Australasian Transport Research Forum 2015 Proceedings
30 September - 2 October 2015, Sydney, Australia
Publication website: http://www.atrf.info/papers/index.aspx

New Approach to Calculation of a Transfer Matrix in a Transit Network

Elnaz Irannezhad[1], Seyed Hossein Mousavi[2], Mark Hickman[3]

[1] PhD student, University of Queensland, Australia

[2] Researcher, Amirkabir University of Technology

[3]Professor and ASTRA Chair of Transport, University of Queensland, Australia

Email for correspondence: e.irannezhad@uq.edu.au

Abstract

The transfer rate is considered a metric for the evaluation of the efficiency of current or proposed transit networks. Assessing transfers is very important in the redesign process, in which it is necessary to have a simple method to calculate the transfer rate in the modelling process. With this transfer information, an unreasonably high transfer rate may require the transit system to consider redesign of the network.

The objective of this research is to develop a simple model for transfer calculation that can assist with transit network design, but without the need for detailed data. Initially, the question to be answered is, how many transfers are necessary for every person to travel from any location to another location in the city? Secondly, how many trips are made with how many transfers in the transit network? One may then explore potential improvements in the transit network. The results of the proposed transfer model were validated with real transfer data obtained from Translink, the transit authority in south east Queensland (SEQ).

Key Words: calculation of transfer, transfer rate, public transportation system

1. Introduction

While transfers present additional access to a large number of potential destinations at an affordable cost to the operator, they also impose inconvenience and additional cost in a non-integrated public transport system. Thus, providing transfer information for users is considered as an effective measure for promoting public transport and encouraging people to consider it in their travel choices. In addition, it is not always easy to say whether the current public transport system covers most of the desired trips just by looking at the physical transit route coverage.

One of the main goals and policies that is stressed by transport managers is that people could reach one of public transport stops within a reasonable walking distance of their origins and could finish their trips with a reasonable transfer rate from every origin to every destination. Thus, at the first step some questions may come to mind, such as, what is the transfer rate in the current public transport system, and how could we improve it with adjusting and redesigning the route network?

Obviously, checking where passengers are boarding or alighting would not give a very full picture of origin-destination (OD) patterns unless supplemented with trip start and end information, or surveys of passenger movements at all key transfer points, or farecard readers on transit vehicles or routes to measure fare card use. Based on the standards categorized in Ceder (2007), which is based on questionnaires of transit agencies, a maximum of 1 to 3 transfers for any OD pair is considered reasonable at a planning level.

In most cases, especially in developing and underdeveloped countries, these automatic fare collection systems are not widely used, and there are little or no valuable data for transfer analysis. As transport engineers move into planning, the limitations of using complicated models become more and more apparent. Also, most policies regarding urban transport systems depend on managers and policy makers who may not wish to complete a time and resource-consuming modelling process with detailed data collection and analysis. Thus, going through a sketch planning approach would be valuable in these cases.

A detailed transit assignment procedure needs various characteristics including the detailed transit network, in-vehicle time, waiting time at stops, transfer paths and failure-to-board probabilities (Ceder, 2007). Traditional assignment procedures were extended by considering some additional important factors such as impacts of inconvenience in crowded vehicles, introduced by Spiess and Florian (1989), Nguyen and Pallotino (1988), and in a transit equilibrium model proposed by De Cea and Fernandez (1993).

Specifically regarding transfers, Stern (2010) explored a cognitive scalar for the public transport network and analysed the relationship between cognizant service levels and connectivity for different transfer waiting times. Hadas and Ceder (2010) focused on connectivity and comfort of transfers while categorizing transfer types into non-adjacent, adjacent and shared transfer points.

Dial (1967), in the Pathfinder algorithm, discussed the necessity of considering transfer penalty in route assignment. More progress was achieved in later researches including Chriqui and Robillard (1975) and Nguyen and Pallotino (1988). Later, passenger-choice behaviour was interpreted using a utility function, which includes a set of passenger attitudes. Nuzzolo et al. (2001) proposed a random utility function for passenger path choice with passenger information.

With the development of geographic information systems (GIS), many pathfinder algorithms used GIS. For example, Koncz et al. (1996) employed transit route connectivity (TRC) matrices and GIS tools for path finding; and, Choi and Jang (2000) described a procedure to develop transit network assignment from a street network, using GIS and spatial analysis, and this method was tested in a small case study.

Liu et al. (2001) proposed a simpler method for calculating the number of transfers by using a power of the connectivity matrix. If A_{ij} indicates whether location i is connected by a route to location j, then A_{ij}^n represents whether anyone can travel from location i to j by passing (n-1) intermediate locations (transfers). Later in 2002, he discovered the other property of the transition matrix: if T_{ij} is a cell of the transition matrix that indicates the number of direct routes one may use from location i to location j, then T_{ij}^n would be the number of ways that a person can travel from i to j by n-1 transfers. So, by computing powers of the transition matrix, he introduced the TPlanning algorithm for finding path plans.

In most of the mentioned studies, transfers were regarded as a limitation in the modelling procedure or path finding algorithms, rather than being seen as a comparative tool between network designs. Moreover, collection of a wide variety of data for these models would be difficult; while, using the proposed algorithm in this research, one only needs to know the number of vehicles in the fleets, the length of transit routes, and a map of those routes. Given that all of these data are available in most cities, the proposed mechanism is simple to use and quick to answer questions on transfers and connectivity.

2. Methodology

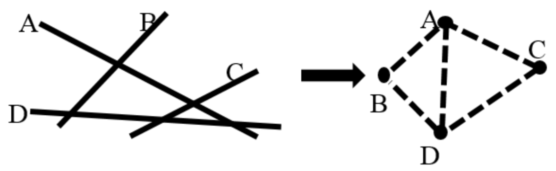
The proposed method for estimating the transfer rate has two major steps. As described below, Step 1 examines coverage without regarding passenger demand, while step 2 gives us the transfer rate with regard to the demand between zones. A simple network is used to illustrate the algorithm.

2.1. Step 1

In the first step, the minimum number of transfers which a person should make to travel between each pair of transit routes is found using the following algorithm:

1. With a map of the transit routes, an adjacency matrix is derived which represents the connectivity among transit routes. The adjacency matrix indicates which transit lines have the possibility of a transfer with each other within a reasonable walking distance. The value of a cell in this matrix is set to 1 if the pair of routes overlap each other and otherwise is set to 0. Actually, each cell of the matrix indicates the possibility of a transfer within a desirable walking distance between the two transit routes. So, a reasonable walking distance is considered as a buffer around each transit route, and connectivity is tested with regard to those buffers, instead of routes.
2. A new synthetic graph based on the obtained adjacency matrix in part 1 is built, in which transit routes are transformed into nodes and the possibility of a transfer between two routes is defined as a link connecting the two nodes. Figure 1 illustrates this general idea.
3. The shortest path on this transformed graph is calculated using Dijkstra's algorithm, which consists of a set N of nodes indicating the number of transit routes and a set A of arcs whose costs were considered one, indicating the number of transfers between each pair of routes.
4. Calculation of the transfer rate matrix

Figure 1: Transforming Graph of the Transit Route System



Adjacency Matrix

	A	B	C	D
A	0	1	1	2
B	1	0	2	1
C	1	2	0	1
D	1	1	1	0

Transfer Matrix

	A	B	C	D
A	0	0	0	0
B	0	0	1	0
C	0	1	0	0
D	0	0	0	0

The value of each cell in the transfer matrix shows the minimum transfers available between each pair of transit routes in given network. Cells with high value represent poor connectivity, but if there is not high demand to pass through these routes, there is no necessity to improve the network.

2.2. Step 2

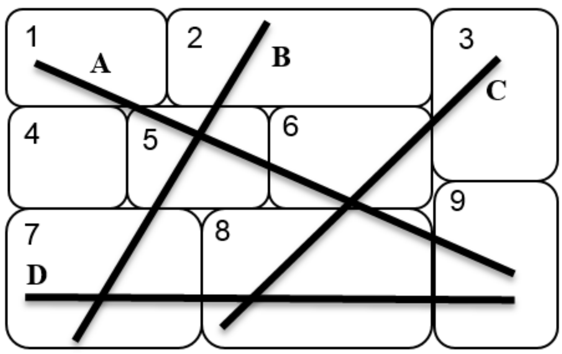
Also, with an origin-destination matrix, the number of passengers who must tolerate these high transfer rates could be found, so this fundamental idea was extended to step 2, described as follows:

1. Find the coverage matrix as a binary matrix which indicates whether each transit line intersects each zone in its 500 metre buffer or not.
2. Build a synthetic graph in which zones are considered as nodes of the graph. Also, if route k (with a buffer of 500 metre around it) passes through both zone i and zone j, then zone i and j are connected with route k.
3. Lacking more detailed data, the cost of each link could be considered as a function of the vehicle fleet per each kilometre of route, which implies the headway and waiting time.

$$cost\ of\ link = \frac{\text{Length of route (Km)}}{Fleet}$$

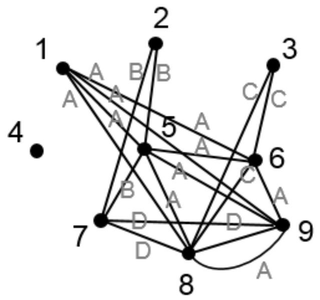
4. Under the assumption that a direct transit route between a pair of zones is considered a better option for passengers, a cost should be regarded for each transfer required. In addition, there is a trade-off between the number of transfers and the headway. For more simplicity, if we do not have any feedback on the passengers' priority, the transfer cost could be considered as the average cost of routes; this means the transfer weight is equal to the cost of a route. Obviously, this assumption depends on the fare policy and the structure of multi-destination stations, which decrease the negative effects of the transfer.
5. Assigning demand between zone i and j to the shortest path between them results in the transfer rate between each pair of routes and in the whole of the network. The following figure illustrates how this idea is demonstrated in a sample network.

Figure 2 Finding the coverage matrix and buildng the connectivity graph of zones



Coverage Matrix

	A	B	C	D
1	1	0	0	0
2	0	1	0	0
3	0	0	1	0
4	0	0	0	0
5	1	1	0	0
6	1	0	1	0
7	0	1	0	1
8	1	0	1	1
9	1	0	0	1



3. Case Study

To validate the proposed algorithm, a complete dataset was needed which includes both touch-on and touch-off stations, route-ID's and times of transactions. Thus, the Go Card dataset for south east Queensland (SEQ) was selected as a case study. In addition, because the results of the model should be tested with real data, some of assumptions should be examined in reality, one of which is the desirable walking distance. One day of transactions related to 5 March 2013 was selected for analysis which, after excluding school buses and other data cleansing resulted in about 400,000 transactions related to only transit routes.

The literature review showed various criteria for transfer identification. Some consider a time threshold for two consecutive boardings; for example, Seaborn et al (2009) used 50 minutes between boardings, or a time threshold for identifying transfers between an alighting and a boarding of 20 min to 45 min, dependent on mode. Hofmann and O'Mahony (2005) used a 90 min interval between boardings, and Bagchi and White (2004) applied 30 min for the time gap between an alighting and a boarding. Munizaga et al. (2012) and Devillane et al. (2012) similarly used a 30 min time interval for identification of consecutive trip legs.

Recently some other researchers proposed further criteria (Chu et al. 2008, Nassir et al. 2011 and Gordon et al. 2013). Nassir et al. (2014) introduced some other criteria for a transfer detection algorithm which was applied to the SEQ Go Card dataset for March 2013.

In this research, for simplicity, 20 min was considered for detection of legs of a journey, and the results of analysis of one-day was similar to one-month analysis of Nassir et al. (2014), as shown in Table 1. Figure 3 illustrates how the transfer rate was determined in this data set.

Figure 3: Flowchart of finding transfer rate in Go card dataset

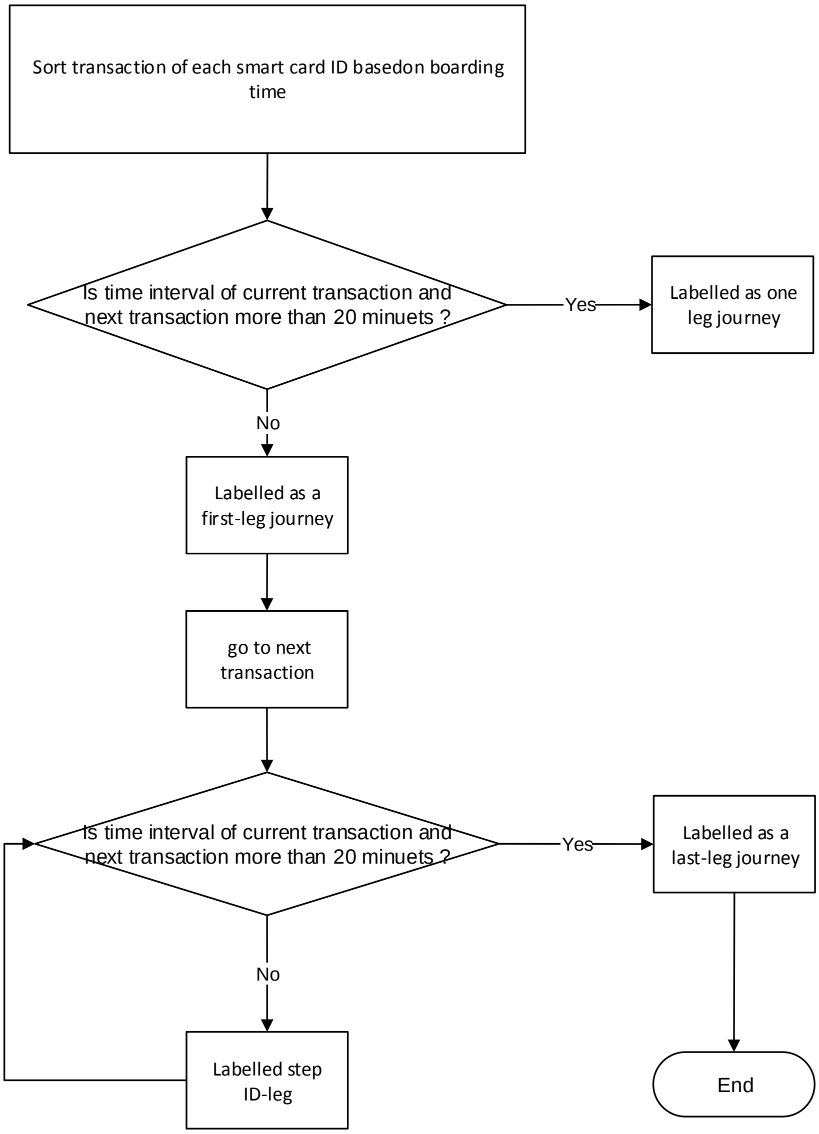


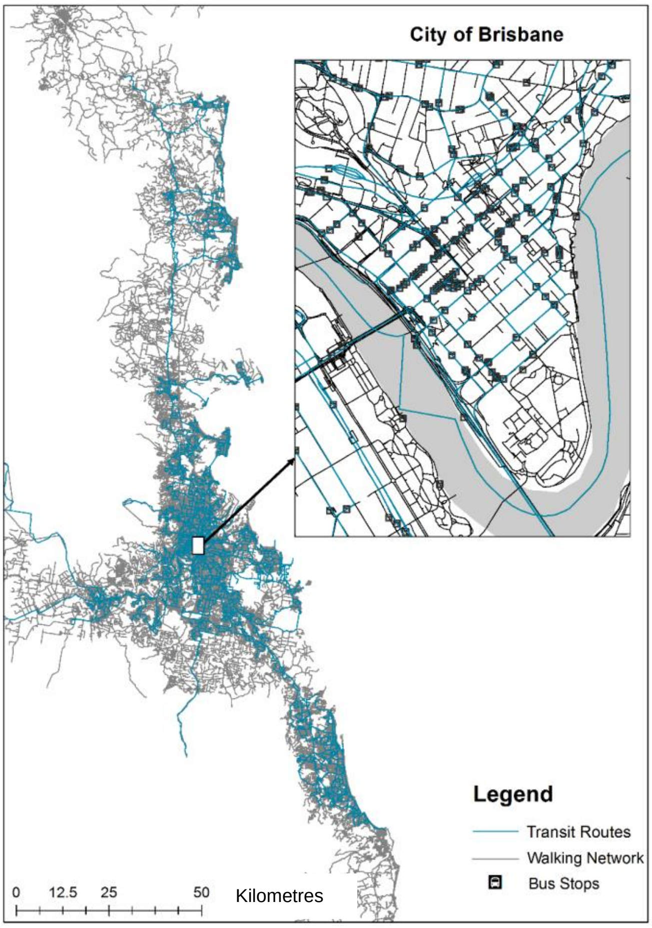
Table 1: Percentage of Transfers from Go Card Dataset for 5 March 2013

	Zero transfer journeys	**One transfer journeys**	**Two transfer journeys**	**Three transfer journeys**	**Four transfer journeys**	**Five transfer journeys**	**Total Journeys**
20 min criteria for one day	274,905	53,274	5,007	577	122	18	333,903
	82.22%	15.95 %	1.50 %	0.17 %	0.04%	0.01%	
Applying several criteria for one month (Nassir et al., 2014)	10,525,147	1,465,796	143,486	4,139	0	0	12,138,568
	86.71 %	12.08 %	1.18 %	0.03 %	0	0	

For finding the comfortable walking distance of Queenslanders, a detailed walking network including all footways was needed. OpenStreetMaps was used for this purpose. OpenStreetMap is an open licensed map of the world being created by volunteers using local knowledge, GPS tracks and donated sources, including a detailed multi-modal network. (www.openstreetmap.org). In comparison with other map sources in South East Queensland, this source was detailed and included all of walking features that could be used for transfers.

The final network was imported into ArcGIS, and after GIS preparation, snapping stops to the street network, and splitting up networks at intersections and transit stops, a walk network was built which includes about 340,000 links and 250,000 nodes in South East Queensland. This is shown in Figure 4. If there was not a detailed network map, this process could be simplified by using a simple Euclidean distance. However, because of complex street geometry and the serpentine nature of the Brisbane River, this Euclidean distance does not match reality.

Figure 4: Walking Network of SEQ



Finally, using the network analysis toolbox in ArcGIS, shortest path distances on this network for each multi-leg journey were calculated. Figure 5 shows the distribution of observed walking distance on the network for the transfers in the 5 March database, after aggregation of all smart card ID's and excluding the 5% top values (as outliers). Based on this finding, 95% of passengers in this case study have transfer walks under 500 metre. So, as discussed before, 500 metre was selected as a buffer around routes for extracting connectivity and coverage matrices.

The full process taken for transfer analysis is summarized in the diagram in Figure 6. The South East Queensland transit network includes train, ferry and bus routes. Coding of the two-step transfer rate model was done in Maple and was run on a Core i7-3.4 GHZ with 16GB RAM in 36 seconds.

Figure 5: Distribution of walking distance in observed multi-leg journeys

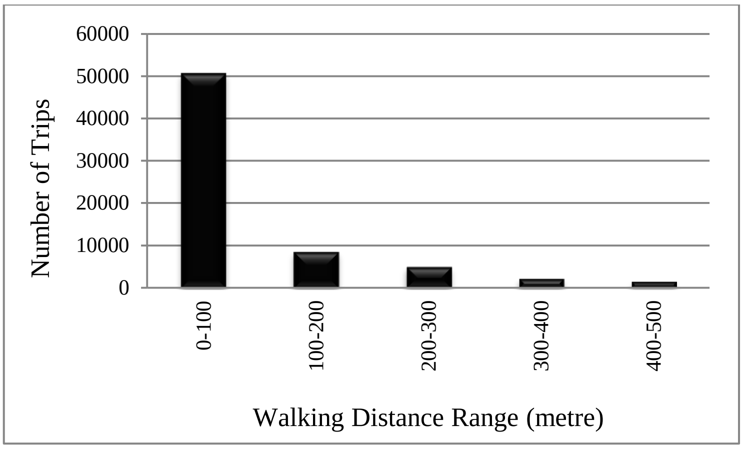
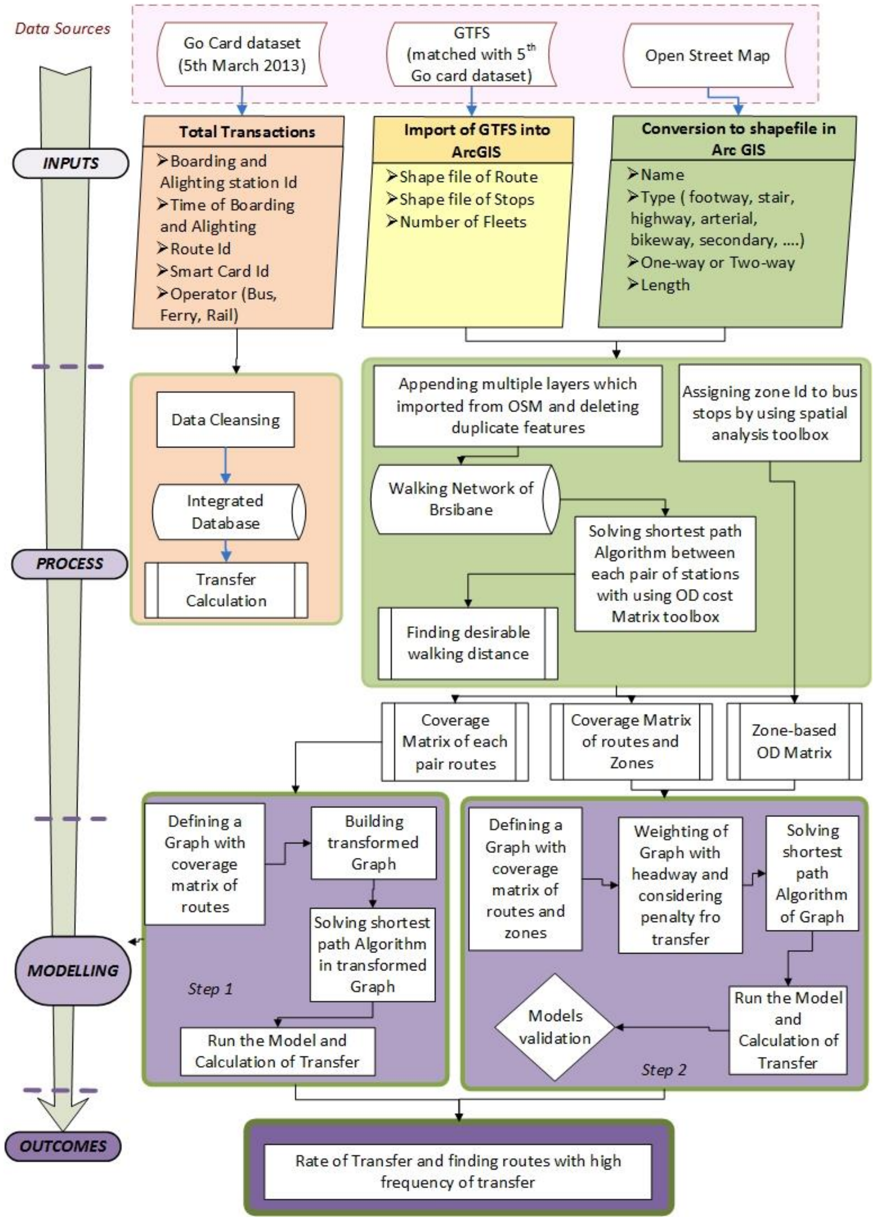


Figure 6: Overview of the research



4. Results

Step 1 of model was tested with the least number of transfers for each pair of routes. Only 1.8% or 189 from 10,296 observed values had a lower transfer rate than the result of the model, which may mean that they either had walked more than 500 metre or had used bike or automobile for access. **Error! Reference source not found.** shows the least transfer rate between each pair of transit routes, in which the red colour cells indicate a high number of transfers between those two lines that should be reviewed when the transit network is re-designed.

Figure 7: Least transfer rate between each pair of transit routes

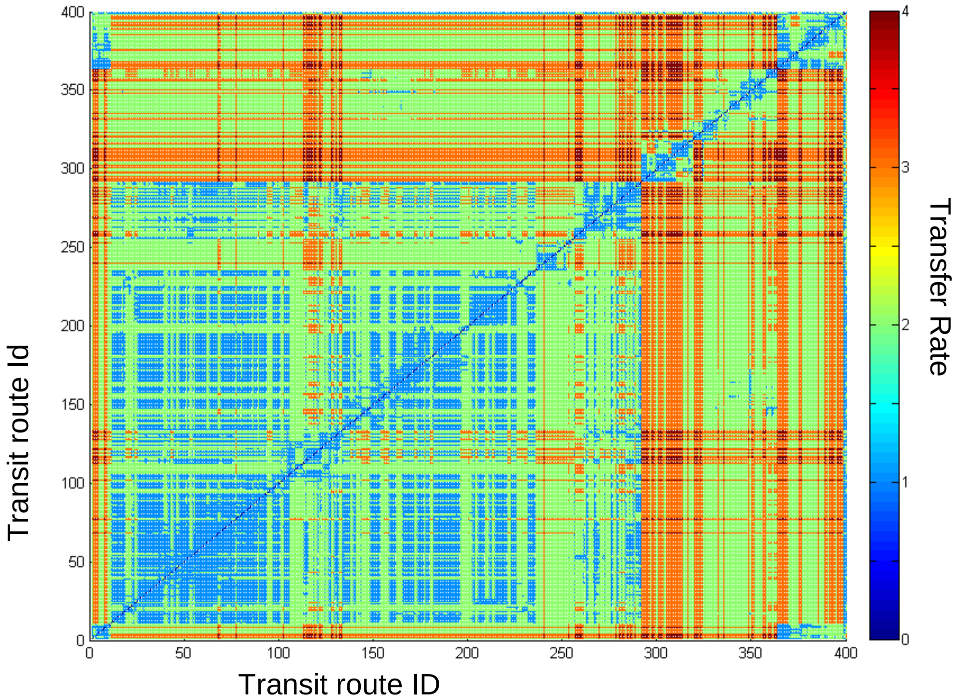


Table 2 represents a comparison of model results and observation, which shows good coincidence. It should be mentioned that the cost function of routes and the transfer weight is totally dependent on the fare structure and coverage of mass transit routes such as train or bus rapid transit (BRT). In this case study, passengers can transfer up to 3 times across all zones, and there is a one hour time limit between transfers without any cost. On the other hand, train covers many trips because the rail network includes multiple north-south and west-east routes that meet at key activity centers and major destinations and increase the percentage of one-leg journeys.

Table 2 Comparison of results of model and observation

Percentage	One leg journeys	One transfer journeys	Two transfer journeys	Three transfer journeys	Four transfer journeys
Model	81.28 %	15.44 %	2.53 %	0.74 %	0.01 %
Go card data	82.22%	15.95 %	1.50 %	0.17 %	0.04 %

5. Conclusion and Further Studies

Obviously, considering more detailed data in the cost function of routes and transfers, the model would give more accurate results. But, lacking these data, this model not only is simple to understand, to develop and to interpret, but it also is able in handling a spatially intensive and geographically extensive network with little computation time. The technique used in this research could answer these questions:

- In the current transit network, for each passenger, how many transfers are necessary to travel from one place to another place?
- How many trips, with what transfer rate, are made in the current public transport network?
- What are the high transfer lines and how could we improve these by designing a new direct line, an integrated station, or by eliminating the transfer cost between them?
- In a new network design, how would transfers change?

As indicated, the results of this research are very encouraging, and the method still can be improved. In this study, 59 zones were used, which was too big as results of detailed transfer rate between each pair of routes could not be compared directly. On the other hand, in the Go Card dataset, all of the transactions using the train routes were recorded as train without mentioning any route ID. Identifying a route ID of train and disaggregating zones would be further steps ahead that will extend the work.

6. Acknowledgements

This paper presents a part of the findings of research carried out in the ITS and Transportation Research Center of Amirkabir University between June 2011 and August 2012 and was developed with data provided by Translink and the Queensland Department of Transport and Main Roads (TMR), under the ASTRA agreement. The authors are grateful for data support and funding of TMR. Any omission or errors are the responsibility of the authors.

7. References

Alfred Chu, K. and R. Chapleau (2008). "Enriching Archived Smart Card Transaction Data for Transit Demand Modeling." *Transportation Research Record: Journal of the Transportation Research Board",* **2063**: 63-72.

Avishai Ceder, "Public Transit Planning and Operation", *Elsevier Ltd*, 2007.

Bagchi, M. and White, P.R. (2004). The Potential of Public Transport Smart Card Data. *Transport Policy*. **Vol. 12**, pp. 464-474.

Choi, K., and W.Jang.(2000) Development of a transit Network from a Street Map Database with Spatial Analysis and Dynamic Segmentation. *Transportation Research Part C*, **Vol. 8**, pp. 129-146.

Chriqui, C. and Robillard, P. (1975). Common bus lines. *Transportation Science*, **9**, 115–121.

De Cea, J. and Fernandez, J. E. (1993). Transit assignment for congested public transport systems: An equilibrium model. Transportation Science, **27**, 133–147.

Devillaine, F., M. Munizaga and M. Trépanier (2012). "Detection of Activities of Public Transport Users by Analyzing Smart Card Data." *Transportation Research Record: Journal of the Transportation Research Board,* **2276**: 48-55.

Hickman, M. and Bernstein, D. H. (1997). Transit service and path choice models in stochastic and time-dependent networks. *Transportation Science*, **31**, 129–146.

Dial, R. B. (1967). Transit pathfinder algorithm. *Highway Research Record*, **205**, 67–85.

Hadas, Y., & Ceder, A. A. (2010). Public transit network connectivity. *Transportation Research Record: Journal of the Transportation Research Board,* **2143**, 1-8.

Hofmann, M., & O'Mahony, M. (2005). Transfer Journey Identification and Analyses from Electronic Fare Collection Data. *IEEE Intelligent Transportation Systems 2005*, *Proceedings*, Vienna.

Munizaga, M. A. and C. Palma (2012). "Estimation of a Disaggregate Multimoda Public Transport Origin–Destination Matrix from Passive Smartcard Data from Santiago, Chile." *Transportation Research Part C: Emerging Technologies,* **24**: 9-18.

N. Koncz, J. Greenfield, and K. Mouskos. A strategy for solving static multiple-optimal-path transit network problems. *ASCE Journal of Transportation Engineering*, **122(3)**:218–225, 1996.

Nassir, N., A. Khani, S. Lee, H. Noh and M. Hickman (2011). "Transit Stop-Level Origin-Destination Estimation Through Use of Transit Schedule and Automated Data Collection System." *Transportation Research Record: Journal of the Transportation Research Board"*, **2263**: 140-150.

Nassir, N., M. Hickman and Z.L. Ma (2015). "Activity detection and transfer identification for public transit fare card data." *Journal of Transportation*, Springer US. 1-23.

Nuzzolo, A., Russo, F. and Crisalli, U. (2001). A doubly dynamic schedule-based assignment model for transit networks. Transportation Science, **35**, 268–285.

Liu, C. L. (2002). Best-path planning for public transportation systems. In *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on* (pp. 834-839). IEEE.

Liu, C. L., Pai, T. W., Chang, C. T., & Hsieh, C. M. (2001). Path-planning algorithms for public transportation systems. In *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE* (pp. 1061-1066). IEEE.

Nguyen, S. and Pallotino, S. (1988). Equilibrium traffic assignment for large-scale transit networks. *European Journal of Operational Research*, **37**, 176–186.

S.A.Baghloee, Avi Ceder. (2011). "Transit-Network Design Methodology for Actual-size Road Networks*", Transportation Research Part B.*

Seaborn, C. and Halcrow Group Ltd (2009). "Smart Card Data For Multi-Modal Network Planning in London Five Case Studies", *European Transport and Contrbutors*.

Seaborn, C., J. Attanucci and N. H. M. Wilson (2009). "Using Smart Card Fare Payment Data To Analyze Multi-Modal Public Transport Journeys in London", MIT Open Access Articles, Transportation Research Board of the National Academies, ISSN: 0361-1981.

Seaborn, C., J. Attanucci and N. H. M. Wilson (2009). "Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data." *Transportation Reseach Record (TRR)* **2121**: 55-62.

Spiess, H. and Florian, M. (1989). Optimal strategies: A new assignment model for transit networks. *Transportation Research Part B*, 83–102.

Stern, E. (1980). A cognitive scalar for the public transport connection matrix. *The Professional Geographer*, ***32*(3)**, 326-334.