collected by switching the route of buses in regular basis.

However APC systems generate robust ridership information at every stop. The information of passengers individually cannot be gathered from the APC systems. Therefore, to build OD matrix for transit passengers by using the records from APC system is not possible.

The ADC systems which are already installed by transit agency have other purposes in public transit. To use the data from these systems in the estimation of OD matrices have no cost. On the other hand, since the installation of the APC system is aimed only to have total ridership information at each stop, the installation and software cost of APC systems can be considered as additional costs by transportation agency.

1.4. Advantages of ADC Data in OD Matrix Estimation

OD matrices can be used for the public transportation planning by the transit planners in a transit agency of a city in many aspects. It is very beneficial for transit planners to know the origins and destinations for every passengers individually. By the help of this information, flow of passengers in a city during different times of the day can be examined. Also, the interchange stations that passengers prefer to make their interchanges in, can be detected.

Conventional way of gathering boarding and alighting location of the passengers is the passenger surveys which are difficult to conduct, have low frequency, give extremely less and unreliable information. To get the required information for the production of OD matrices from the conventional surveys, it should be conducted in a high scale. Main goal of these surveys; however, normally is not to produce OD matrices primarily in practice.

Production of OD matrices by using ADC data has several advantages over traditional surveys as:

- Significant reduction in cost of obtaining OD matrices,
- Obtaining individual trip information of passengers by the help of passenger smartcard,
- Easiness to update the data and run the process more frequently.
- Providing continuous trip information of passengers in transit system in a larger scale.
- Conveniently utilizing in more comprehensive studies.

One of the main and the most important advantages of ADC systems is the ability of them to provide much more information about the travels of passengers at significantly lower costs. This is mainly because the information needed to be gathered to generate OD matrices is already being stored in the systems which are installed primarily for different reasons. For example, AFC systems are already installed to the every public transportation systems in İstanbul to collect the fares and the passengers are given the smartcard to use in public transit. AFC systems already records most of the needed information to generate OD matrices. Therefore, with an additional data process OD matrices can be easily obtained from the ADC data.

It is hard to picture the whole system without extremely large sample sizes using conventional surveys. Since the larger sample sizes in surveys means the higher costs, it is not preferred by the transit agencies to conduct high scale passenger surveys. For this reason, most of the time surveys give biased data. However, if the system is installed to the whole transit system and work accurately, ADC systems records the passenger’s information with no sampling error and provides dramatic increase in sample size (Cui 2006).

One the main disadvantages of the conventional surveys is the infrequency of applications. It is preferred by the transit agencies to conduct surveys before and after extensive system changes only. Hence, OD matrices produced with the help of passenger on-board surveys are estimated very infrequently, normally every 5 to 10 years (Barry 2002).

ADC systems, however, store continuous data of the transit systems and it can be easily extracted from the systems. Therefore, it is quite possible to conduct OD estimation at any time since the raw data is always readily available in the system. Transit planners can make analysis before and after every changes in transit systems by the help of continuous ADC data.

1.5. Disadvantages of ADC Data

Normally, ADC systems are not installed and designed to be used to produce OD matrices; therefore, data extracted from the ADC systems are needed to be processed and converted into useful format.

Moreover, unlike the conventional surveys ADC data have no information about the passengers' intentions or purposes of the trips. Therefore, transit planners need to analyze the ADC data further to estimate the purposes and types of passengers' trips. Even then the reliability of the acquired results should be checked with the conventional surveys.

Also as it is seen in this study, ADC system sometimes provide inaccurate data. Especially for the AVL systems, it is very common to have biased data because of the defective records of the installed equipments. Thus, the accuracy of the data gathered from the ADC systems should be checked further.

1.6. Type of ADC Systems

1.6.1. Systems Which Record Only Boarding Location

In some cases, neither the origin nor the destination of individual trips are stored in ADC systems. Among these, there are some cases having Automated Vehicle Location systems. For this case, passengers boarding locations can be found by matching the AFC data and AVL data, which gives the location of the bus from AVL data at the time when passengers enters the system.

For the cases which don’t have AVL systems installed to vehicles, locations of the bus in its route can be found by the schedule of the buses. Since the departure time of the buses are scheduled and the estimated travel time is known between the stops, scheduled arrival time of a bus at the stops on route can be derived. Istanbul Transportation Agency give this information for all stops of every single route in its website. With the known arrival time and the AFC data which has the information of the passengers’ boarding time, by matching these two information, boarding location of each passenger can be found. However this method can give very biased and misleading results because of the deviation in schedule and travel times of buses.

1.6.2. Systems Recording Both Boarding and Alighting Location

In some systems, a distance-based fare collection is used. To achieve that, both the entry and exit location of each passenger should be recorded. After that, the distance between these two points are calculated. Each passenger is charged according to the distance he traveled.

Building the Origin and Destination matrices in these systems is very easy since both the boarding and alighting locations and times are recorded.

There are some examples of this systems in Turkey and abroad. In Seoul, ADC systems records for every trip boarding and alighting locations. This allowed Jang (2010) to analyze the systems in a very detailed way and to determine the interchange stations. This system is used in İstanbul also. In Metrobüs line which is the most commonly used and one of the most congested routes in İstanbul, distance-based fare collection system is introduced in 2009 (ibb.gov.tr). In Metrobus route, the highest fare is taken from the passengers’ smartcards when they enter the system. When they leave the Metrobus, they use their smartcards again in their alighting locations and collect the surplus in fare taken from smartcards at the boarding stops. However, passengers tend to forget to use their smartcards in the alighting locations and this results absence in ADC data.

2. METHODOLOGY

2.1. Origin Inference

Cui (2006), Zhao (2007) and Wang (2010) made the origin inference for the passengers by integrating the AVL data and the related AFC data recorded in the systems in their studies.

Main problem here is the examination of the missing parts of GPS data. The same problem is present in AFC data also. Hence, to get rid of this problem these two datasets should be examined together so that missing data can be found using the relevant information from each dataset. In this analysis, because the municipality records the boarding locations in AFC data, we can derive the boarding locations of the particular passenger from AFC data only. The AVL data of the buses operate in the studied route could be provided only for a small portion of trips. Therefore, AVL data is used in this study only when a direction error is found for the inferences made. However, if all the required AVL data is available, GPS data of the relevant bus at the closest time of the passengers` boarding times can be used for the cases where AFC data is missing boarding locations. By this way, locations of a bus from GPS data can be derived and assigned as the boarding location of the passengers.

Since the records in AFC and AVL generally don't match perfectly, several assumptions can be made to infer the absent boarding locations.

Wang (2010) proposed a three different rule for finding the closest AVL data against AFC data in London. First, the previous stop rule assigned the previous stop before the time of AFC data as boarding stop. Second, the next stop rule assigned the next stop after the time of AFC data as boarding stop. Finally, the closest stop rule assigned the closest stop to the time of AFC data as boarding stop. After the calculation, he obtained the best results from closest stop rule. Wang (2010) applied this rule for assigning the boarding locations of the passengers, because in London AFC data is stored in Oyster system and AVL data is recorded in iBus system separately.

In this study, even though the AVL data is used for the cases in which a direction error is found, AVL data and recorded boarding locations in AFC data is compared for some circumstances to check the consistency. To achieve this, the closest stop of the selected bus route (11 L) to the relevant coordinates recorded in AVL data is to be determined. After this process, some contradictions between the assigned bus stop to AVL data and the recorded bus stop in AFC data is observed. To check the accuracy of the process, the stop of 11 L closest to the coordinates in AVL data and relevant coordinates were determined in the map.

One example is illustrated in Table 2.1. For the specific bus (C-1722) at the given time, the boarding location is recorded as Kısıklı in the AFC dataset. On the other hand, when the relevant recorded coordinates in AVL data (Table 2.2) is located in the map it is found that the closest stop for this coordinates is Dostluk Parkı bus stop.

Table 2.1. Recorded Boarding Location in AFC Data and Calculated Boarding Location from AVL Data.

Date	Stop ID	Bus No	Stop Name	Stop Name from GPS of Bus
15.09.2014 18:40:23	A0291A	C-1722	KISIKLI	DOSTLUK PARKI
15.09.2014 18:40:26	A0291A	C-1722	KISIKLI	DOSTLUK PARKI
15.09.2014 18:40:28	A0291A	C-1722	KISIKLI	DOSTLUK PARKI

Table 2.2. Coordinates in AVL Data and Closest Bus Stop in 11 L Route.

Date	Longitude	Latitude	Stop Name from GPS of Bus
15.09.2014 18:40:28	29,08107	41,01688	B_DOSTLUK PARKI

Figure 2.1. Location of the Coordinates Taken from AVL Data.

Since used coordinates for the assigned stop in AFC data are not available in the ADC data, it is hard to decide which dataset stores the most accurate records. In this study; however, it was observed that for the cases where direction errors occurred in the inferences of alighting locations, AVL data set, if available, gave much more reasonable results than AFC data.

2.2. Destination Inference

Zhao (2007), Cui (2006), Trepanier *et al.* (2007) and Wang (2010) all made the same assumptions for the inference of destination methodology as:

- Passengers don't use private transportation modes between the recorded trips.
- The distance between alighting location of the previous trip and the boarding location of the next trip cannot exceed predetermined level for these consecutive trips to be considered as transit trips.

- Passengers return to boarding location of their first trip with their last trip on that day.

First two assumptions are made in the interchange method which is used to infer the alighting location of the trips with interchanges to other routes. With the help of these two assumptions, boarding location of the next trip is considered as the alighting locations for the previous trips. This method is called as Interchange or Next Trip Method. On the other hand, for the last trips of the days, the third assumption is also taken into consideration and the first trip of the day is taken as the next trip of the studied trip. By this way, passengers are assumed to come back to the location where they start their first trip on the day. For the last trips of the day, other assumptions used in the interchanges above are also still valid. This method is named as Last Trip Method.

Wang (2010) showed the process for destination inference in his study with Figure 2.2.

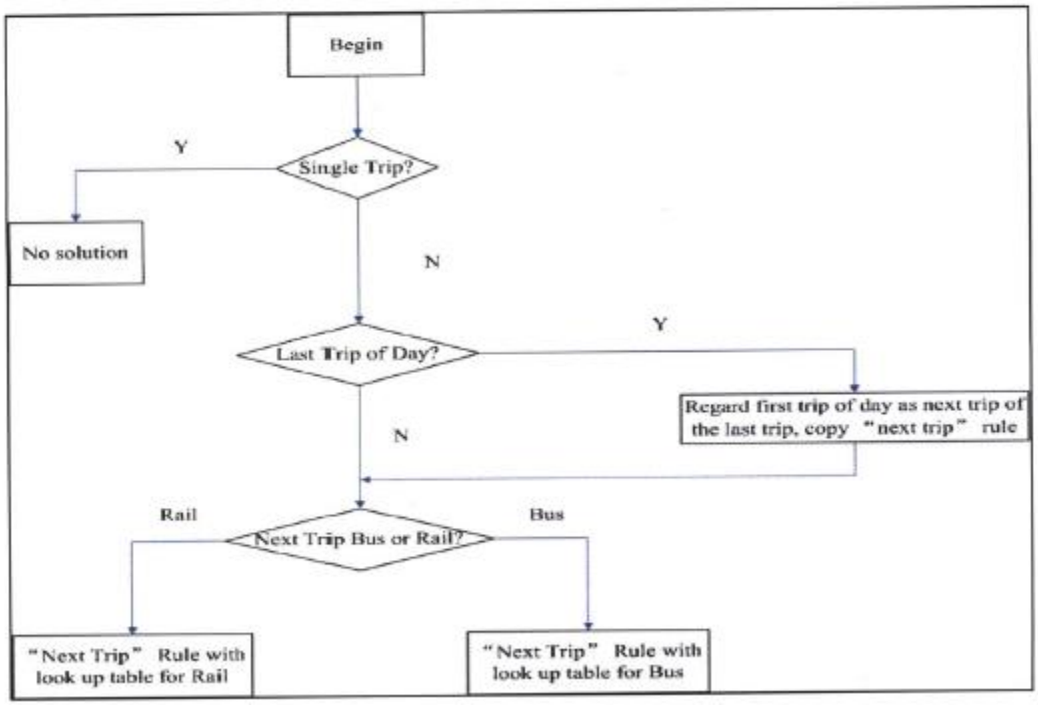


Figure 2.2. Process for Destination Inference (Wang 2010).

In this thesis; however, the methods used in the previous studies for both interchanges and last trips were modified in the proposed methodology. The process carried out in this study for the inference of alighting locations can be summarized as:

- Next trips of the studied trips are found.
- In terms of their next trips, trips were classified into groups.
- For the trips which have no other trip on the analyzed days, no alighting location is inferred.
- For the trips which have a next trip in 2 hours, boarding location of the next trip is checked in terms of the distance to the stops of studied route. If distance to the closest stop of the studied route is below some specified limit or the next trip is made at one of the stops in the route of studied trip, then these stops are considered as the alighting location. In this group, destinations of the main possible routes which is assumed to be missing in the ADC data are also taken into account and the closest stop of the studied route to the origins of these trips are taken as the alighting location if the next trip is made at the location near the destination of these main routes. For this cluster, if at the end of alighting inference, direction error occurs then recorded direction of the studied trip is changed. These additional two assumptions are made for the trips in this group because of the short time interval between the consecutive trips.
- For the interchanges made beyond 2 hours, the same procedure with the previous item is followed. However, for the trips in this group the missing trip assumption and the correction in direction are not made. This is because the time interval between the trips are relatively long and the passenger can reach to his/her recorded next boarding by using different routes.
- For the interchanges made in the same buses are studied in a different group to detect the repetitive use of the same smartcard. If the passenger makes his/her next trip in 60 minutes at the same direction and bus then it is assumed that the cardholder use his/her smartcard for another passenger. Hence for these records no alighting location is inferred.
- For the trips which are the last trips of the day, boarding location of the first trip is checked whether it is one of the studied route’s stations or close enough to the stations of the studied trips. If it is below limits, as in the previous studies also, the

boarding location of the first trip of the day is taken as alighting location of the last trip.

- For the trips which are last trips of the day and no result found with the methods described in the previous item another method is proposed in this study. Since the next trip of all trips are found at the beginning of the study, next trips of these trips are analyzed in terms of their day. If the next trip is made on a day close enough to the day of studied trip then the same procedure used for trips in interchange after 2 hours cluster is applied to infer the alighting location. By this method, many of the trips which are not studied in the methods suggested in the previous studies can be analyzed further. For example, as it is seen in the process for destination inference in Wang's (2010) study for single trips which have no other trips on that day, no result can be inferred. However with the help of proposed method, further analysis can be made for these trips also. But, it should be noted that the number of days between the day of the studied trip and next weekday is taken as the limits for the difference between the day of studied trip and the next trip. The next trip day must satisfy this rule.

The proposed algorithm of the procedure described above is explained in the next chapters.

3. ISTANBUL CASE STUDY

3.1. Characteristics of 11 L (Bulgurlu-Uskudar) Route

In this thesis, 11 L (Üsküdar-Bulgurlu) bus route which runs between Üsküdar and Bulgurlu in the Anatolian part of İstanbul is selected for the analysis. Total length of the route is 9-10 km for each direction with 12 minutes headways during early mornings and 15 minutes daytime headways. Routes of the buses in each direction are shown in Figure 3.1 and Figure 3.2.

11 L route is analyzed in this study because of the following reasons:

- This route runs between the location, mainly consisting of residences and the location which is the one of the most commonly visited places in İstanbul. Therefore, in this route not only the commuters but also the irregular users of this route are expected to be recorded.
- The route has intersections with Metrobüs BRT line, Marmaray subway and the ferries runs in the Bosphorus between the Anatolian and the European part of the İstanbul. These are commonly used public transportation systems in İstanbul. By analyzing the 11 L route, the interchanges to these main routes can be studied.
- Since Üsküdar is a location which has many historical places and shopping centers, it is quite possible to make comments about the interchanges in the location with these features after the analysis of 11 L route.

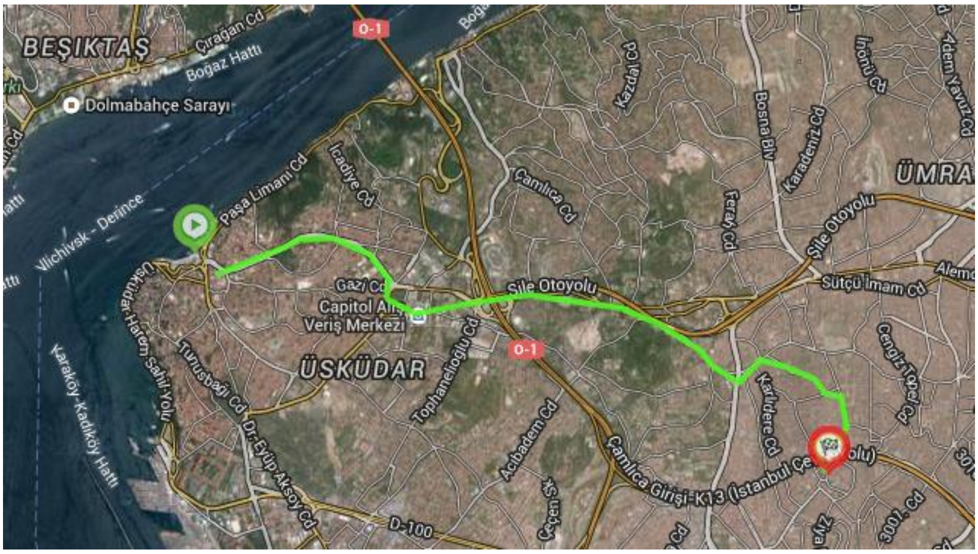


Figure 3.1. Route Schematic of 11 L in Bulgurlu Direction.

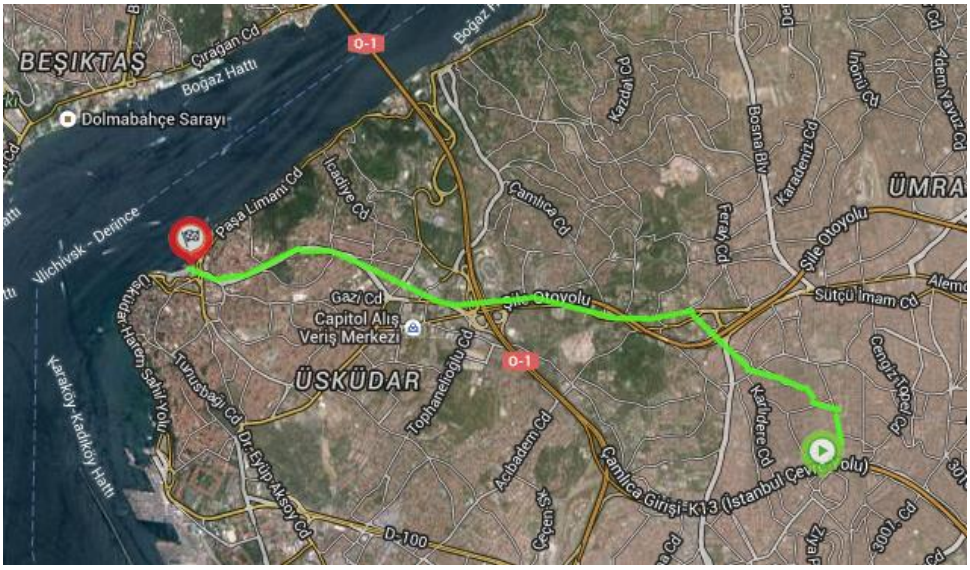


Figure 3.2. Route Schematic of 11 L in Üsküdar Direction.

Buses in 11 L route start their trips in Esatpaşa to Üsküdar direction. After reaching the Üsküdar they turn back to Bulgurlu direction and finish their trip in Esatpaşa. They make a ring trip which means they don’t stop and wait in Üsküdar stations.

Since different routes are selected for each direction (Üsküdar and Bulgurlu), in some parts of the routes buses pass through different locations. Thus, the number and the names of bus stops are different for each direction, illustrated in Table 3.1.

Table 3.1. Bus Stops of 11 L Route in Both Directions.

BULGURLU DIRECTION			ÜSKÜDAR DIRECTION		
No	STOP ID	STOP NAME	No	STOP ID	STOP NAME
1	A0001B	ÜSKÜDAR CAMİ ÖNÜ	1	A0475A	ESATPAŞA
2	A0279A	ÜSKÜDAR MARMARAY	2	A0525C	DOĞAN SOKAK
3	A0280A	HORHOR	3	A2484B	ÜÇYOL
4	A0281A	BÜLBÜL DERESİ	4	A0424B	AÇAN SOKAK
5	A0282A	SETBAŞI	5	A0423B	TUFAN SOKAK
6	A0283A	FISTIKAĞACI	6	A2483B	DÖRTYOL
7	A0284A	KURUÇEŞME	7	A0422B	ALTINKÖY
8	A1158A	KÜLTÜR MERKEZİ	8	A0421B	DOSTLUK PARKI
9	A0285A	BAĞLARBAŞI	9	A0420B	BAĞLARİÇİ
10	A0286B	CAPİTOL	10	A0294B	ALVARLIZADE CAMİİ
11	A0287B	ALTUNİZADE	11	A0413C	FERAH CADDESİ
12	A0288B	MİLLET BAHÇESİ	12	A1841A	TURİSTİK ÇAMLICA TES
13	A0291A	KISIKLI	13	A0291B	KISIKLI
14	A0292A	ÇAMLICA İÖOKULU	14	A0288A	MİLLET BAHÇESİ
15	A0293A	BULGURLU	15	A3685A	METROBÜS ALTUNİZADE
16	A0419A	GAZİLER	16	A0287A	ALTUNİZADE
17	A1450A	BULGURLU CADDESİ	17	A0284B	KURUÇEŞME
18	A0420A	BAĞLARİÇİ	18	A0283B	FISTIKAĞACI
19	A0421A	DOSTLUK PARKI	19	A0282B	SETBAŞI
20	A0422A	ALTINKÖY	20	A0281B	BÜLBÜL DERESİ
21	A2483A	DÖRTYOL	21	A0280B	HORHOR
22	A0423A	TUFAN SOKAK	22	A0279A	ÜSKÜDAR CAMİ ÖNÜ
23	A0424A	AÇAN SOKAK	23	A0001B	ÜSKÜDAR MARMARAY
24	A2484A	ÜÇYOL			
25	A0525D	DOĞAN SOKAK			
26	A0475A	ESATPAŞA			

Because the number and the name of the bus stops are not the same, for each bus stop the closest bus stop in the opposite direction is assigned to eliminate the direction errors.

Kültür Merkezi bus stop is not included in the bus stop list for Üsküdar direction. However, it is known that Kültür Merkezi bus stop is located in the both direction of 11 L and the 11 L buses stops at this stop in the trips to Üsküdar direction. This stop is not identified in İETT's system for 11 L to Üsküdar direction. As seen in Figure 3.3, the route of 11 L to Üsküdar direction is shown at the website of agency as green line. However 11 L buses follow the route shown in blue when they go to Üsküdar from Bulgurlu.

For this reason in this study inference of Kültür Merkezi bus stop as alighting location in Üsküdar direction is taken as a valid inference. However, since there is no boarding records for Kültür Merkezi in ADC data to Üsküdar direction, boardings made at this stop couldn't be inferred.

There were also difference between the recorded stops in ADC data and stated stops for 11 L route in the website transportation agency. In practice, buses stops at 2 bus stops in Üsküdar mainly Üsküdar Cami Önü and Üsküdar Marmaray stops. However, in the ADC dataset there were some records for a stop name as "Üsküdar". Even though the names are different, these stops have the same stop ID, namely Üsküdar Marmaray stop. Therefore, the boardings with these two stops are summed and taken as the boardings at the Üsküdar Marmaray bus stop. In the inference of alighting location also, Üsküdar Marmaray bus stop is used.

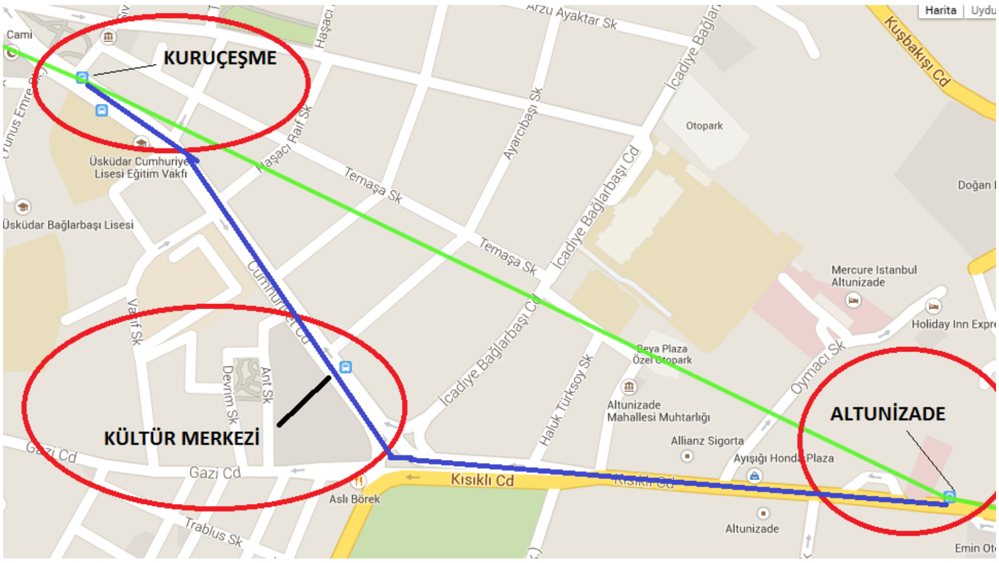


Figure 3.3. Location of Kültür Merkezi Bus Stop and Its Corrected Route.

Even though it is not a very long route, 11 L intersects with several important and different transportation systems as follows:

Metrobüs; which is the longest and commonly used transportation system in İstanbul (Figure 3.4). It has 52 km long distance from Söğtülüçeşme to Beylikdüzü, which are the terminal stops. Over 700,000 passengers are using Metrobüs in weekdays. Intersection point of 11 L and Metrobüs is at Altunizade (IBB, 2014).

Marmaray; which is a subway and it runs across the Bousphorus in a very short period of time (Figure 3.5). It connects with Yenikapı-Hacıosman, Aksaray-Kirazlı and Aksaray-Atatürk Airport Metro Lines in European part of İstanbul. It also reaches the Kadıköy-Kartal Metro in its terminal station (Ayrılıkçeşme) in Anatolian part of İstanbul. Terminal station of 11 L in Üsküdar direction is the interchange station to Marmaray from 11 L.

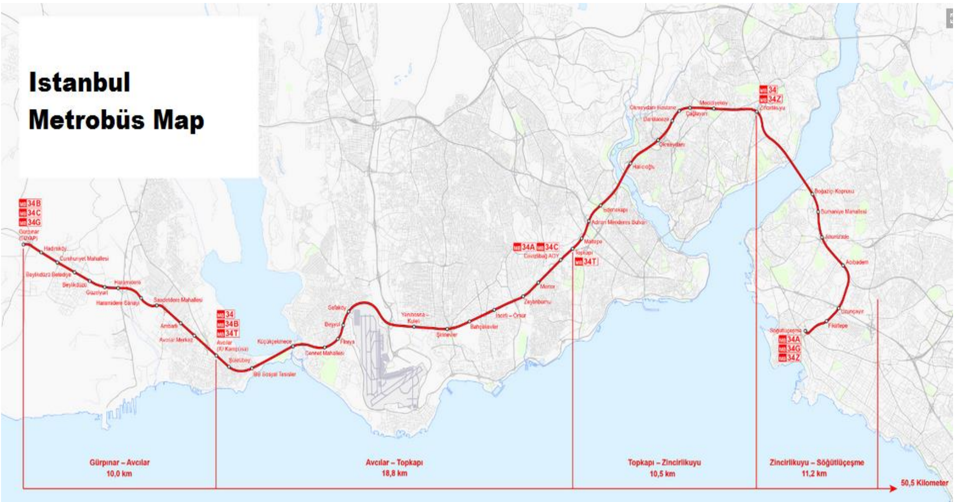


Figure 3.4. Istanbul BRT Line Map 2014 (IETT).

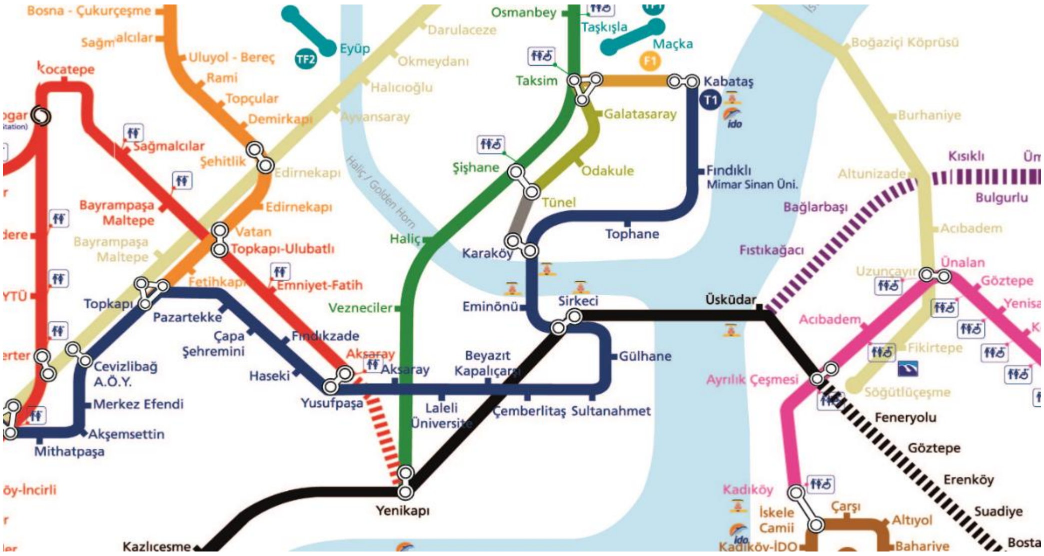


Figure 3.5. Routes of Subways in Istanbul.

Ferries; which runs between Üsküdar and the ports in European part of İstanbul as illustrated in Figure 3.6. Even there are ferries to several ports from Üsküdar, mostly used destinations are Beşiktaş, Karaköy, Kabataş and Eminönü Ports. 11 L reaches the ferry ports in its latest stops.

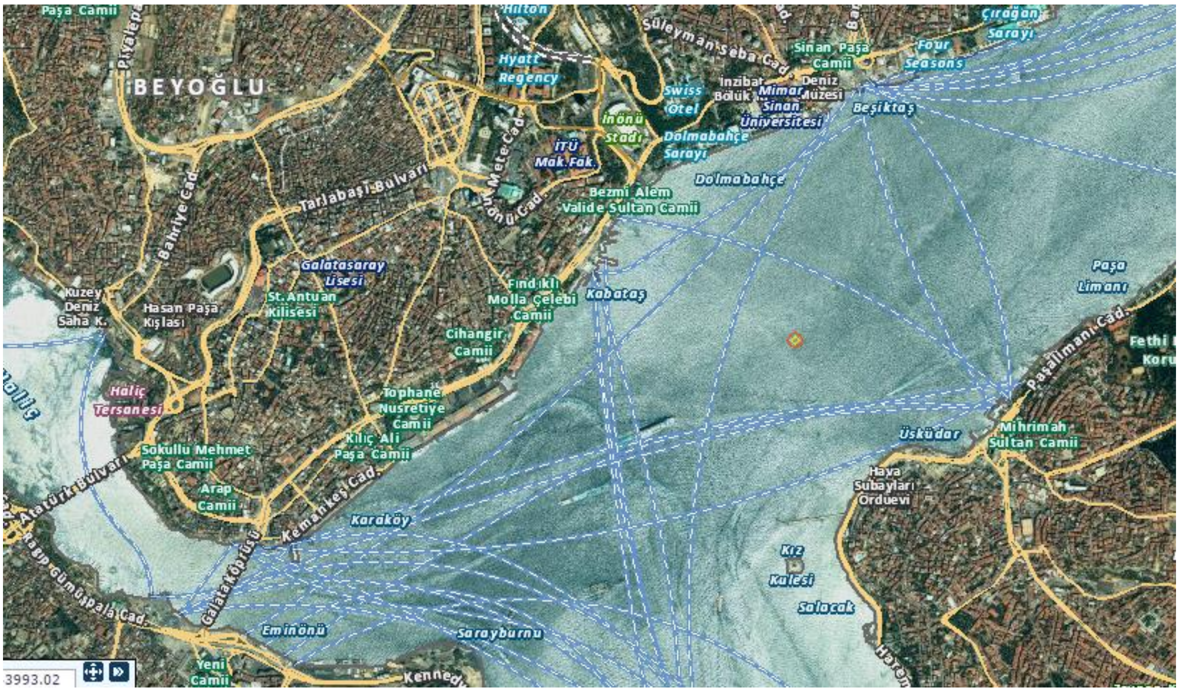


Figure 3.6. Routes of Ferries in Bousphorus.

3.2. ADC Data Analysis

3.2.1. Properties of ADC Data

In this study, ADC data of 11 L for September 15-23, 2014 is used. A sample from data is shown in Table 3.2. ADC data mainly contains;

- Date; which is the date and time in second precision.
- Route: is the name of the line that ADC data records.
- ID; is the unique ID number for each smartcard.
- Type of Ticket; is the type of smartcard given to the passenger according to his/her status in terms of age, education, etc.
- Stop ID; is the unique ID of each bus stops for each direction.
- Gate No; is the unique ID of bus travels in that route. Since in ferries, Metrobüs and subways fares are collected in stations, “Gate No” refers to stations in these transportation systems.
- Name of Stop; is the name of each bus stop. There is no “Stop ID and Name of Stop” information for ferries, Metrobüs and subways because of the explained reason.

Table 3.2. An Example ADC Data.

English	Turkish	Data
Date	Tarih	15.09.2014 06:01:33
Operation Group	Operatorgrubu	Özel Halk Otobüsü
Operator	Operator	Otobüs A.Ş.
Route	Hat	11L
ID	medyaserino	046***************
Type of Ticket	BiletTipi	İndirimli Bilet / Discount Ticket
Type of Fare	GecisTipi	Kontürlü / With Credit
Type of Interchange	AktarmaTipi	Normal / Normal
Stop ID	DurakId	A0424B
Gate No	KapiNo	C-1709
Name of Stop	NoktaAdi	AÇAN SOKAK

Main objective of this study is to find the boarding and alighting locations of the passengers in their 11 L trips. As seen in Table 3.2, ADC data in Istanbul Transportation System unlike many other transportation systems all around world, records the AFC data along with AVL data and contains the boarding locations of the passengers. However, to infer the alighting locations of the passengers, all the information about the consecutive trips is also needed. ADC data for all other trips of 11 L passengers are also studied.

ADC data of 11 L for September 15-23, 2014 have 13.530 records while ADC data which contains the information about 11 L and other trips have 69.195 records for that period. Both datasets are used for this study.

3.2.2. Ticket Types

As explained above, there are different types of cards used in Istanbul Transportation Systems. These cards are given to the passengers by the Transportation Agency (İETT) according to different status of the passengers. For example, passengers whose ages are over 65 are given smartcards to use public transportation for free. Some public officials also have different type smartcards. This classification enable transit planners to analyze the smartcard records according to the type of cardholders.

In Table 3.3, the number of different ticket types in 11 L trips and all trips are shown. As seen in Table 3.3, in 11 L trips dataset, there are 4 trips using missing type of tickets. All trips dataset also has these missing data because all trips dataset contains 11 L trips dataset also.

The percentages of ticket types' usage are shown in Figure 3.7 and Figure 3.8 for 11 L and all trips made in the analysis period. As seen in figures, "Discount Ticket" has the biggest portion in each dataset. This may result from the students of the schools and universities on the route of 11 L.

Table 3.3. Number of Trips by Ticket Types in All Trips and 11 L Trips.

Type of Ticket	BiletTipi	All Trips	11L Trips
Discount Ticket	İndirimli Bilet	31209	5830
Blue Card	Mavi Kart	12771	1964
Full Rate Ticket	Tam Bilet	12751	3570
Over 65 age	65 Yaş Üstü Ücretsiz	4809	753
Handicapped	Özürlü	2555	318
Teacher	Öğretmen	1585	321
Elder	Yaşlı	1522	393
Police	EHS Polis	1376	288
Yellow Press	Sarı Basın	210	24
Companion of Disabled	Özürlü Refakatçi	131	21
War Veteran Wife	Gazi Eşi	81	8
War Veteran	Gazi	79	12
Martyr Family	Şehit Ailesi	39	1
Disabled	Malul	18	5
PTT	PTT Görevli	13	3
Martyr Wifes	Şehit Eşleri	13	1
Military Police	EHS Jandarma	12	3
Travel Expenses	Harcırah	10	9
Retired Personnel	Emekli Personel	5	1
NULL	NULL	4	4
Scholarship Student	Burslu Öğrenci	2	1
	Total	69195	13530

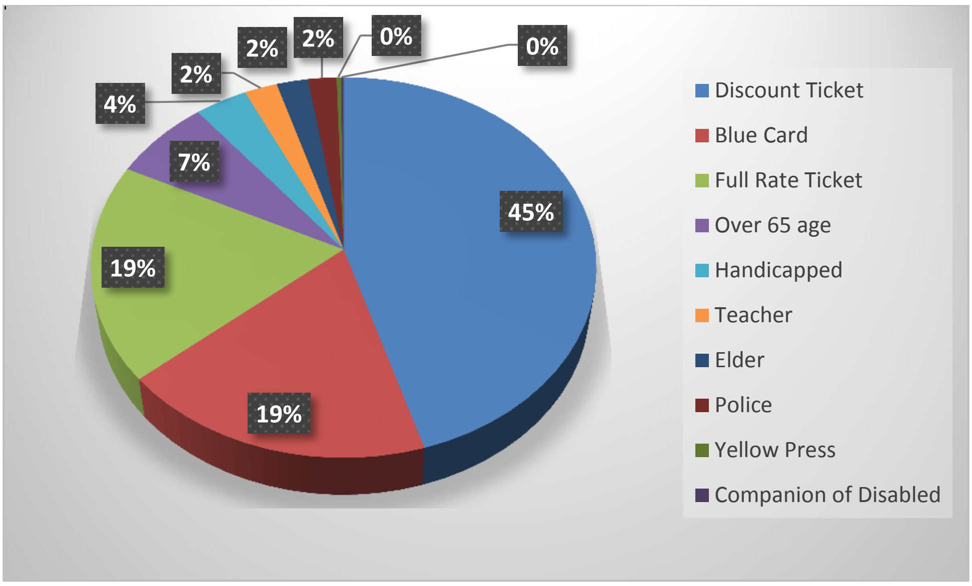


Figure 3.7. Percentages of Different Ticket Types in All Trips.

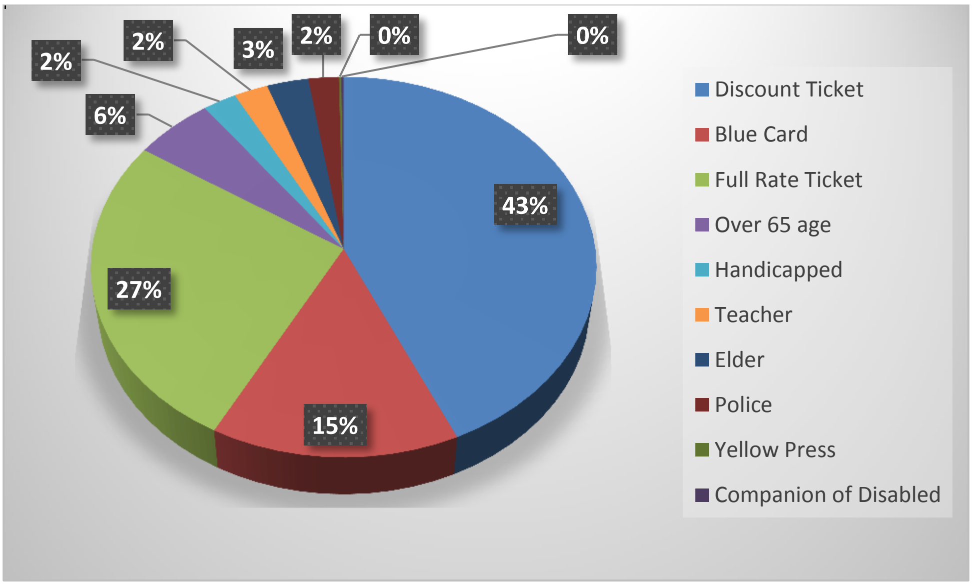


Figure 3.8. Percentages of Different Ticket Types in 11 L Trips.

These cardholders use the transportation systems frequently in different times of the day. One example is passengers over the age of 65. Unlike discount ticket users elder people tend to make their daily trips between morning and evening rush hours when the traffic congestion is relatively less.

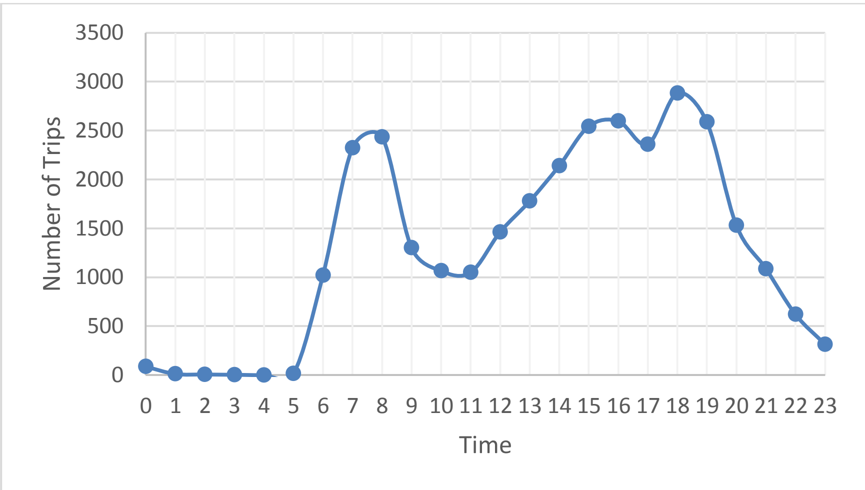


Figure 3.9. Number of Trips of the Passengers Who Use Discount Ticket vs. Time.

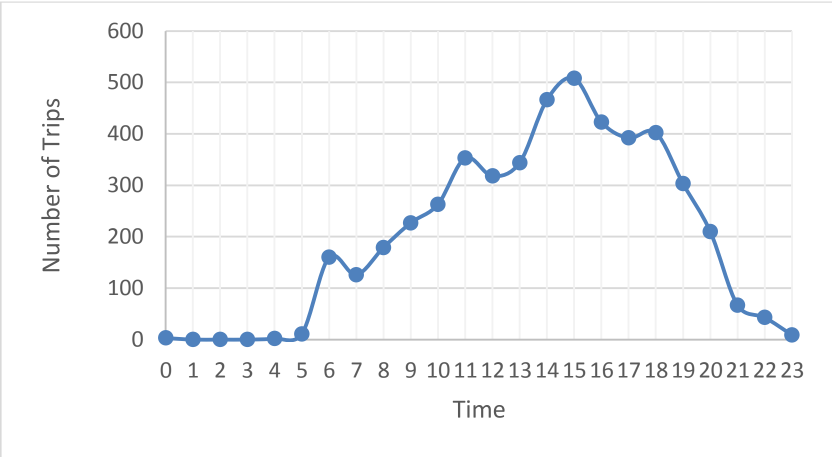


Figure 3.10. Number of Trips of the Passengers over 65 Age vs. Time.

3.2.3. The Difference of Ridership on Weekdays and Weekends

Ridership of 11 L in weekdays and weekend are different as expected. Especially in morning peak hours of weekends, ridership dramatically decreases. Figure 3.11 demonstrates the ridership of 11 L in different days and its distribution on the day.

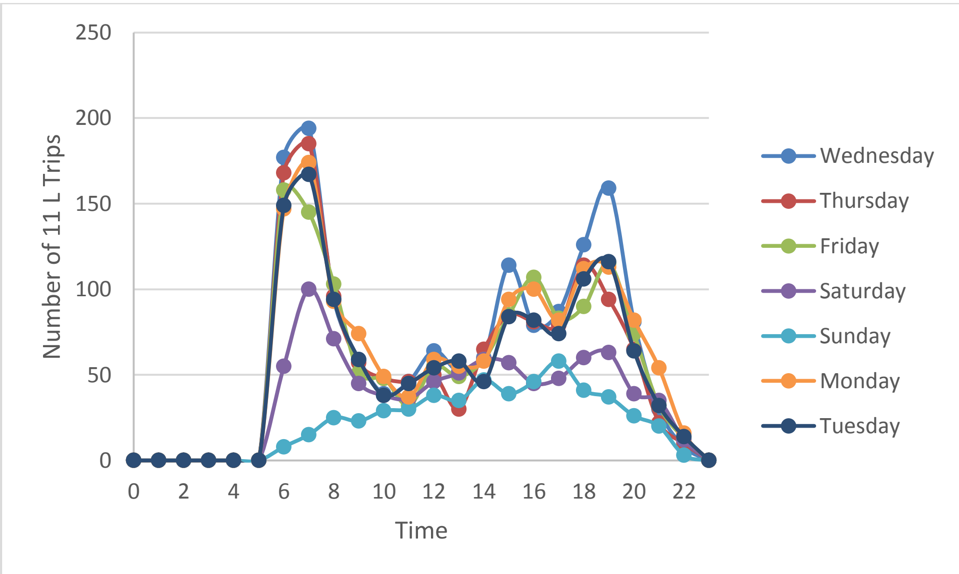


Figure 3.11. Number of 11 L Trips on Each Day vs Time.

3.3. Inference of Alighting Locations

To start the analysis, first very next trip of every single 11 L trips is extracted from "All Trips" dataset. The information about the date, stop name, stop ID, route name and gate no are taken from the dataset and named as "Next Trip" of the relevant 11 L trips.

After extracting all next trips, 11 L ADC dataset is clustered to perform the analysis with different assumptions.

Table 3.4. Type of 11 L Trips and Their Percentages.

Type of 11 L Trips	Number of Trips	%
Interchange in that day	6849	51%
Last trip of the day	5067	37%
Single trip in that day	1533	11%
Single trip in all days	81	1%
Total	13530	100%

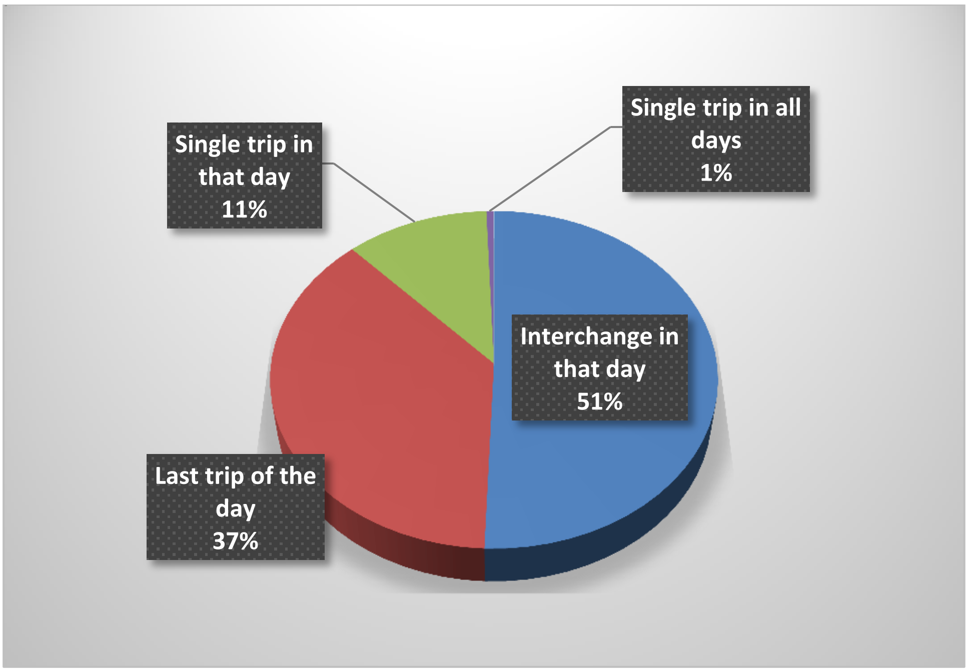


Figure 3.12. Type of 11 L Trips and Their Percentages.

4. ALIGHTING INFERENCE FOR INTERCHANGES

In this cluster, 11 L trips which have recorded next trip on that day as an interchange are analyzed. To study the interchanges in detail, they are classified into three groups;

Table 4.1. Type of Interchanges and Number of Trips in Each Group.

Type of Interchange	Number of Trips
Interchange in 2 hours	3943
Interchange after 2 hours	2218
Interchange to the same bus	688
Total	6849

- Interchanges in 2 hours; are the interchanges made within 2 hours after the studied 11 L trip.
- Interchanges after 2 hours; are the interchanges made after 2 hours.
- Interchanges to the same bus; are the interchanges to the same “Gate Number” which means the same bus and same route after 11 L trip.

4.1. Interchange in 2 Hours

In the interchange study, 2 hours is taken as the threshold because the Transportation Agency (İETT) in İstanbul defined the interchange as the following 5 trips made within 2 hours after the first trip. Passengers are charged less in their interchange trips and the trips made after 2 hours are not accepted as interchange trip. In Table 4.2, fares for discount tickets and full-rate tickets according to the interchange is shown.

Table 4.2. Fare of Interchanges in 2 Hours. (İETT, 2014).

Smartcard Fares	Full-Rate Ticket	Discount Ticket
First Trip	2.15 TL	1.10 TL
1st Interchange	1.45 TL	0.45 TL
2nd Interchange	1.15 TL	0.40 TL
3rd Interchange	0.85 TL	0.40 TL
4th Interchange	0.85 TL	0.40 TL
5th Interchange	0.85 TL	0.40 TL

Also as seen in Figure 4.1, most of the interchanges after 11 L trips are made within the first 2 hours. Also from Figure 4.1, it is understood that passengers make their return trips mostly after 7-8 hours from their 11 L trips.

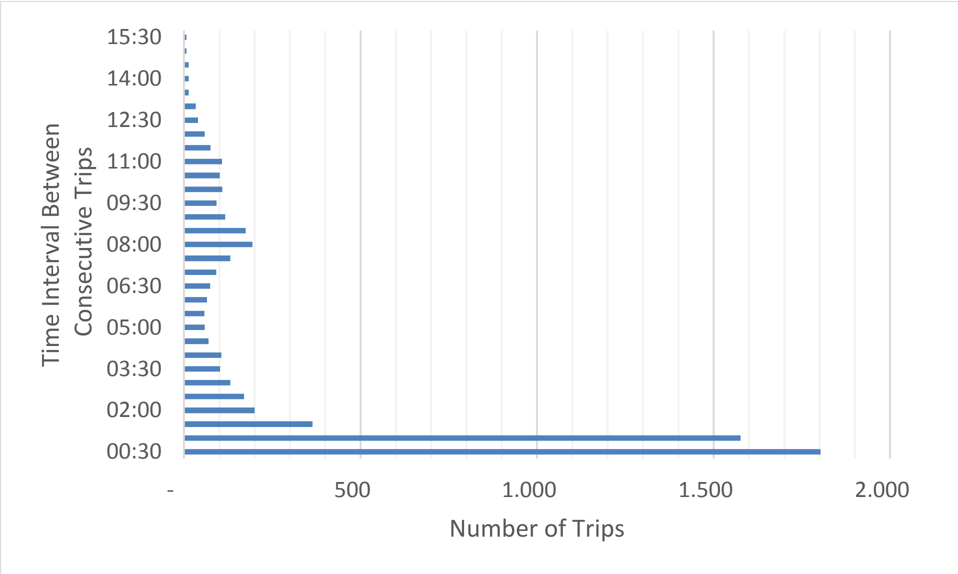


Figure 4.1. Time Interval between 11 L Trip and Next Trip.

Because of the 2 hours limit in interchanges, passengers in İstanbul tend to make their interchange trips in 2 hours. For this reason, interchanges within and after 2 hours were studied separately.

In dataset of interchange in 2 hours, there are some missing boarding stops both in 11 L trips and next trips after 11 L trips. 257 of 3943 11 L trips have no boarding stop information while 73 of these 11 L trips' next trips have no boarding stop data. Since there are no stop information in both 11 L trip and the next trip of the same passengers for 13 trips, total 317 of 3943 11 L trips have missing boarding location in ADC data whether in 11 L trips or next trips.

11 L trips which have no boarding information in their next trips are extracted from data. Because with no information about the interchange location it is not possible to assign an alighting location for the previous 11 L trips.

4.1.1. Bus to Metro Interchanges

In this subset, all the transportation systems which collect the fares at the stops like Marmaray, Metrobüs, Ferries and other subways, are taken as “Metro”, and the subset is named as “Interchange to Metro”. 62% of (2453 of 3943) interchanges made within 2 hours are the interchanges to Metro.

11 L route intersects with Metrobüs in Altunizade. However, the name of stops are different for each direction. To Üsküdar direction Metrobüs, interchange stop is “Metrobüs Altunizade”, while to Bulgurlu direction it is “Altunizade” bus stop. As a result, for the 11 L trips having the interchanges to Metrobüs these 2 bus stop are assigned as alighting location according to the direction.

Since 11 L route reaches to ferries and Marmaray in its terminal stations, “Üsküdar Cami Önü” and “Üsküdar Marmaray” respectively, these 2 bus stops are considered as alighting stops for 11 L trips followed by ferries or Marmaray.

These two stops are very close to each other. Passengers usually alight in the previous bus stop, namely Üsküdar Cami Önü” bus stop when there is traffic congestion. Moreover, in practice passengers tend to alight even in “Horhor” bus stop, the second stop before the terminal station to Üsküdar direction, mostly in rush hours because of the heavy traffic congestion.

As it is mentioned, intersection with ferries and Marmaray is the terminal stations of 11 L in Üsküdar direction. In practice, for the trips of 11 L to Bulgurlu directions it is impossible to have an interchange in Üsküdar because these stops are the first stops of 11 L route in Bulgurlu direction. Therefore, 11 L trips which have the direction to Bulgurlu and next trip of Marmaray or ferries in 2 hours are studied in detail for further correction purposes.

In Table 4.3, there are examples of above mentioned problem. All of these passengers started their 11 L trips in “Doğan Sokak“ and make an interchange in approximately 1 hour to ferries or Marmaray. Trips with USE-UDR, USE-USA and USE-

BSB gate numbers are ferries while the USE-MR1 refers to the Marmaray trip. “Doğan Sokak” is a bus stop very close to the terminal station of 11 L in Bulgurlu direction. However, all these 11 L trips have direction number “1” which is the Bulgurlu direction. Therefore, these ADC data are clearly inaccurate. Also from their Gate No it is understood that they are the ADC data of the same bus. It is known that sometimes AVL system of the buses doesn’t work accurately. This might be the reason behind these inaccurate records. The most adequate way to correct this is to change their direction to the opposite direction. However, in each direction of 11 L there some different stops. When the direction is changed there is a possibility that a bus stop might be assigned to a direction which it doesn’t belong to. To eliminate this problem, for every bus stop of 11 L, closest bus stop in the opposite direction is defined. While correcting the direction, bus stop is also changed with its assigned stop in the opposite direction.

Table 4.3. Examples of Inaccurate Direction Record in ADC Data.

Time	Route	Direction	Stop ID	Stop Name	Gate No	Next Trip Time	Next Trip Name	Next Trip Gate	Time Difference
07:24:08	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:35:42	TY-İST	USE-UDR	01:11:34
07:26:01	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:32:30	DT-İST	USE-USA	01:06:29
07:27:01	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:29:09	TY-İST	USE-UDR	01:02:08
07:27:09	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:33:11	DT-İST	USE-USA	01:06:02
07:27:11	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:28:09	ŞH-İST	USE-BSB	01:00:58
07:28:15	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:27:15	BC1	USE-MR1	00:59:00
07:30:56	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:29:38	DT-İST	USE-USA	00:58:42
07:33:13	11L	1	A0525D	DOĞAN SOKAK	C-1709	08:27:48	BC1	USE-MR1	00:54:35

Above mentioned procedure is also followed for the interchanges to Metrobüs. If the interchange station is not on the way of the direction taken from the ADC data then it is changed to its opposite direction.

In Interchange to Metro dataset, 8 % of11 L trips (187 of 2453) have no recorded boarding location. On the other hand, alighting locations for these 11 L trips were assigned from the interchanges after 11 L trips.

In Table 4.4, results of interchange to metro analysis are summarized. There are 9 direction errors, 187 missing stop ID in 11 L trips, 1 interchange with the boarding location same with the boarding location of the studied trip and 5 undefined interchanges. Alighting locations are inferred using the remaining part of 11 L trips.

Table 4.4. Inference Results of Interchanges to Metro within 2 Hours.

Interchange to Metro in 2 Hours		
Direction Error	9	0.4%
Missing Stop ID in 11 L Trip	187	7.6%
Same Stop	1	0.0%
Interchange Stop Not Close to Metro	5	0.2%
Inference of Alighting	2251	91.8%
Total	2453	

4.1.2. Bus to Bus Interchanges

11 L trips having a next trip in 11 L route or in other routes in 2 hours are studied in this subset. Below procedure is followed:

- If the passenger makes his next trip in 11 L or other routes which share the same “Stop ID”s with 11 L routes, then the boarding location of the next trip is taken as the alighting location of the studied 11 L trip.
- If the passenger makes his next trip in other routes which don’t have common Stop ID’s with 11 L route, then the closeness of the stop to the stops in the route of 11L is checked. If the boarding location of next trip is close enough to any of the stops of 11 L then the closest stop of 11 L to that stop is determined and taken as alighting location of the 11 L trip. In this analysis, 1 km threshold is taken as the maximum walking distance and the stops which have less than 1 km distance to any bus stop in 11 L route are named as “Close” while others are marked as “Not close”.

In the analysis of transfers to buses from 11 L route in 2 hours, some trips made in the European part of the İstanbul within a very short period of time are detected. With further studies it is understood that there are some missing trips between the 11 L trips and the recorded next trips. As it seen in Table 4.5, some passengers start their next trip from places close to the destination of ferries or the stations of the Marmaray.

Table 4.5. Examples of Missing Transfer Trips.

Time	Route	Stop ID	Stop Name	Next trip time	Next Trip Name	Next Trip Boarding Name	Close?	Closest Stop
13:32:31	11L	A0288A	MİLLET BAHÇESİ	14:16:29	30M	BEŞİKTAŞ İSKELE	CLOSE	BEŞİKTAŞ İSKELE
13:40:36	11L	A0423B	TUFAN SOKAK	15:12:06	99A	EMİNÖNÜ	CLOSE	EMİNÖNÜ İSKELE
14:12:30	11L	A0287A	ALTUNİZADE	14:57:10	32T	EMİNÖNÜ	CLOSE	EMİNÖNÜ İSKELE
14:14:57	11L	A0420B	BAĞLARİÇİ	15:13:31	EM1	EMİNÖNÜ	CLOSE	EMİNÖNÜ İSKELE
14:30:52	11L	A0283B	FISTIKAĞACI	16:10:32	40B	BEŞİKTAŞ İSKELE	CLOSE	BEŞİKTAŞ İSKELE
08:12:39	11L	A0283B	FISTIKAĞACI	09:18:51	31Y	YENİKAPI İSTASYON	CLOSE	YENİKAPI MARMARAY
10:46:56	11L	A0294B	ALVARLIZADE CAMİİ	12:39:49	99A	EMİNÖNÜ	CLOSE	EMİNÖNÜ İSKELE
08:01:09	11L	A0424B	AÇAN SOKAK	09:01:35	TURİSTİK HAT	YENİKAPI İSTASYON	CLOSE	YENİKAPI MARMARAY

It is very clear that these passengers have missing records in ADC dataset. This may be due to the inaccurate ADC recordings or these passengers didn't use their smartcard in the transfer trips. Most adequate explanation to this with the consideration of smartcard usage in practice is that, when these passengers reach the station of Marmaray or ferries, they realized that in their smartcards there is no credits and there is no time to load credit to their smartcards before the ferry or Marmaray leave the station. Therefore, they asked to their friends or other passengers to use their smartcards. After they reached to destination of ferry or Marmaray, they load credits to their cards and continued their trips.

The main ferry trips from Üsküdar are to Eminönü, Beşiktaş, Karaköy and Kabataş while Marmaray goes through Sirkeci, Yenikapı and Kazlıçeşme after Üsküdar station. Hence the stops in Table 4.6 are also considered as stops of 11 L and matched with Üsküdar stations of 11 L. If the passengers start their next trip in 1 km distance to any of these stops, alighting location of previous 11 L trip is assigned accordingly. This correction made only in the interchanges within 2 hours. In the analysis of other clusters these destinations aren't included in the list of close stops to 11 L route.

Table 4.6. Destination of Ferries and Marmaray & Assigned Bus Stops on 11 L route.

Destinations Of Ferry Or Marmaray	11 L Stop
Kabataş İskele	Üsküdar Cami Önü
Karaköy İskele	Üsküdar Cami Önü
Beşiktaş İskele	Üsküdar Cami Önü
Eminönü İskele	Üsküdar Cami Önü
Sirkeci Marmaray	Üsküdar Marmaray
Yenikapi Marmaray	Üsküdar Marmaray
Kazliçeşme Marmaray	Üsküdar Marmaray

Because of the long route of Metrobüs, very detailed study should be made for Metrobüs for defining the acceptable bus stops and different transfer times for each station of Metrobüs. Therefore, Metrobüs route is not analyzed in this thesis. It can be studied in further research in this subject.

With the same procedure of “Bus to Metro” subset, boarding and assigned alighting locations of the 11 L trips are further analyzed in “Bus to Bus” subset also. During the analysis, some inaccurate records were detected. One example is shown in Table 4.7. From the ADC data of bus C-1722, Bulgurlu Caddesi is recorded as boarding locations of every transactions for approximately 50 minutes. This is obviously not possible in practice. Therefore, from the AVL data of bus C-1722 GPS coordinates of the bus are taken and closest stop in the route of 11 L is assigned to the AVL records. Assigned stops from AVL data for each ADC data seems much more accurate and thereby the continuity of the bus in its route is achieved. These correct AVL records are switched with the inaccurate boarding locations.

Since the AVL data available didn’t contain all 11 L trips, these corrections couldn’t be made for every inaccurate ADC data.

With the same procedure followed in “interchange to metro” cluster, the inaccurate direction records in ADC data is corrected by changing to the opposite direction.

Table 4.7. Inaccurate Record of Boarding in AFC Data and Assigned Boarding Location from AVL Data.

TIME	ROUTE	GATE NO	DIRECTION	STOP NAME	DIRECTION from AVL	STOP INFO from AVL
14:24:47	11L	C-1722	1	BULGURLU CADDESİ	2	ÜÇYOL
14:24:52	11L	C-1722	1	BULGURLU CADDESİ	2	ÜÇYOL
14:25:51	11L	C-1722	1	BULGURLU CADDESİ	2	AÇAN SOKAK
14:25:54	11L	C-1722	1	BULGURLU CADDESİ	2	AÇAN SOKAK
14:25:56	11L	C-1722	1	BULGURLU CADDESİ	2	AÇAN SOKAK
14:28:24	11L	C-1722	1	BULGURLU CADDESİ	2	DOSTLUK PARKI
14:30:29	11L	C-1722	1	BULGURLU CADDESİ	2	DOSTLUK PARKI
14:30:37	11L	C-1722	1	BULGURLU CADDESİ	2	BAĞLARİÇİ
14:32:58	11L	C-1722	1	BULGURLU CADDESİ	2	ALVARLIZADE CAMİİ
14:33:27	11L	C-1722	1	BULGURLU CADDESİ	2	ALVARLIZADE CAMİİ
14:52:39	11L	C-1722	1	BULGURLU CADDESİ	2	ALTUNİZADE
14:54:39	11L	C-1722	1	BULGURLU CADDESİ	2	KÜLTÜR MERKEZİ
14:54:43	11L	C-1722	1	BULGURLU CADDESİ	2	KÜLTÜR MERKEZİ
14:54:47	11L	C-1722	1	BULGURLU CADDESİ	2	KÜLTÜR MERKEZİ
14:54:51	11L	C-1722	1	BULGURLU CADDESİ	2	KÜLTÜR MERKEZİ
14:55:02	11L	C-1722	1	BULGURLU CADDESİ	2	KURUÇEŞME
15:13:45	11L	C-1722	1	BULGURLU CADDESİ	1	HORHOR
15:14:11	11L	C-1722	1	BULGURLU CADDESİ	1	HORHOR
15:14:15	11L	C-1722	1	BULGURLU CADDESİ	1	HORHOR

As it is seen in Table 4.8, for approximately 86 percent of the 11 L trips alighting locations are inferred in the interchanges to 11 L routes or the routes which have common bus stops with 11 L route.

Even though there is a procedure to eliminate the direction errors, some of the errors couldn’t be corrected. This mainly results from the terminal station error. For some 11 L trips alighting location is inferred as terminal station of the route and it is not logical to change the direction because of the fact that the boarding location of the 11 L trip is very close to the last station of 11 L in the opposite direction.

Table 4.8. Inference Results of Interchanges to 11 L Route or Other Routes with Common Stops within 2 Hours.

Interchange To Bus (11 L Or Routes With Common Stops) In 2 Hours		
Direction Error	83	8.7%
Missing Stop Id In 11 L Trip	40	4.2%
Same Stop	15	1.6%
Interchange Stop Is Not Close	0	0.0%
Inference Of Alighting	816	85.5%
Total	954	

Table 4.9. Inference Results of Interchanges to Other Buses within 2 Hours.

Interchange To Bus (Other Routes) In 2 Hours		
Direction Error	0	0.0%
Missing Stop Id In 11 L Trip	16	3.5%
Same Stop	8	1.7%
Interchange Stop Is Not Close	127	27.5%
Inference Of Alighting	310	67.2%
Total	461	

For the interchanges to the other routes which don't share any bus stop with 11 L route but have bus stops in 1 km distance to the route of 11 L, below results are found. Inference rate is relatively less because in some interchanges the distance between the boarding location of next trip and the closest 11 L stop exceeds the limit. Almost 30 percent of the interchanges are made in locations which are not close enough to the 11 L route. This percentage may decrease if the interchange time is limited for a shorter period of time.

Total number of trips seen in Table 4.8 and Table 4.9 are less than the number of trips in this cluster. This is mainly because 11 L trips which have no boarding information in their next trip are not analyzed in this study.

4.2. Interchange After 2 Hours

In the analysis of this cluster, the assumption made for the missing transfer trips is not taken into account because of the time difference between the 11 L trip and the next trip. It is quite possible for passengers to reach the locations within the time of interchanges in this cluster.

Also no correction is done for the direction of the 11 L trips. In this cluster, since the time difference between 11 L trips and the next trip is long enough, it is not preferable to make a change in 11 L trip data according to the boarding location of next trip. There are many possible ways for a passenger to reach the boarding location of the next trip in that period of time.

In conclusion, for 11 L trips in this cluster, boarding location of the next trip is taken and checked whether it is one of the 11 L route's stops or the close stops to the 11 L route that are determined at the beginning of the study.

The results found after the analysis for the interchanges after 2 hours to 11 L or routes which have common stops with 11 L (Table 4.10) are quite similar to the results of interchanges in 2 hours cluster. On the other hand, for the interchanges to the other routes (Table 4.11), the percentage of the inferred alighting location for 11 L trips is dramatically decreased from 84.6 % to 38.2 %.

Table 4.10. Inference Results of Interchanges to 11 L Route or Other Routes with Common Stops after 2 Hours.

Interchange To Bus (11 L Or Routes With Common Stops) After 2 Hours		
Direction Error	113	7.2%
Missing Stop Id In 11 L Trip	108	6.9%
Same Stop	21	1.3%
Interchange Stop Is Not Close	0	0.0%
Inference Of Alighting	1329	84.6%
Total	1571	

Table 4.11. Inference Results of Interchanges to Other Routes after 2 Hours.

Interchange To Bus (Other Routes) After 2 Hours		
Direction Error	22	3.8%
Missing Stop Id In 11 L Trip	50	8.6%
Same Stop	3	0.5%
Interchange Stop Is Not Close	286	49.0%
Inference Of Alighting	223	38.2%
Total	584	

4.3. Interchange to the Same Bus

As shown before, interchanges are divided into three groups in this study; interchanges in 2 hours, interchanges after 2 hours and interchanges to the same bus. The last cluster is introduced for a particular reason. In public transportation systems of İstanbul, passengers need to use their smartcards to pay the fare of their trips. Most of the passengers who use the public transportation of Istanbul in daily basis have their own smartcards which have unique IDs. In Metrobüs, people who don't have smartcards are asked to buy one-time or multiple time tickets to make the payments. In subways and ferries, passengers can buy a token from machines or pay desks in the stations. However, bus riders in İstanbul need to pay their fees in buses by swiping their smartcards to the machines installed to buses. Machines can read only the smartcards and there is no other payment options serviced to the passengers. For this reason if a passenger boards a bus with his smartcard having not enough credit in, he either need to get off the bus, look for a place to load credit to his smartcard and wait the next bus or ask other passengers to use their smartcards and make the payment in cash to the passenger lending his smartcard.

In the last cluster of interchanges, 11 L trips which have a record of 11 L trip as next trip in the same bus are studied.

In the study 60 minutes threshold is taken to detect the records of above explained situation because estimated trip duration for 11 L route is given as 88 minutes for a round by the transportation agency. This duration can change during different periods of a day. It may increase in morning and evening peak hours while it is quite possible that during the very early morning and late nighttime a round takes less time. Therefore, interchanges to the same bus of 11 L within 60 minutes having the same direction, are analyzed in details.

As seen in Table 4.12, most of the next records are made in a very short period of time. The reason behind this is that passengers having smartcards with no required credit in it, firstly ask to the passengers who board at the same bus stop with them for the usage of their smartcards. Therefore, the time difference of these records are usually very small and the boarding locations are the same.

On the other hand, some passengers use their smartcard for other passengers at the next stations in the route, like the passenger in the 10th row of Table 4.12. That passenger began his or her trip at the first station of 11 L in Üsküdar direction and used his or her smartcard for another passenger after two stops.

There are 507 records of this situation. This is considerably high number since 4 % of 11 L (507 of 13530) records are repeated records.

Since most of the records have the same boarding location, these repeated records aren´t excluded from the dataset. However, in the proposed algorithm these records are suggested to be eliminated from the dataset.

Other interchanges to the same buses are considered as normal trips and processed with the same procedure of interchanges after 2 hours subset.

Inference results of this cluster is shown in Table 4.13. Since this cluster is introduced to detect the repetitive use of the same smartcard, most of the records in this cluster are these repeated records. Even though it is not expected to have high inference rate in this cluster, 22.4 % of 11 L trips are successfully explained. The passengers of these trips made their next trips luckily in the same bus and opposite direction or in the same bus and direction but after at least 1 hour.

Table 4.12. Examples of Repetitive Use of the Same Smartcard.

No	Time	Route	ID	Direction	Gate No	Stop Name	Next Trip Time	Next Trip Name	Direction	Next Trip Gate	Next Trip Boarding	Time Difference
1	06:23:34	11L	042*********	2	C-1715	AÇAN SOKAK	06:23:37	11L	2	C-1715	AÇAN SOKAK	00:00:03
2	06:24:40	11L	046*********	2	C-1715	TUFAN SOKAK	06:26:33	11L	2	C-1715	TUFAN SOKAK	00:01:53
3	06:29:39	11L	043*********	2	C-1715	BAĞLARİÇİ	06:29:43	11L	2	C-1715	BAĞLARİÇİ	00:00:04
4	06:38:02	11L	047*********	2	C-1720	AÇAN SOKAK	06:38:16	11L	2	C-1720	AÇAN SOKAK	00:00:14
5	06:39:14	11L	042*********	2	C-1720	TUFAN SOKAK	06:39:31	11L	2	C-1720	TUFAN SOKAK	00:00:17
6	06:41:47	11L	046*********	2	C-1715	KURUÇEŞME	06:42:00	11L	2	C-1715	KURUÇEŞME	00:00:13
7	06:46:57	11L	043*********	2	C-1722	ÜÇYOL	06:47:01	11L	2	C-1722	ÜÇYOL	00:00:04
8	06:49:55	11L	047*********	2	C-1722	TUFAN SOKAK	06:50:06	11L	2	C-1722	TUFAN SOKAK	00:00:11
9	06:51:25	11L	042*********	2	C-1722	DÖRTYOL	06:51:28	11L	2	C-1722	DÖRTYOL	00:00:03
10	12:59:53	11L	043*********	2	C-1732	ESATPAŞA	13:08:22	11L	2	C-1732	ÜÇYOL	00:08:29
11	13:01:24	11L	047*********	1	C-1722	KURUÇEŞME	13:01:30	11L	1	C-1722	KURUÇEŞME	00:00:06
12	13:01:41	11L	042*********	1	C-1722	KURUÇEŞME	13:01:49	11L	1	C-1722	KURUÇEŞME	00:00:08
13	13:01:49	11L	046*********	1	C-1722	KURUÇEŞME	13:02:11	11L	1	C-1722	KURUÇEŞME	00:00:22

Table 4.13. Inference Results of the Interchanges to the Same Bus.

Interchange To The Same Bus		
Direction Error	11	1.6%
Missing Stop Id In 11 L Trip	15	2.2%
Same Stop	1	0.1%
Repeated Records	507	73.7%
Inference Of Alighting	154	22.4%
Total	688	

5. ALIGHTING INFERENCE FOR LAST TRIPS

In this cluster, 11 L trips which have no next trip in that day are examined. These 11 L trips are the last trips of the day. As it is discussed in the methodology section, passengers are assumed to come back to the places where they make their first trips in that day.

5.1. First Method

It is known that next trips of these 11 L trips are not made in the same day. Therefore, first trips and boarding locations of the passengers are determined. If the first trips are made in 11 L route or the routes which have common stops with 11 L, boarding location is taken directly. On the other hand, if the boarding location is not one of the 11 L bus stops then the distance of the stop to the nearest 11 L stop should be calculated. Since the nearby bus stops to the 11 L is determined at beginning of the analysis, it is an easy process to determine whether the first boarding record of the passengers is in the distance limit or not. First, the boarding stop ID is taken and searched in the list which contains the bus stops in 1 km distance to 11 L route. If it is in the list, the closest 11 L stop for that bus stop is taken from the list and assigned to the passenger.

For 62 % of the last trips (3134 of 5067), the first boarding location is determined successfully. For 80 % of these 11 L trips (2502 of 3134), the alighting locations were successfully inferred while 85 of the 11 L trips have no boarding information (Table 5.1).

For 546 11 L trips, the assigned alighting location resulted in direction errors. However, since the time difference of these trips are relatively high, to change the direction of recorded 11 L to the opposite direction and eliminate these errors is not preferred. This method is used in “Interchange in 2 hours” cluster because the time difference between the trips are relatively short.

Table 5.1. Inference Results of the Last Trips Using the First Method.

Last Trip Of The Day (First Trip 11 L or Routes With Common Stops)		
Direction Error	546	17.4%
Missing Stop Id In 11 L Trip	85	2.7%
Same Stop	1	0.0%
First Stop Of The Day Is Not Close	0	0.0%
Inference Of Alighting	2502	79.8%
Total	3134	

5.2. Second Method

The other 11 L trips for which no alighting location could be assigned from the first trip of the day are extracted from the cluster. These trips and trips in “Single Trip in that day” cluster are combined. Using the first trips of passengers on that day, it is not possible to infer an alighting location of the 11 L trips which are the last trips of the day for these two clusters.

To infer an alighting location for these trips, next trip analysis is performed. If the next boarding of the passenger is in the route of 11 L or close enough to the route of 11 L then it is assumed that boarding location of the next trip is the alighting location of 11 L trip.

However, since the next trip of that passenger is made in the next day or in the upcoming days, there should be a limit for time difference in days. Thus, the limits in Table 5.2 are used. For weekdays except for Friday and for Sunday, the limit is taken as 1 day. For Saturday and Sunday 2 and 3 days thresholds are assigned, respectively. The reason behind this is the public transportation usage habits of the passengers. Passengers tend to use public transportation in weekends less often. Therefore, the number of days between the day of trip and next weekday is compared with the assumed limits. For example, if a passenger has a record of 11 L as his last trip of the day on Friday and his next trip is recorded in the next week on Monday then the continuity of the consecutive trips is assumed to be accomplished.

Table 5.2. Limits for the Time Differences between the Day of 11 L Trip and the Day of the Next trip.

Next Trip Limits	The day of 11 L trip
1 day	Monday
1 day	Tuesday
1 day	Wednesday
1 day	Thursday
3 days	Friday
2 days	Saturday
1 day	Sunday

After the implementation of the described methodology, below results are found. Only about 36 percent of the alighting locations could be inferred. The biggest portion in errors is found in the closeness of the next trips' boarding location.

Table 5.3. Inference Results of the Last Trips from Second Method.

Last Trip Of The Day (No Result From First Trip Of The Day)		
Direction Error	429	12.7%
Missing Stop Id In 11 L Trip	120	3.5%
Same Stop	382	11.3%
Interchange Stop Is Not Close	911	26.9%
Day Of Next Trip Is Not Close	334	9.9%
Inference Of Alighting	1210	35.7%
Total	3386	

5.3. Last Trip and No Trip on the Other Days

There are also some passengers which have records only in one day. These passengers are studied separately with the same procedure of first method in last trip cluster. The results are shown in Table 5.4. As it is seen, the inference rate is very low. This may result from the irregular public transportation use or the ticket type of these passengers o. The transportation agency offers passengers tickets which can be used up to 10 times. These cards are preferred mostly by visitors and tourists in İstanbul. It is clear that these passengers aren't commuters or use the public transportation systems regularly. From the overall results, it is concluded that for the commuters or the passengers who use the public transportation systems in daily basis, it is much easier to infer the alighting locations of the trips.

Table 5.4. Inference Results of Last Trips (No Other Trips on Other Days).

Last Trip of the Day (No Trip on Other Days)		
Direction Error	16	21.9%
Missing Stop ID in 11 L Trip	3	4.1%
Same Stop	8	11.0%
First Stop of the Day IS Not Close	29	39.7%
Inference of Alighting	17	23.3%
Total	73	

6. PROPOSED ALGORITHM

Methodology used in this study is based on the several assumptions described in the previous studies like Cui (2006) and Wang (2010). As it is explained in the methodology section of the thesis; however, some additional assumptions are made in this study to improve the previous methods. All the steps of the methodology are described using some examples in Chapters 4 and 5.

The proposed algorithm (Figure 6.1) for the alighting inference of the studied trips are summarized in the below algorithm schemes. The proposed algorithm is deemed to give accurate results in the inference of alighting locations of the trips with the consideration of key points stated below:

- In the search for closeness of the next boarding in the interchanges made within 2 hours, look up table for the closest stops should be generated with the consideration of possible missing trips.
- Threshold for the maximum walking distance should be determined for the studied routes before the analysis.
- For the steps of the algorithm introduced to detect the repetitive use of the same smartcard, temporal threshold (1 hour in this study) should be determined according to the minimum duration of the route for a single round.

Also it should be noted that since most of the repetitive use of smartcard is seen at the same stop, these repetitive records aren't extracted from the records in the studied ADC dataset. However, in the proposed algorithm it is suggested that at the beginning of the study these records should be detected and discarded.

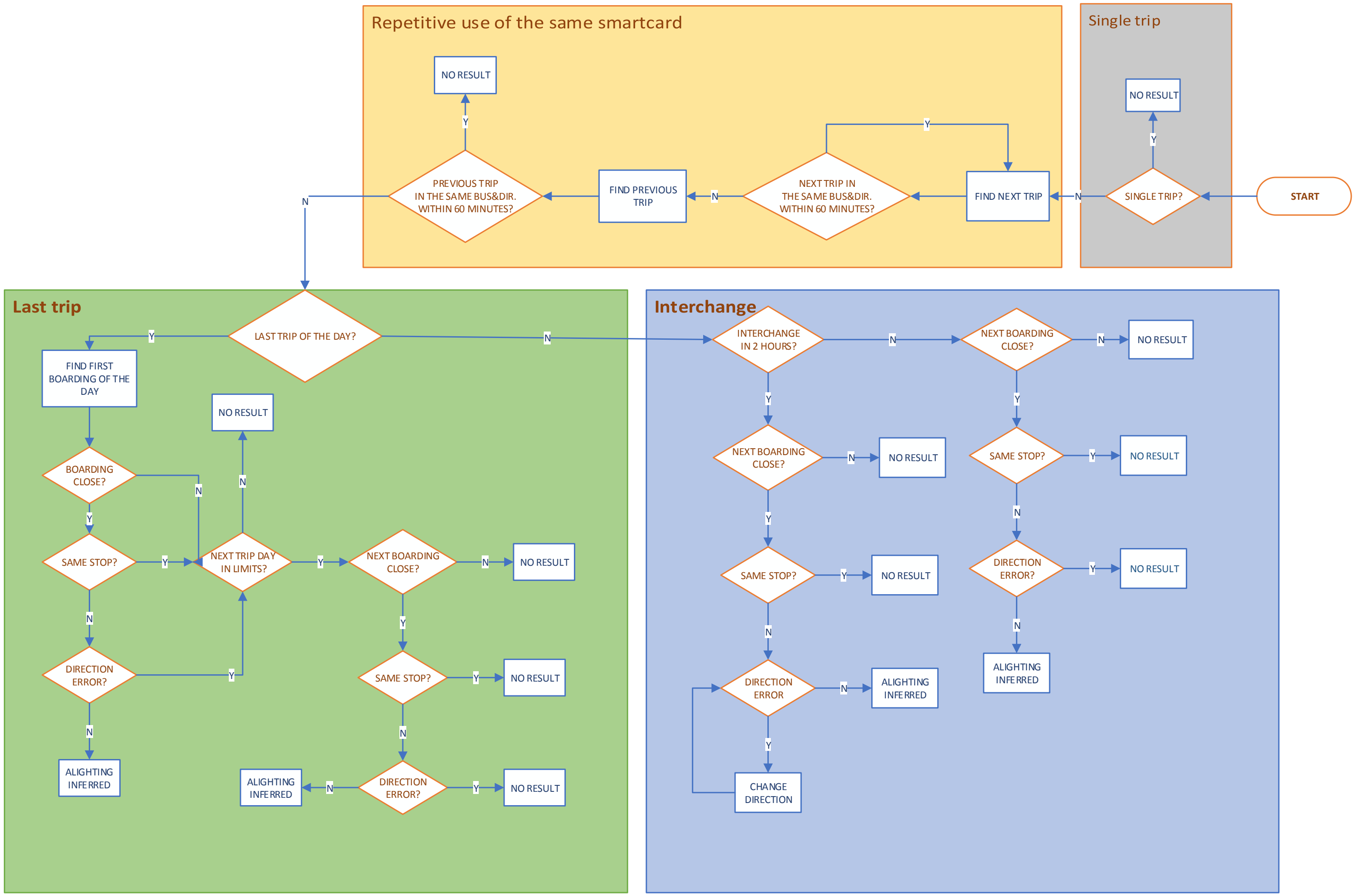


Figure 6.1. The Proposed Algorithm for Inference of Alighting Location.

The proposed algorithm starts with the initial step that checks whether there is any other trip of the passengers being evaluated. If there is not any other trip of those passengers on studied days, then no inference can be made (Figure 6.1).

For repetitive use of the same smartcard, the algorithm checks the first next trip to determine whether it is a repetitive use of the same smartcard (Figure 6.2). If it is so, then the algorithm goes back in the process and finds the next record of that passenger. This goes until finding the real next trip that passenger made after the studied trip. The proposed algorithm also checks in its second step whether the studied trip is actually the repetitive use of the same smartcard. For this, algorithm asks the previous trip of the studied trip with the same conditions. If the previous trip of the studied trip made under the given conditions, then it is concluded that the studied trip is the repetitive use of the same smartcard and no inference is made for the studied trip. This step is introduced to eliminate the misleading alighting inferences for the passengers who used other passengers' smartcard.

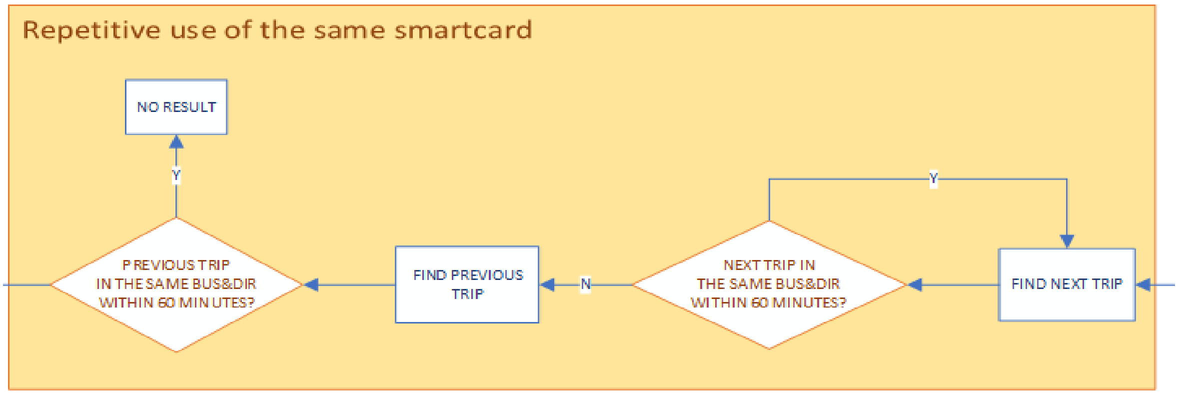


Figure 6.2. The Proposed Algorithm Steps for Determining the Repetitive Use of the Same Smartcard.

After pinning down the repetitive uses of the same smartcard, the algorithm checks whether it is the last trip of the day or there is any trip after that. According to the type of the studied trip algorithm goes whether last trip part or interchange part.

Both in the last trip and interchange inferences, the algorithm first checks the closeness of the inferred alighting location. If it is close enough to the studied route then the algorithm proceeds and checks whether the boarding stop and the inferred alighting stop are the same or not. If they are the same, there is obviously no result. If not, the algorithm further checks for the direction error. The direction error means that inferred alighting location is not on the route of recorded direction.

In the last trip section (Figure 6.3), the proposed algorithm goes to the step where the day of next trip is checked for the trips that no inference could be made from the first boarding of the passenger on that day. After that, the same procedure as described in the previous paragraph is followed.

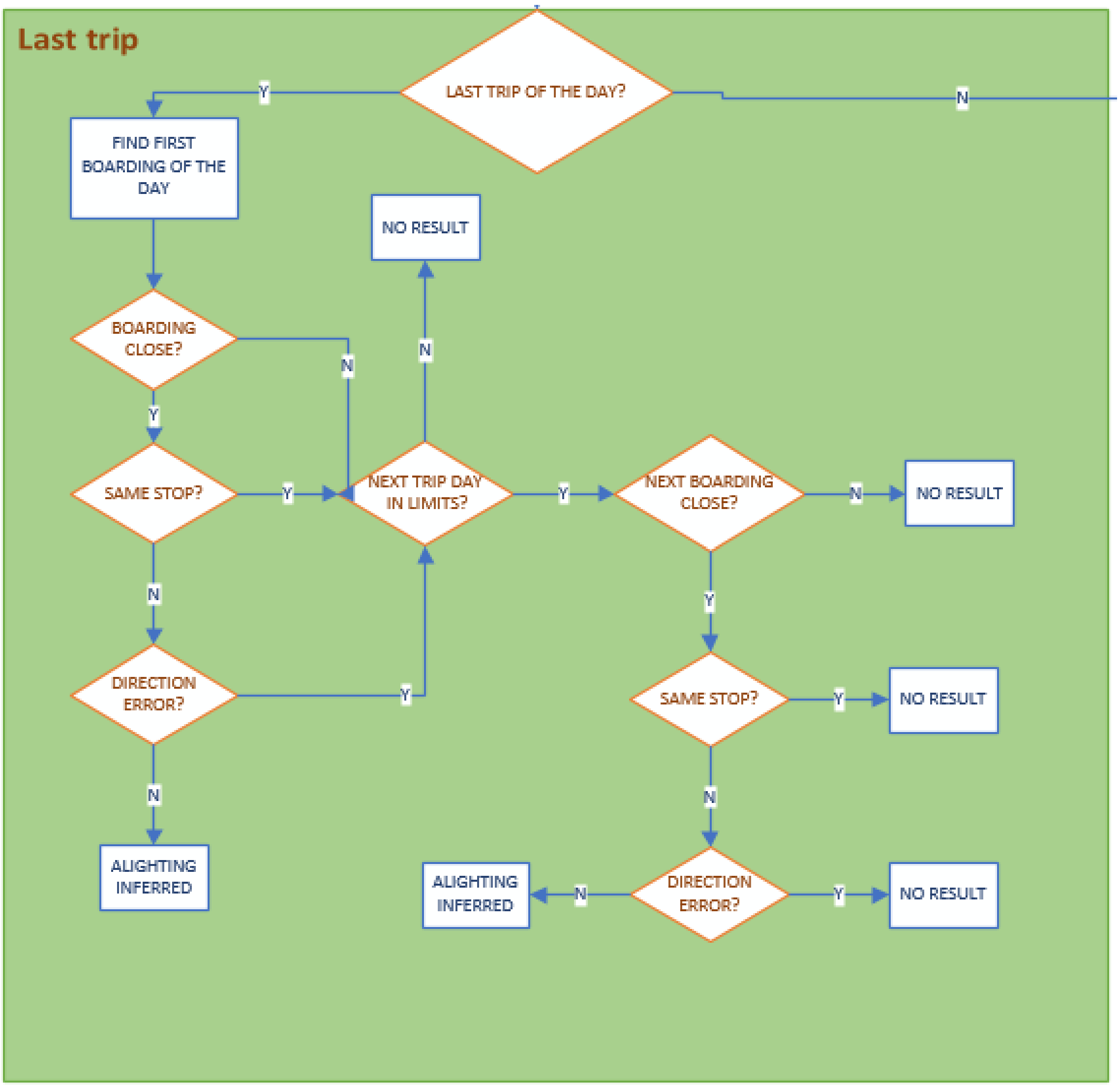


Figure 6.3. The Proposed Algorithm Steps for the Last Trips of the Day.

The interchange portion of the proposed algorithm (Figure 6.4) is different from the previous algorithms in the literature in that it changes the recorded direction of the studied trip if a direction error occurs for the trips which have an interchange in 2 hours.

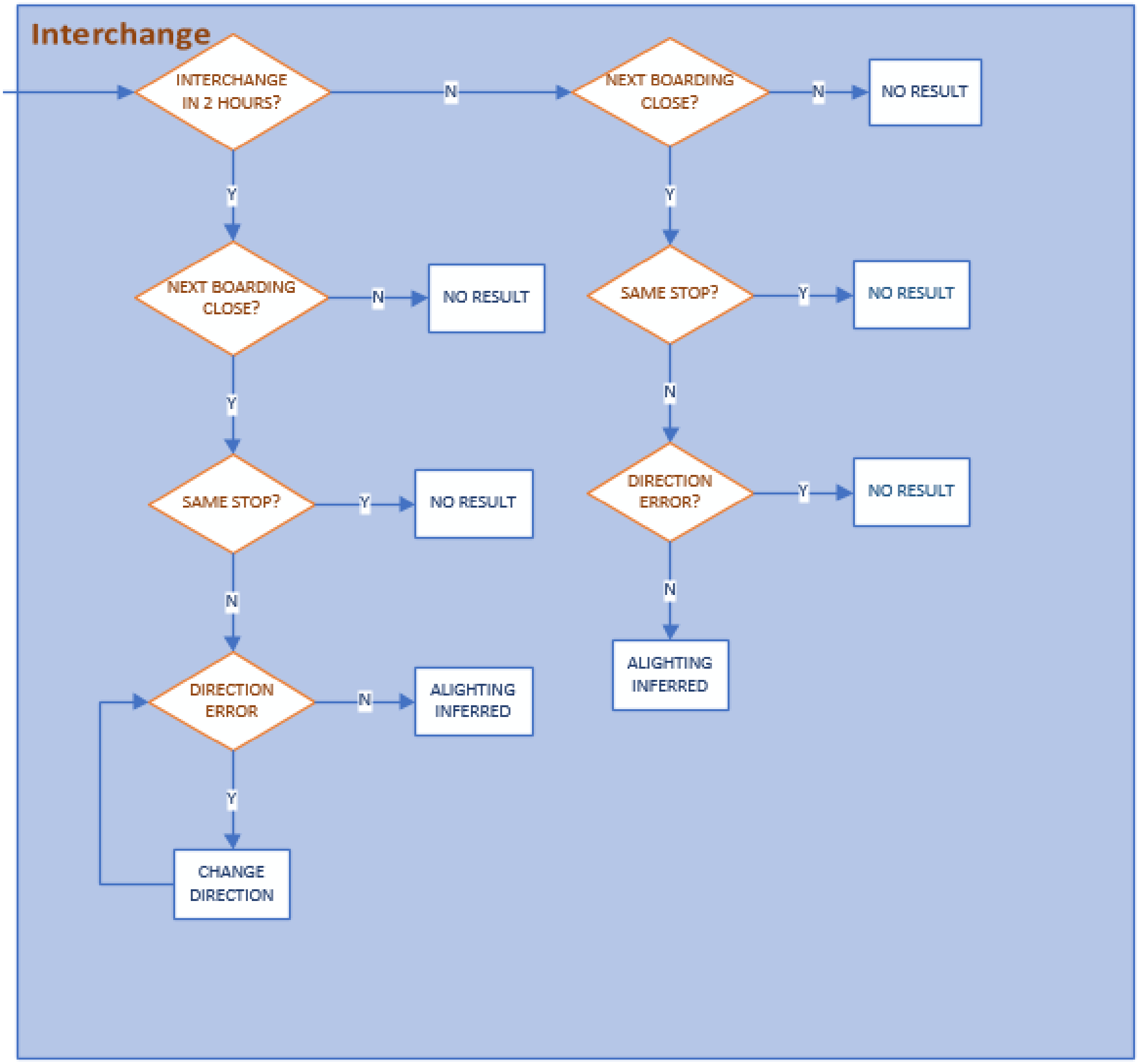


Figure 6.4. The Proposed Algorithm Steps for the Trips Having an Interchange Trip on the Same Day.

7. INFERRED ORIGIN & DESTINATIONS

With the modifications to the previous algorithms proposed in this study, the inference process gives very fulfilling results as shown in Table 7.1. The total number of studied 11 L trips is 13,304. This number reaches to 13,530 with the inclusion of the examined 11 L trips of the passengers that have no other trips in the studied days and the 11 L trips which have no boarding information in their next trip. Total of 8,812 out of 13,304 11 L trips, the alighting locations are inferred. For 65.13% of all 11 L trips ODs are successfully inferred. This rate increases to 66.24% when the 11 L trips, not studied, are excluded.

Table 7.1. Overall Inference Results.

Reason of Error	Total	Percentage
Direction Error	1,229	9.24%
Missing STOP ID In 11 L Trip	624	4.69%
Same Stop	440	3.31%
Interchange Stop Is Not Close	1,329	9.99%
Repeated Records	507	3.81%
First Stop Of The Day Is Not Close	29	0.22%
Day Of Next Trip Is Not Close	334	2.51%
Results		
Inference OF Alighting	8,812	Success Rate
Total Studied 11 L Trips	13,304	66.24%
Total 11 L Trips	13,530	65.13%

On the other hand, 33.76% of the 11 L trips` ODs could not be inferred:

- 9.24% of the inferred alighting location causes the direction error. Since the direction of the buses is determined using the boarding location of 11 L trips, inferred alighting location is supposed to be one of the next stops in the recorded direction. As it is discussed in Chapter 4, recorded direction of the passenger at boarding location is changed to opposite direction if the direction error occurs when the interchanges are made within a very short period of time.

- 4.69% of the studied 11 L trips has no boarding information recorded. Therefore, the direction of the trips are also unknown.
- 3.31% of the inferred alighting location happens to be the same station as recorded boarding location for some 11 L trips. Since the name and the location of the stations are not always the same in the route of 11 L, the alighting locations are analyzed and assigned to an 11 L trip according to its direction.
- 9.99% of the studied 11 L trips have interchanges with boarding locations which have distance to the route of 11 L above the limit of walking distance. It should be noted that in the last trip cluster, 11 L trips are studied further and the next trip analysis is carried out. Hence, these errors, including also the cases analyzed further in last trip cluster, occur when the boarding location of the next trips are not close enough.
- 3.81% of the studied 11 L trips includes the same bus information with the previous 11 L trips within 60 minutes. These records are considered as the results of the repetitive use of the same smartcard.
- 0.22% of the inferred alighting locations for the 11 L trips in last trip cluster are the first trips of the day and they are not close to any station of 11 L route. This rate is relatively low because alighting locations for 11 L trips in the last trip cluster are inferred first from the first trip of the day. If the boarding location of the first trip is not close to the 11 L route then the next trip made in the upcoming days are studied. These errors are detected in the 11 L trips which have no other trips on the other days. The number of trips in that cluster is 73.
- As explained in the previous section, 11 L trips in the last trip cluster are studied in 2 steps. If no meaningful information is extracted from the first trip of the day, then the next trips after 11 L trips are studied. To eliminate the misleading inference, the day difference between next trips and the studied 11 L trips are limited. 2.51% of the inferences has day difference above the limits.

7.1. Inference Rates of Methods

In this study, mainly two clusters are introduced, namely 11 L trips with interchanges in that day and the 11 L trips as the last trips of that day. Inference rate of the interchange cluster is found to be much higher than the last trip cluster, because it is easy to track the passenger in interchange cluster with his/her next trip. On the other hand, the last trip rule applied at the first step of last trip method is based on the assumption that a passenger returns his/her first boarding location with his/her last trip. This requirement couldn't be met for most of the 11 L trips in the last trip cluster. Hence, to analyze these trips another method used at second step is introduced. This method uses the passenger next trip on the upcoming days to infer the alighting location for the last trip of the day with certain assumptions.

It is hard to track the passenger after his/her last trip of the day. Also, the possibility of the passenger to be recorded in his/her next trip on the next days in a different location is quite high. Therefore, the proposed algorithm successfully worked for only small portion of the trips.

Table 7.2. Inference Results of Each Group.

Reason of Error	Interchange	Last Trip	Total
Direction Error	238	991	1229
Missing Stop Id In 11 L Trip	416	208	624
Same Stop	49	391	440
Repeated Records	507	0	507
First Stop Of The Day Is Not Close	0	29	29
Day Of Next Trip Is Not Close	0	334	334
Interchange Stop Is Not Close	418	911	1329
Inference Of Alighting	5083	3729	8812
Total	6711	6593	13304
Inferred	76%	57%	66%
Not Inferred	24%	43%	34%

As explained in the previous chapters, these two clusters are also divided into groups. A total 13,530 11 L trips are analyzed in 9 different groups. The success rates of these clusters are demonstrated in Figure 7.1. As seen in Figure 7.1, interchange to metro cluster has the highest inference rate. Inferences made using the 11 L route or routes with common stops as the next trip or the first trip of the day, are also quite successful. Because in these cases continuity of the trips are satisfied. Since “interchange to the same bus” cluster is introduced to detect the repetitive use of the same smartcard, it is expected to have low success rate in this cluster. Inferences made using the next trips made in other routes is quite high in the “interchange in 2 hours” cluster. However, the inference rate of the interchanges made after 2 hours dramatically decreases. The reason behind this is that when the time interval between the trips increases the possibility of the passengers to make private trips or walk above the limits increases also. One of the lowest inference rate is determined for the 11 L trips which are the last trips of the day and no other trips for the passengers are recorded on the other days. Even though the rate of successfully inferred alighting locations is relatively low in the cluster introduced for analyzing the 11 L trips which no inference could be made from the first boarding of the day, almost 40% percent of the trips are explained by the method proposed in this thesis.

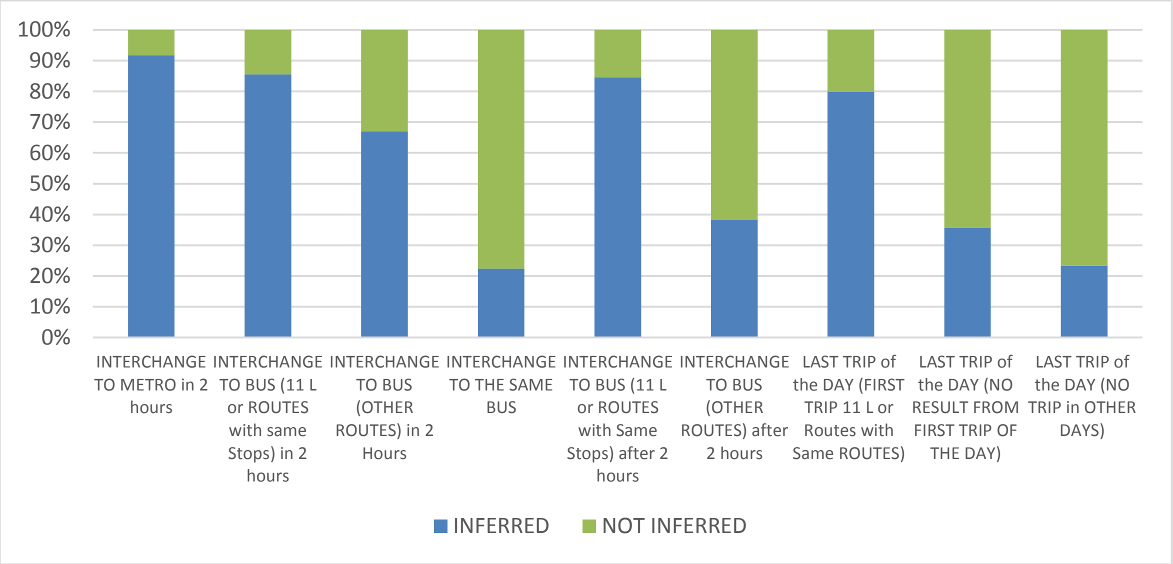


Figure 7.1. Inference Rate of Each Cluster.

7.2. Inference Rate in Each Direction

Inference rate of the trips in each direction is almost equal to each other. These rates are higher than the overall inference rate of the study because some 11 L trips have no records in terms of boarding location or direction of the trips. This is shown as “missing stop id in 11 L trip” in Table 7.1.

Table 7.3. Inference Results of Each Direction.

	Bulgurlu Direction		Üsküdar Direction	
Total Trips	6259		6598	
Alighting Inferred	4253		4559	
Inference Rate	68%		69%	
Inferred By Interchange	830	20%	4253	93%
Inferred By Last Trip	3423	80%	306	7%

Table 7.3 also shows by which cluster these trips are inferred. As the Bulgurlu direction is considered to be the return trip of the passengers, for most of the trips in Bulgurlu direction, the alighting locations are inferred by the last trip methods while for the Üsküdar direction they are inferred by the interchange methods.

Based on these results, it can be concluded that each algorithm might be customized with assumptions which are mostly valid for only one direction to improve the overall inference rate.

7.3. Origin and Destination Matrices

The number of passengers boarding and alighting is shown in Table 7.4 and Table 7.5 for each direction. The rows show the boarding while the columns refer to the alighting. Therefore, the total number of boarding is in the very right column of the tables and the total alighting is provided at the bottom of the tables. Since all the direction errors were checked and if possible corrected during the process with the proposed algorithm, no alighting is seen below the diagonal.

The number of the total inferred alighting and boarding at each stop is illustrated in Figure 7.2 and Figure 7.3. For the inferences in Bulgurlu direction, alighting locations were distributed uniformly at the section of the route close to the terminal station of Esatpaşa in Bulgurlu direction. These stops are located in a territory largely consists of residences and most of the passengers alighted at these stops live around the bus stops. Hence, there is not any distinctive feature of these stops that differentiate one from the other. There is a comparatively large dominance of Horhor bus stops among the boarding locations recorded in the ADC data. Even though the inferred alighting proportions were different in the opposite direction, Horhor is found to be the most congested bus stop of 11 L route in the Bulgurlu direction. As Wang (2010) stated the same problem in his study for London, the reason behind this might be that passengers tend to walk in places between the stops if the stops are close to the shopping centers. In Üsküdar case, there is not only shopping centers between the Horhor station and the terminal station of 11 L route to the Üsküdar direction but also many historical places. Therefore, passengers tend to alight at Horhor station even though it is not the closest stop to the stations of ferries and Marmaray.

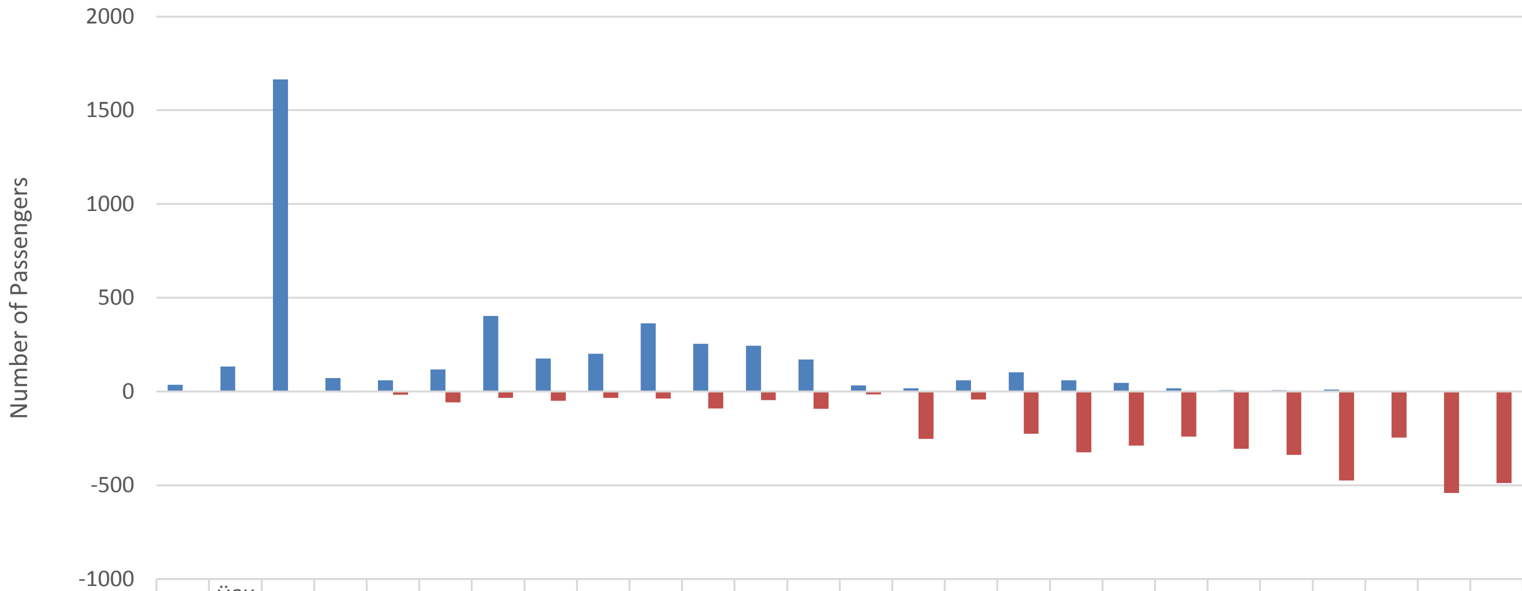
As discussed in the previous chapters, alighting locations in Üsküdar direction mostly inferred by the interchanges after 11 L trips. And since in the inference for interchanges closest 11 L stops to the boarding of next trips is assigned as the alighting location, most of the alighting inferences were found at Metrobüs Altunizade, Üsküdar Marmaray and Üsküdar Cami Önü bus stops which are the closest bus stops to Metrobüs BRT line, Marmaray subway and the ferries, respectively.

Table 7.4. Origin and Destination Matrix of 11 L Route in Bulgurlu Direction.

		ALIGHTING																									
		ÜSKÜDAR CAMİ ÖNÜ	ÜSKÜDAR MARMARAY	HORHOR	BÜLBÜL DERESİ	SETBAŞI	FISTIKAĞACI	KURUÇEŞME	KÜLTÜR MERKEZİ	BAĞLARBAŞI	CAPİTOL	ALTUNİZADE	MİLLET BAHÇESİ	KISIKLI	ÇAMLICA İÖOKULU	BULGURLU	GAZİLER	BULGURLU CADDESİ	BAĞLARİÇİ	DOSTLUK PARKI	ALTINKÖY	DÖRTYOL	TUFAN SOKAK	AÇAN SOKAK	ÜÇYOL	DOĞAN SOKAK	ESATPAŞA	TOTAL
BOARDING	ÜSKÜDAR CAMİ ÖNÜ	0	0	1	0	0	0	0	2	2	0	1	0	0	0	2	0	3	5	4	0	3	1	3	2	5	2	36
	ÜSKÜDAR MARMARAY	0	0	0	0	0	3	2	3	0	0	2	8	1	0	8	0	11	13	6	9	11	15	19	8	12	2	133
	HORHOR	0	0	0	6	18	55	31	28	19	24	36	21	46	4	100	20	104	126	116	86	116	111	189	97	180	131	1664
	BÜLBÜL DERESİ	0	0	0	0	0	0	0	0	3	2	3	1	1	0	3	0	2	4	7	7	6	4	6	3	13	6	71
	SETBAŞI	0	0	0	0	0	0	0	7	0	0	5	1	2	0	6	0	5	1	3	3	4	2	6	4	4	6	59
	FISTIKAĞACI	0	0	0	0	0	0	1	1	2	4	14	2	2	0	8	1	3	10	10	8	13	9	9	5	8	8	118
	KURUÇEŞME	0	0	0	0	0	0	0	9	8	5	15	8	15	1	31	1	15	30	31	23	25	27	39	29	54	37	403
	KÜLTÜR MERKEZİ	0	0	0	0	0	0	0	0	1	3	4	1	7	1	12	2	9	17	24	7	11	11	14	12	29	11	176
	BAĞLARBAŞI	0	0	0	0	0	0	0	0	0	0	7	2	4	1	15	4	12	19	10	9	13	19	22	12	24	28	201
	CAPİTOL	0	0	0	0	0	0	0	0	0	0	4	2	11	4	29	6	14	40	39	20	28	30	39	22	44	31	363
	ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	1	2	2	14	3	12	22	10	25	22	19	34	14	50	24	254
	MİLLET BAHÇESİ	0	0	0	0	0	0	0	0	0	0	0	0	1	0	11	4	22	11	17	19	16	40	26	12	36	29	244
	KISIKLI	0	0	0	0	0	0	0	0	0	0	0	0	0	2	10	1	5	13	3	3	13	12	19	7	18	64	170
	ÇAMLICA İÖOKULU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	3	1	1	0	2	1	0	0	8	13	33
	BULGURLU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	2	0	2	0	0	4	6	17
	GAZİLER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	2	1	6	4	3	4	3	13	18	59
	BULGURLU CADDESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	4	8	8	17	23	5	15	13	102
	BAĞLARİÇİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	7	7	16	3	9	13	60
	DOSTLUK PARKI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	2	1	3	3	9	25	46
	ALTINKÖY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	1	4	3	3	17
	DÖRTYOL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	1	2	7
	TUFAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	5	6
	AÇAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	7	10
	ÜÇYOL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	DOĞAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
	ESATPAŞA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	TOTAL	0	0	1	6	18	58	34	50	35	38	91	47	92	15	253	43	226	324	288	241	305	339	474	246	541	488	4253

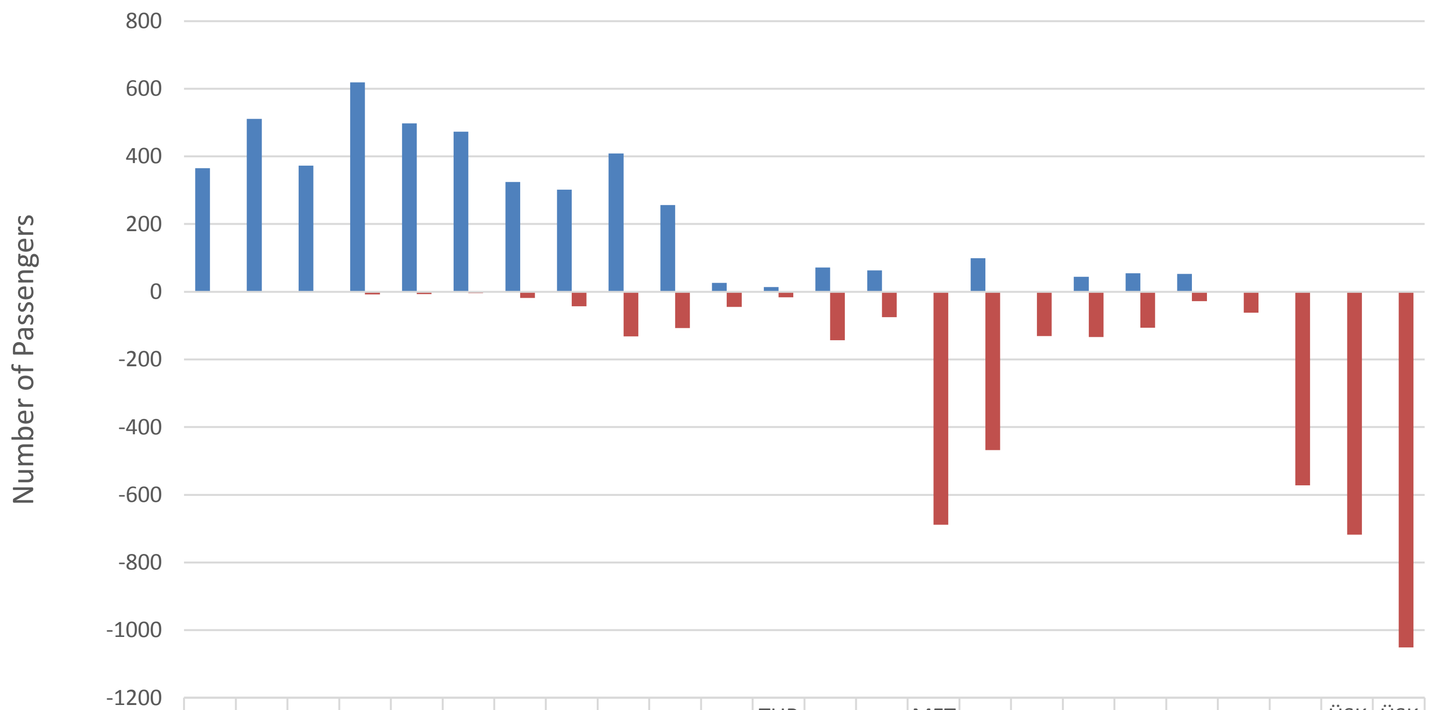
Table 7.5. Origin and Destination Matrix of 11 L Route in Üsküdar Direction.

		ALIGHTING																								
		ESATPAŞA	DOĞAN SOKAK	ÜÇYOL	AÇAN SOKAK	TUFAN SOKAK	DÖRTYOL	ALTINKÖY	DOSTLUK PARKI	BAĞLARİÇİ	ALVARLIZADE CAMİİ	FERAH CADDESİ	TURİSTİK ÇAMLICA TES	KISIKLI	MİLLET BAHÇESİ	METROBÜS ALTUNİZADE	ALTUNİZADE	KÜLTÜR MERKEZİ	KURUÇEŞME	FISTIKAĞACI	SETBAŞI	BÜLBÜL DERESİ	HORHOR	ÜSKÜDAR MARMARAY	ÜSKÜDAR CAMİ ÖNÜ	**TOTAL**
BOARDING	ESATPAŞA	0	2	1	2	2	1	0	10	26	23	11	1	39	11	37	41	14	6	4	1	2	34	36	61	**365**
	DOĞAN SOKAK	0	0	0	5	1	0	2	7	11	24	1	6	13	10	56	63	10	19	10	2	3	63	74	131	**511**
	ÜÇYOL	0	0	0	1	1	0	2	3	8	4	0	0	9	4	65	31	13	16	3	2	4	46	43	118	**373**
	AÇAN SOKAK	0	0	0	0	3	0	2	5	30	12	5	0	11	18	120	57	11	14	11	4	5	82	102	127	**619**
	TUFAN SOKAK	0	0	0	0	0	3	6	9	21	19	6	3	20	7	84	41	9	10	12	3	5	48	70	122	**498**
	DÖRTYOL	0	0	0	0	0	0	6	7	23	8	13	0	11	6	79	61	16	9	9	2	8	48	51	116	**473**
	ALTINKÖY	0	0	0	0	0	0	0	2	9	9	4	2	5	6	85	38	4	11	11	0	7	42	28	61	**324**
	DOSTLUK PARKI	0	0	0	0	0	0	0	0	4	3	3	3	10	1	40	33	19	13	9	3	3	52	51	55	**302**
	BAĞLARİÇİ	0	0	0	0	0	0	0	0	0	5	2	0	15	3	69	45	18	12	15	5	6	53	93	68	**409**
	ALVARLIZADE CAMİİ	0	0	0	0	0	0	0	0	0	0	0	1	7	8	32	39	9	7	7	4	8	24	57	53	**256**
	FERAH CADDESİ	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	5	2	0	0	0	0	2	7	4	**26**
	TURİSTİK ÇAMLICA TES	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	1	1	0	0	0	3	2	2	**14**
	KISIKLI	0	0	0	0	0	0	0	0	0	0	0	0	0	1	15	7	4	3	1	0	1	10	19	11	**72**
	MİLLET BAHÇESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	0	12	5	1	1	12	9	18	**63**
	METROBÜS ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	**1**
	ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	8	1	3	20	22	43	**99**
	KÜLTÜR MERKEZİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
	KURUÇEŞME	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	10	12	17	**44**
	FISTIKAĞACI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	18	28	**55**
	SETBAŞI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	13	24	15	**53**
	BÜLBÜL DERESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	**2**
	HORHOR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
	ÜSKÜDAR CAMİ ÖNÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
	ÜSKÜDAR MARMARAY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
	TOTAL	0	2	1	8	7	4	18	43	132	107	45	16	143	75	688	468	131	134	106	28	62	572	718	**1051**	**4559**



	ÜSKÜDAR CAMİ ÖNÜ	ÜSKÜDAR MARMARAY	HORHOR	BÜLBÜLDERESİ	SETBAŞI	FISTIKAĞACI	KURUÇEŞME	KÜLTÜRMERKEZİ	BAĞLARBAŞI	CAPİTOL	ALTUNİZADE	MİLLET BAHÇESİ	KISIKLI	ÇAMLICA İÖOKULU	BULGURLU	GAZİLER	BULGURLU CADDESİ	BAĞLARİÇİ	DOSTLUK PARKI	ALTINKÖY	DÖRTYOL	TUFAN SOKAK	AÇAN SOKAK	ÜÇYOL	DOĞAN SOKAK	ESATPAŞA
BOARDING	36	133	1664	71	59	118	403	176	201	363	254	244	170	33	17	59	102	60	46	17	7	6	10	1	3	0
ALIGHTING	0	0	-1	-6	-18	-58	-34	-50	-35	-38	-91	-47	-92	-15	-253	-43	-226	-324	-288	-241	-305	-339	-474	-246	-541	-488

Figure 7.2. The Number of Boarding and Alighting at Each Stop of 11 L Route in Bulgurlu Direction.



	ESATPAŞA	DOĞAN SOKAK	ÜÇYOL	AÇAN SOKAK	TUFAN SOKAK	DÖRTYOL	ALTINKÖY	DOSTLUK PARKI	BAĞLARİÇİ	ALVARLIZADE CAMİİ	FERAH CADDESİ	TURİSTİK ÇAMLICA TES	KISIKLI	MİLLET BAHÇESİ	METROBÜS ALTUNİZADE	ALTUNİZADE	KÜLTÜR MERKEZİ	KURUÇEŞME	FISTIKAĞACI	SETBAŞI	BÜLBÜLDERESİ	HORHOR	ÜSKÜDAR CAMİ ÖNÜ	ÜSKÜDAR MARMARAY
BOARDING	365	511	373	619	498	473	324	302	409	256	26	14	72	63	1	99	0	44	55	53	2	0	0	0
ALIGHTHTING	0	-2	-1	-8	-7	-4	-18	-43	-132	-107	-45	-16	-143	-75	-688	-468	-131	-134	-106	-28	-62	-572	-718	-105

Figure 7.3. The Number of Boarding and Alighting at Each Stop of 11 L Route in Üsküdar Direction.

7.4. Validation

7.4.1. Validation of the Proposed Algorithm

As discussed in the previous chapters, inference of alighting location for the 11 L trips which are the last trips of the day, are made in two steps. First step is to find the first trip of the day and its alighting location. Passengers are assumed to come back their starting location and last trip of the day ends at the boarding location of the first trip. If no alighting location is inferred by this method then it is checked whether it is possible to infer an alighting location from the boarding location of the next trip. Validation of these two methods used in “the last trip” cluster for inference of alighting location is performed and they are compared with each other.

The method used in the second step is applied to 2502 11 L trips for which alighting locations are successfully inferred by the first method. However, to apply the second method some limitations are set for the next trip of 11 L, Therefore, only 1834 of 2502 11 L trips satisfy the limits for the next trip. After determining these 11 L trips, second method is applied to them for inference of the alighting locations. At the end of the process, 1001 of 1834 11 L trips provided the same alighting location in both methods. This means in 54% of 11 L trips both methods are applicable.

7.4.2. Validation of the Inferred Results with the Surveys

To validate the inferred boarding and alighting locations with passengers’ actual origins and destinations, trip surveys were conducted at several trips of 11 L at peak hours. Four 11 L (two in each direction) trips were surveyed in morning and evening peak hours on the 5th of January, 2015. About 250 passengers were counted. The arrival time of the buses at the terminal station and Esatpaşa, was 08:10 and 17:25 which are during the peak hours.

. Even though the sample size was small, it still gave the same trend with the results determined by the origin and destination inference especially for the main stations on the route using the proposed algorithm. The alighting location of the passengers in the surveyed trips was slightly different from the inferred alighting locations. This might results from the time of surveyed trips. To increase the sample size, additional surveys were conducted at peak hours. However, it is observed that some passengers got off the bus before their destinations because of the traffic congestion. Instead of waiting in the bus, they preferred to walk. It is quite possible in the route of 11 L because the distances between the bus stops are relatively short.

There is a considerable difference between the surveyed and inferred ratios at the terminal stations. Figure 7.4 and Figure 7.5 demonstrate the inferred and surveyed alighting and boarding locations of the passengers at each bus stop as the ratio to the total ridership of the trip. Since the survey was conducted on a weekday, the inferences of the weekdays is shown in the figures. The differences can be explained as:

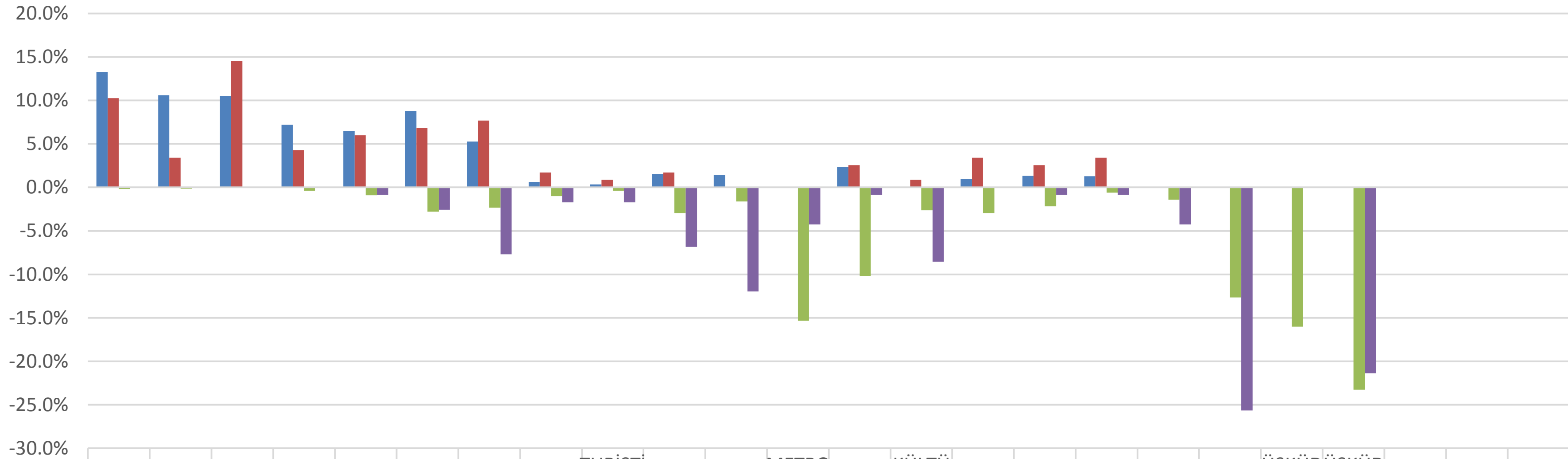
- To the Üsküdar direction, "Üsküdar Marmaray" station is inferred as alighting location in the study for the interchanges to Marmaray because it is the closest station of 11 L route to Marmaray station in Üsküdar. Since in the proposed algorithm the closest stop was searched and assigned as the alighting location for these interchanges, for all the 11 L trips that had an interchange to Marmaray, Üsküdar Marmaray bus stop was inferred as the destination. However, in practice it takes 11 L buses longer to reach the Üsküdar Marmaray bus stop, so most of the passengers are getting off the bus at the previous bus stop, namely Horhor. Therefore, alighting at Horhor bus stop is actually the sum of inferred alightings at Horhor and Üsküdar Marmaray stations. As it is in Figure 7.4, the ratio of surveyed alightings at Horhor is approximately equal to the sum of these two ratios.
- Inferred boardings to the Bulgurlu direction gave different results especially at Üsküdar Marmaray. Initially, no difference was expected between the inferred boarding and survey results since most of the boarding data were inferred from recorded actual ADC data. The difference might be resulted from the passengers' transit usage habits in peak hours. At these hours, it is quite possible that the trip durations might reach undesirable levels because of the congestion. For this reason, passengers tend care much to sit in buses during peak hours. Thus, they prefer to wait in the first stops to find a comfortable place in bus. However, during the daytime the comfort is not the primary concern of the passengers. Also, none of the surveys was

conducted during off-peak hours in this study. So, it is assumed that the differences between the surveyed and recorded alighting are due to the difference in the habits of transit riders during a day.

- There are differences also in the stations where the alighting and boarding ratios are low. This is also considered to be a consequence of the small sample size of the conducted survey.

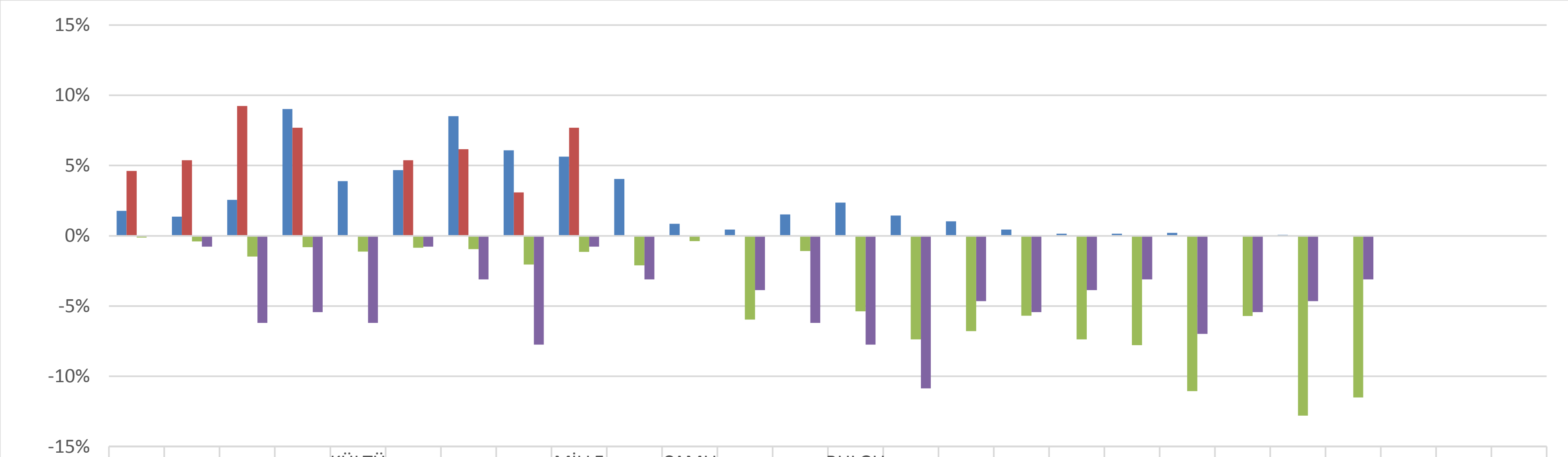
The correlations between the surveyed and inferred boarding are calculated as 0.85 for both direction. This means that there is a high correlation between the inferred and surveyed results. Since most of the boardings are recorded in the ADC systems, this result is not surprising. However for the alightings, correlations between inferred and surveyed results are relatively low. For Üsküdar direction, the correlation is found as 0.58. However, if a correction is made for alighting at Üsküdar Marmaray stop in Üsküdar direction, because of the aforementioned reasons, the correlation reaches almost 0.70. On the other hand, for the alightings in Bulgurlu direction, correlation drops below 0.5. Actually, this result is compatible with the inference rate of last trip cluster. Since the inference rate in last trip cluster is less than the inference rate in interchange cluster, it is expected to see such result in correlation also.

.



	AÇAN SOKAK	TUFAN SOKAK	DÖRTY OL	ALTINK ÖY	DOSTL UK PARKI	BAĞLA RİÇİ	ALVARL IZADE CAMİİ	FERAH CADDE Sİ	TURİSTİ K ÇAMLI CA TES	KISIKLI	MİLLET BAHÇE Sİ	METRO BÜS ALTUNİ ZADE	ALTUNİ ZADE	KÜLTÜ R MERKE Zİ	KURUÇ EŞME	FISTIKA ĞACI	SETBAŞ I	BÜLBÜ L DERESİ	HORHO R	ÜSKÜD AR MARM ARAY	ÜSKÜD AR CAMİ ÖNÜ		
■ 8,3%	13.3%	10.6%	10.5%	7.2%	6.5%	8.8%	5.3%	0.6%	0.3%	1.5%	1.4%	0.0%	2.3%	0.0%	1.0%	1.3%	1.3%	0.0%	0.0%	0.0%	0.0%		
■ 6,0%	10.3%	3.4%	14.5%	4.3%	6.0%	6.8%	7.7%	1.7%	0.9%	1.7%	0.0%	0.0%	2.6%	0.9%	3.4%	2.6%	3.4%	0.0%	0.0%	0.0%	0.0%		
■ 0,0%	-0.2%	-0.1%	0.0%	-0.4%	-0.9%	-2.8%	-2.3%	-1.0%	-0.4%	-2.9%	-1.6%	-15.3	-10.2	-2.6%	-2.9%	-2.2%	-0.6%	-1.4%	-12.7	-16.0	-23.3		
■ 0,0%	0.0%	0.0%	0.0%	0.0%	-0.9%	-2.6%	-7.7%	-1.7%	-1.7%	-6.8%	-12.0	-4.3%	-0.9%	-8.5%	0.0%	-0.9%	-0.9%	-4.3%	-25.6	0.0%	-21.4		

Figure 7.4. Rates of Inferred vs Surveyed Boarding and Alighting at Each Stop of 11 L Route in Üsküdar Direction.



	BÜLBÜL DERESİ	SETBAŞI	FISTIK AĞACI	KURUÇEŞME	KÜLTÜR MERKEZİ	BAĞLARBAŞI	CAPİTOL	ALTUNİZADE	MİLLET BAHÇESİ	KISIKLI	ÇAMLICA İÖOKULU	BULGURLU	GAZİLER	BULGURLU CADDESİ	BAĞLARİÇİ	DOSTLUK PARKI	ALTINKÖY	DÖRTYOL	TUFAN SOKAK	AÇAN SOKAK	ÜÇYOL	DOĞAN SOKAK	ESATPAŞA			
8,3%	2%	1%	3%	9%	4%	5%	9%	6%	6%	4%	1%	0%	2%	2%	1%	1%	0%	0%	0%	0%	0%	0%	0%			
6,0%	5%	5%	9%	8%	0%	5%	6%	3%	8%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%			
0,0%	0%	0%	-1%	-1%	-1%	-1%	-1%	-2%	-1%	-2%	0%	-6%	-1%	-5%	-7%	-7%	-6%	-7%	-8%	-11%	-6%	-13%	-12%			
0,0%	0%	-1%	-6%	-5%	-6%	-1%	-3%	-8%	-1%	-3%	0%	-4%	-6%	-8%	-11%	-5%	-5%	-4%	-3%	-7%	-5%	-5%	-3%			

Figure 7.5. Rates of Inferred vs Surveyed Boardings and Alightings at Each Stop of 11 L Route in Bulgurlu Direction.

8. CONCLUSIONS AND RECOMMENDATIONS

8.1. Comments on the Key Findings

8.1.1. Times of the Last and First Trips of Commuters

Successfully inferred alighting locations in the first step of the last trip method are expected to belong to the passengers who use the 11 L route in regular basis. In general, this kind of passengers mostly consists of commuters. Trips of commuters are largely concentrated in the morning and evening peak hours because these passengers use the public transportation to go to work, school or universities all along weekdays. Hence, it is expected to observe these passengers using the 11 L route more frequently.

After the analysis of the passengers of 2502 11 L trips which were successfully explained in the first step of last trip method, average number of 11 L trips of these passengers during the studied 9 days found to be approximately 9 trips, while the average number of 11 L trips for all passengers was about 4 trips.

When the first and last trip times of these passengers are studied, it is seen that the first trips were mainly made in morning peak hours while the last trips were made evening peak hours. And the number of trips at off-peak hours are extremely low.

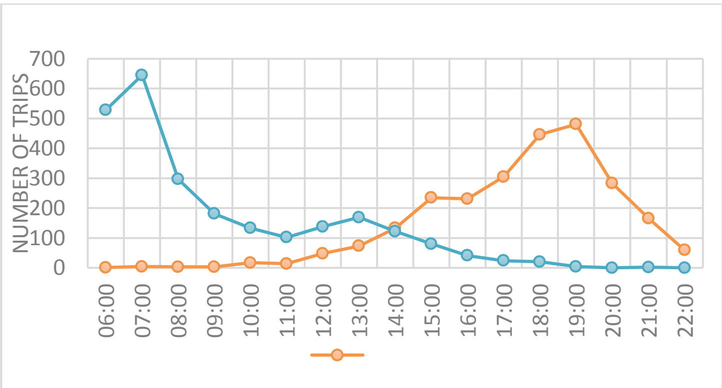


Figure 8.1. Times of the Last and First Trips of Commuters.

Also the time interval between the first and last trips of these passengers are analyzed and the average duration is calculated as 8.5 hours. This value is decreased to 7.5 hours when all other passengers in last trip cluster are taken into account.

Time interval between the last trip and the first trip of the day for the records of passengers whose alighting locations were inferred at the first step of last trip method and considered to be the commuters and the passengers who are studied at the second step of the last trip rule, are demonstrated in Figure 8.2. There is a huge difference between these two groups in this sense. Second group shows a very irregular trend in terms of difference between the last and first trip times of the day, while the commuters mostly have the time differences which approximately equal to the average working time with the consideration of the travel times for the first trips made at the morning before arriving at workplaces or schools. Large time differences seen at the upper part of Figure 8.2 can be explained by the after work activities.

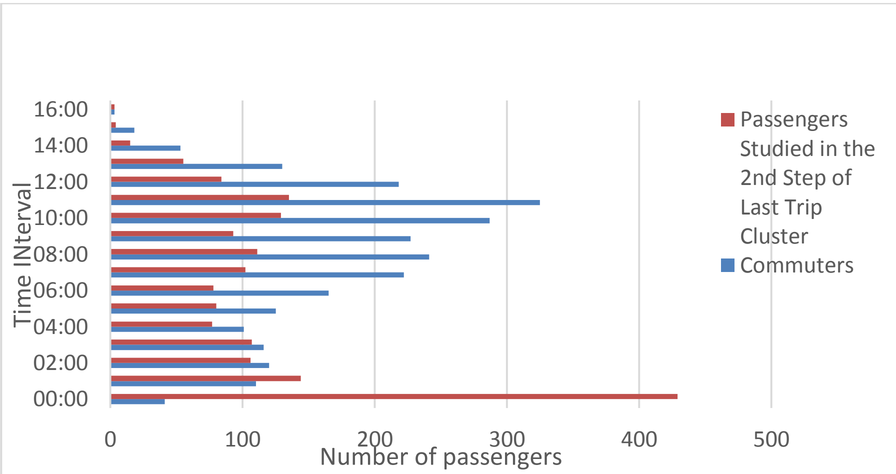


Figure 8.2. Time Intervals between the Last and First Trip of the Day.

8.1.2. Times of Repetitive Use of the Same Smartcard

As it is discussed in the previous chapters, in ADC dataset there were some records which were obviously resulted from the repetitive use of the same smartcard by the cardholder for the other passengers who don’t have tickets or required credit in their smartcards. These records are extremely misleading and needed to be extracted from the dataset. The passengers who used other passengers’ smartcard to make the payments probably don't use the public transportation regularly. Time of these repetitive records are analyzed and it is found that most of these records were seen before and during the evening peak hours. It might be claimed that irregular users of public transportation make their trips during these hours.

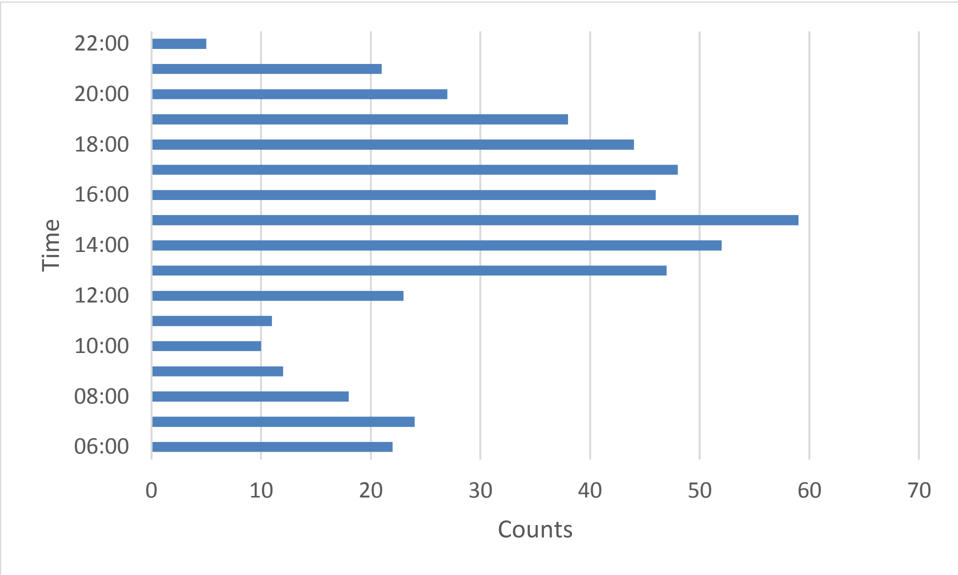


Figure 8.3. Times of Repetitive Use of the Same Smartcard.

The total ridership in hours for all examined trips is demonstrated in Figure 8.4. The evening hours, have much larger ridership than that of the morning hours. In fact the ridership before the evening peak hours is almost as high as the ridership in the morning peak hours. When Figure 8.3 and Figure 8.4 are considered together one can conclude that the increase in the ridership before and during evening the peak hours might be the product of the trips made by the passengers who use public transportation systems in an irregular basis. This conclusion is compatible with the result of the study made by Jun *et al.* (2014). For the analysis made on the AFC data of the route in Nanning City, China, they found that during the morning peak hours, the percentages of commuting card use is about 50% while for the evening peak hours, this rate is generally lower.

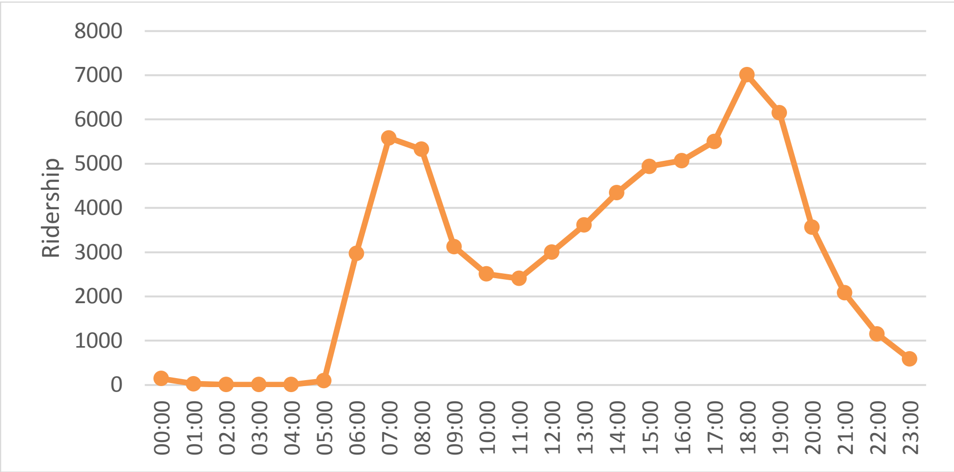


Figure 8.4. The Total Ridership in Hours for All Trips.

8.1.3. Interchange Time

The time differences between the successive trips are shown in Figure 8.5. As demonstrated in Figure 8.5, most frequently observed time intervals between consecutive trips are between 20 and 30 minutes. Hofmann and Mahony (2005) also made an analysis for the time difference between the successive trips in their study. Even though different from our study they set 90 minutes threshold for interchange time interval, in their histogram of time differences between consecutive trips also mostly seen interchanges are between the same time intervals as in this study.

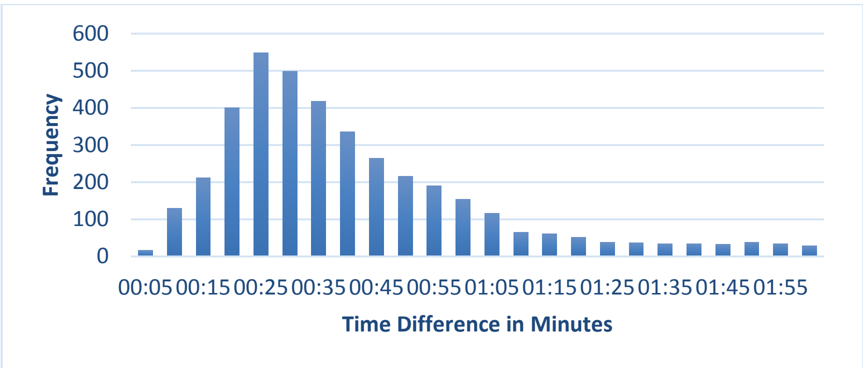


Figure 8.5. Consecutive Boarding Times Difference.

8.1.4. Passengers of Single Trip on All Day

In the single trip cluster, there are 81 trips, all of which are 11 L trips and no further trips are recorded for these passengers. In Table 8.1, the type of tickets in this category is shown. As it is seen, unlike the percentages of ticket types for all ADC data, in this cluster mostly elder people and full rate ticket users are recorded. From this analysis it is understood that especially for elder passengers who don't use public transportation very often, much detailed analysis should be made and the sample size should be much larger to collect required number of records to make a statistical analysis.

Table 8.1. Type of Ticket Used by the Passengers Recorded Only Once on All Days.

Type of Ticket	Counts	Percentage
Over 65 Age	8	10%
Full Rate Ticket	50	62%
Discount Ticket	10	12%
No Data	2	2%
Teacher	2	2%
Martry Wife	1	1%
Yellow Press	1	1%
Elder	5	6%
Police	1	1%
Blue Card	1	1%
Total	81	100%

8.2. Conclusions

OD matrices with wide-range of usage can help transit agencies to improve the quality of transportation systems, in several aspects. With the information of origins and destinations of the passengers, transit planners can detect the critical and mostly used interchange stations. With the help of this information, required improvements can be introduced into these locations.

If the OD matrices are generated on the network level, they also give very useful outputs. For instance, passenger flows during any day can be explained by the results found in the generation of ODs for transit ridership. These results can be used by not only transit planners but also the city planners. To monitor the origin and destination of the passengers directly gives the very informative data to detect the residential locations and the locations mainly consisting of the work places. With this information, city planners can improve their decision making process about the possible locations of the planned new industrial zones or the new residential areas.

In this analysis OD estimation for a single route in İstanbul is made. With the help of previous studies about this topic, basic assumptions were taken to infer the alighting locations of the studied trips of September 15-23, 2014. With the proposed algorithm, using new methods and assumptions, the trips which cannot be studied under the assumptions of the previous studies, were also analyzed.

The overall results showed that it is quite possible to estimate the destinations of passengers for the transit agency of İstanbul with only the analysis of recorded ADC data. Using the proposed algorithm, over 65% of all recorded trips alighting location was inferred. If the ADC data with missing boarding information or the repetitive use of the same card were excluded from the dataset, this success rate would reach to 70 percent.

As it is seen in this study, unique characteristics of the transportation systems should be further analyzed to eliminate the misleading outputs of any proposed algorithm. Since the passengers' public transportation usage habits are different from each other in different metropolises, the most appropriate assumptions and methods should be set with the consideration of these information. It is also known that transportation agencies in different cities apply different rules and systems. Therefore, properties of the studied public transportation systems should be analyzed carefully before the OD inference study is started.

With the comparison made between the result of this study and the observations made in the studied route, it is concluded that even if the assumptions made in the analysis are quite consistent and rational, it is highly possible to see different results in practice.

This shows the difficulty of analyzing a system whose main component is the passengers. Human factors sometimes cannot be explained even by the accurate assumptions and methods. Therefore, it is suggested to further continue to conduct surveys in transportation systems not to collect data about the alighting and boarding location of the passengers but to understand the behavior and priorities of the passengers when using public transportation systems.

APPENDIX A: ORIGIN AND DESTINATION MATRICES

This page is intentionally left blank.

Table A.1. Origin and Destination Matrix of 11 L Route in Bulgurlu Direction for the Weekdays.

BOARDING \ ALIGHTING	ÜSKÜDAR CAMİ ÖNÜ	ÜSKÜDAR MARMARAY	HORHOR	BÜLBÜL DERESİ	SETBAŞI	FISTIKAĞACI	KURUÇEŞME	KÜLTÜR MERKEZİ	BAĞLARBAŞI	CAPİTOL	ALTUNİZADE	MİLLET BAHÇESİ	KISIKLI	ÇAMLICA İÖOKULU	BULGURLU	GAZİLER	BULGURLU CADDESİ	BAĞLARİÇİ	DOSTLUK PARKI	ALTINKÖY	DÖRTYOL	TUFAN SOKAK	AÇAN SOKAK	ÜÇYOL	DOĞAN SOKAK	ESATPAŞA	TOTAL
ÜSKÜDAR CAMİ ÖNÜ	0	0	1	0	0	0	0	2	2	0	1	0	0	0	2	0	3	5	4	0	3	1	3	2	5	2	36
ÜSKÜDAR MARMARAY	0	0	0	0	0	3	2	3	0	0	2	7	0	0	8	0	10	12	6	7	10	14	15	6	9	2	116
HORHOR	0	0	0	5	16	55	30	25	19	24	35	21	45	4	94	19	101	112	109	82	110	102	175	91	165	119	1558
BÜLBÜL DERESİ	0	0	0	0	0	0	0	0	3	2	3	1	1	0	3	0	2	3	7	7	6	4	5	3	13	6	69
SETBAŞI	0	0	0	0	0	0	0	7	0	0	5	1	2	0	5	0	4	1	3	2	3	1	6	4	4	5	53
FISTIKAĞACI	0	0	0	0	0	0	0	1	1	3	9	2	2	0	5	1	3	9	7	8	13	9	8	5	7	7	100
KURUÇEŞME	0	0	0	0	0	0	0	6	7	5	13	7	11	1	27	1	13	23	28	21	24	22	32	27	51	33	352
KÜLTÜR MERKEZİ	0	0	0	0	0	0	0	0	1	3	1	1	7	1	12	2	7	15	21	6	11	8	10	8	27	11	152
BAĞLARBAŞI	0	0	0	0	0	0	0	0	0	0	7	2	3	1	14	4	10	16	9	7	13	19	20	10	24	23	182
CAPİTOL	0	0	0	0	0	0	0	0	0	0	4	2	8	4	27	6	12	38	36	18	25	28	37	21	39	27	332
ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	1	2	2	13	3	10	20	10	24	20	19	32	10	49	22	237
MİLLET BAHÇESİ	0	0	0	0	0	0	0	0	0	0	0	0	1	0	10	4	22	8	15	18	15	36	24	11	29	27	220
KISIKLI	0	0	0	0	0	0	0	0	0	0	0	0	0	2	9	1	4	13	2	3	13	8	17	7	16	63	158
ÇAMLICA İÖOKULU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	3	1	1	0	2	1	0	0	8	13	33
BULGURLU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	2	0	2	0	0	4	6	17
GAZİLER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	2	1	6	4	3	4	3	13	18	59
BULGURLU CADDESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9	4	7	7	13	23	5	13	11	92
BAĞLARİÇİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	7	5	14	3	9	13	56
DOSTLUK PARKI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	3	2	8	24	40
ALTINKÖY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	1	4	3	3	17
DÖRTYOL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	1	1	6
TUFAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	5	6
AÇAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	5	8
ÜÇYOL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DOĞAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
ESATPAŞA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TOTAL	0	0	1	5	16	58	32	44	33	37	80	45	82	15	233	42	210	288	265	222	288	304	431	223	499	449	3902

Table A.2. Origin and Destination Matrix of 11 L Route in Bulgurlu Direction for the Weekend.

		ALIGHTING																									
		ÜSKUDAR CAMİ ÖNÜ	ÜSKUDAR MARMARAY	HORHOR	BÜLBÜL DERESİ	SETBAŞI	FISTIKAĞACI	KURUÇEŞME	KÜLTÜR MERKEZİ	BAĞLARBAŞI	CAPİTOL	ALTUNİZADE	MİLLET BAHÇESİ	KISIKLI	ÇAMLICA İÖOKULU	BULGURLU	GAZİLER	BULGURLU CADDESİ	BAĞLARİÇİ	DOSTLUK PARKI	ALTINKÖY	DÖRTYOL	TUFAN SOKAK	AÇAN SOKAK	ÜÇYOL	DOĞAN SOKAK	ESATPAŞA	TOTAL
BOARDING	ÜSKÜDAR CAMİ ÖNÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	ÜSKÜDAR MARMARAY	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	2	1	1	4	2	3	0	17
	HORHOR	0	0	0	1	2	0	1	3	0	0	1	0	1	0	6	1	3	14	7	4	6	9	14	6	15	12	106
	BÜLBÜL DERESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	2
	SETBAŞI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	0	0	0	1	6
	FISTIKAĞACI	0	0	0	0	0	0	1	0	1	1	5	0	0	0	3	0	0	1	3	0	0	0	1	0	1	1	18
	KURUÇEŞME	0	0	0	0	0	0	0	3	1	0	2	1	4	0	4	0	2	7	3	2	1	5	7	2	3	4	51
	KÜLTÜR MERKEZİ	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	2	2	3	1	0	3	4	4	2	0	24
	BAĞLARBAŞI	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	2	3	1	2	0	0	2	2	0	5	19
	CAPİTOL	0	0	0	0	0	0	0	0	0	0	0	0	3	0	2	0	2	2	3	2	3	2	2	1	5	4	31
	ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	2	2	0	1	2	0	2	4	1	2	17
	MİLLET BAHÇESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	3	2	1	1	4	2	1	7	2	24
	KISIKLI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	4	2	0	2	1	12
	ÇAMLICA İÖOKULU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	BULGURLU	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	GAZİLER	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	BULGURLU CADDESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	4	0	0	2	2	10
	BAĞLARİÇİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2	0	0	0	4
	DOSTLUK PARKI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	0	1	1	1	6
	ALTINKÖY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	DÖRTYOL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	TUFAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	AÇAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	2
	ÜÇYOL	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	DOĞAN SOKAK	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	ESATPAŞA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	TOTAL	0	0	0	1	2	0	2	6	2	1	11	2	10	0	20	1	16	36	23	19	17	35	43	23	42	39	351

Table A.3. Origin and Destination Matrix of 11 L Route in Üsküdar Direction for the Weekdays.

		ALIGHTING																							
		ESATPAŞA	DOĞAN SOKAK	ÜÇYOL	AÇAN SOKAK	TUFAN SOKAK	DÖRTYOL	ALTINKÖY	DOSTLUK PARKI	BAĞLARİÇİ	ALVARLIZADE CAMİİ	FERAH CADDESİ	TURİSTİK ÇAMLICA TES	KISIKLI	MİLLET BAHÇESİ	METROBÜS ALTUNİZADE	ALTUNİZADE	KÜLTÜR MERKEZİ	KURUÇEŞME	FISTIKAĞACI	SETBAŞI	BÜLBÜL DERESİ	HORHOR	ÜSKÜDAR MARMARAY	ÜSKÜDAR CAMİ ÖNÜ	TOTAL
BOARDING	ESATPAŞA	0	2	0	2	2	0	0	9	23	23	10	1	37	10	33	37	13	5	3	1	2	30	35	60	338
	DOĞAN SOKAK	0	0	0	5	0	0	2	7	10	22	1	6	12	9	54	60	7	19	9	2	3	57	69	120	474
	ÜÇYOL	0	0	0	1	1	0	2	3	7	3	0	0	9	3	55	26	12	16	3	2	3	44	39	111	340
	AÇAN SOKAK	0	0	0	0	3	0	2	2	25	12	4	0	5	15	115	51	8	10	10	4	5	72	92	110	545
	TUFAN SOKAK	0	0	0	0	0	1	4	9	19	18	6	3	12	6	77	35	5	10	11	2	5	42	61	109	435
	DÖRTYOL	0	0	0	0	0	0	6	5	18	8	11	0	11	5	73	55	15	9	9	2	7	45	46	106	431
	ALTINKÖY	0	0	0	0	0	0	0	2	9	5	4	2	5	6	82	32	4	10	9	0	7	39	27	52	295
	DOSTLUK PARKI	0	0	0	0	0	0	0	0	4	1	3	3	7	1	35	28	15	11	7	3	3	49	48	48	266
	BAĞLARİÇİ	0	0	0	0	0	0	0	0	0	4	2	0	14	3	60	41	13	10	11	4	4	47	90	58	361
	ALVARLIZADE CAMİİ	0	0	0	0	0	0	0	0	0	0	0	1	6	7	26	34	9	6	7	3	8	21	43	45	216
	FERAH CADDESİ	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	5	2	0	0	0	0	2	6	4	25
	TURİSTİK ÇAMLICA TES	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	3	1	1	0	0	0	3	2	2	14
	KISIKLI	0	0	0	0	0	0	0	0	0	0	0	0	0	1	14	7	3	2	1	0	1	8	15	11	63
	MİLLET BAHÇESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	4	0	11	3	1	1	10	9	18	58
	METROBÜS ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
	ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	5	1	3	20	22	43	96
	KÜLTÜR MERKEZİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	KURUÇEŞME	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	10	12	14	41
	FISTIKAĞACI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	18	28	54
	SETBAŞI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	13	24	15	53
	BÜLBÜL DERESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
	HORHOR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	ÜSKÜDAR MARMARAY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	ÜSKÜDAR CAMİ ÖNÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	TOTAL	0	2	0	8	6	1	16	37	115	96	41	16	121	66	630	418	108	121	89	25	58	520	658	955	4107

Table A.4. Origin and Destination Matrix of 11 L Route in Üsküdar Direction for the Weekend.

		ALIGHTING																								
		ESATPAŞA	DOĞAN SOKAK	ÜÇYOL	AÇAN SOKAK	TUFAN SOKAK	DÖRTYOL	ALTINKÖY	DOSTLUK PARKI	BAĞLARİÇİ	ALVARLIZADE CAMİİ	FERAH CADDESİ	TURİSTİK ÇAMLICA TES	KISIKLI	MİLLET BAHÇESİ	METROBÜS ALTUNİZADE	ALTUNİZADE	KÜLTÜR MERKEZİ	KURUÇEŞME	FISTIKAĞACI	SETBAŞI	BÜLBÜL DERESİ	HORHOR	ÜSKÜDAR MARMARAY	ÜSKÜDAR CAMİ ÖNÜ	TOTAL
BOARDING	ESATPAŞA	0	0	1	0	0	1	0	1	3	0	1	0	2	1	4	4	1	1	1	0	0	4	1	1	27
	DOĞAN SOKAK	0	0	0	0	1	0	0	0	1	2	0	0	1	1	2	3	3	0	1	0	0	6	5	11	37
	ÜÇYOL	0	0	0	0	0	0	0	0	1	1	0	0	0	1	10	5	1	0	0	0	1	2	4	7	33
	AÇAN SOKAK	0	0	0	0	0	0	0	3	5	0	1	0	6	3	5	6	3	4	1	0	0	10	10	17	74
	TUFAN SOKAK	0	0	0	0	0	2	2	0	2	1	0	0	8	1	7	6	4	0	1	1	0	6	9	13	63
	DÖRTYOL	0	0	0	0	0	0	0	2	5	0	2	0	0	1	6	6	1	0	0	0	1	3	5	10	42
	ALTINKÖY	0	0	0	0	0	0	0	0	0	4	0	0	0	0	3	6	0	1	2	0	0	3	1	9	29
	DOSTLUK PARKI	0	0	0	0	0	0	0	0	0	2	0	0	3	0	5	5	4	2	2	0	0	3	3	7	36
	BAĞLARİÇİ	0	0	0	0	0	0	0	0	0	1	0	0	1	0	9	4	5	2	4	1	2	6	3	10	48
	ALVARLIZADE CAMİİ	0	0	0	0	0	0	0	0	0	0	0	0	1	1	6	5	0	1	0	1	0	3	14	8	40
	FERAH CADDESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
	TURİSTİK ÇAMLICA TES	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	KISIKLI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	2	4	0	9
	MİLLET BAHÇESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	0	0	2	0	0	5
	METROBÜS ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	ALTUNİZADE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	3
	KÜLTÜR MERKEZİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	KURUÇEŞME	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3
	FISTIKAĞACI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
	SETBAŞI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	BÜLBÜL DERESİ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
	HORHOR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	ÜSKÜDAR MARMARAY	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	ÜSKÜDAR CAMİ ÖNÜ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	TOTAL	0	0	1	0	1	3	2	6	17	11	4	0	22	9	58	50	23	13	17	3	4	52	60	96	452

REFERENCES

Barry, J., R. Newhouser, A. Rahbee, and S. Sayeda, 2002, "Origin and Destination Estimation in New York City Using Automated Fare System Data", *Transportation Research Record*, Vol. 1817, pp. 183-187.

BELBIM, İstanbul Municipality Information Technologies, 2014, http://www.belbim.com.tr, [Accessed December 2014].

Chen, J., Y. Dongyuan, 2013, "Estimating Smart Card Commuters Origin Destination Distribution Based on APTS Data", *Journal of Transportation Systems Engineering and Information Technology,* Vol. 13, pp. 47-53.

Cui, A., 2006, "Bus Passenger Origin-Destination Matrix Estimation Using Automated Data Collection Systems", M.S. Thesis*, Massachusetts Institute of Technology*.

Hofmann, M., M. O'Mahony, 2005, "Transfer Journey Identification and Analyses from Electronic Fare Collection Data", *Proceedings of the 8th International IEEE Conference on Intelligent Transportation Systems Meeting.*

IBB, İstanbul Metropol Municipality, 2014, http://www.ibb.gov.tr, [Accessed December 2014].

IETT, İstanbul Electric Tramway and Tunnel Establishments, 2014, http://www.iett.gov.tr, [Accessed December 2014].

Infodev, Automatic Passenger Counting System, 2014, http://www.infodev.ca, [Accessed December 2014].

Jang, W., 2010, "Travel Time and Transfer Analysis Using Transit Smart Card Data", *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2144, pp. 142-149.

Ma, X., Y.-J. Wu, Y. Wang, F. Chen, J. Liu, 2013, "Mining Smart Card Data for Transit Riders' Travel Patterns", *Transportation Research Part* C, Vol. 36, pp. 1-12.

Trépanier, M., N. Tranchant, R. Chapleau, 2007. "Individual Trip Destination Estimation in a Transit Smart Card Automated Fare Collection System", *Journal of Intelligent Trasnportaton Systems: Technology, Planning and Operations,* Vol. 11, pp. 1-14.

Yang, F., J. Jin, Y. Cheng, B. Ran, 2014, "Origin-Destination Estimation for Non-Commuting Trips Using Location-Based Social Networking Data", *International Journal of Sustainable Transportation*, Accepted.

Wang, W., 2010, "Bus Passenger Origin-Destination Estimation And Travel Behavior Using Automated Data Collection Systems In London, UK", M.S. Thesis, *Massachusetts Institute of Technology*.

Zhao, J., A. Rahbee, N. H. M. Wilson, 2007, "Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems", *Computer- Aided Civil and Infrastructure Engineering*, Vol. 22, No. 5, pp. 376-387.