

Monte Carlo Simulation for Transit Transfer Volumes: TRB data analysis competition

Gregory S. Macfarlane^{a,b,*}, Candace Brakewood^a, Jamie M. Fischer^a, Alex Poznanski^a

^a*School of Civil and Environmental Engineering, Georgia Institute of Technology
790 Atlantic Drive, Atlanta GA 30332-0355*

^b*School of Economics, Georgia Institute of Technology
221 Bobby Dodd Way, Atlanta, GA 30332*

Abstract

Transfers are an integral element of urban transit systems because they enable increased network coverage. Despite their necessity, passengers often dislike transfers due to various factors such as increased travel time, decreased reliability, and lack of infrastructure at transfer facilities, among others. In this paper, we aim to estimate the transfer volume at the hub in a small radial transit network. Our overall approach is based on the well-known gravity model from the urban transportation literature. The variables used in this analysis are grounded in literature pertaining to transit network design, transfer penalties, and friction factors in public transportation. These factors are analyzed using Monte Carlo simulation. The results of the analysis show a range of transfer volumes from XX to YY, depending on the random conditions.

Keywords: TRB data analysis competition, trip distribution, gravity model, transfer penalty, public transit

1. Introduction

The purpose of this project is to estimate the transfer volumes in a given transit network, shown in Figure # and Table #.

The approach uses an adapted gravity model, using randomly generated parameters. The gravity model is not the only alternative for trip distribution; alternatives include the entropy model (?). The gravity model was chosen for this analysis because it is simple and most commonly used.

*Corresponding author. Tel.: +1 801 616 9822

Email addresses: gregmacfarlane@gatech.edu (Gregory S. Macfarlane), candace.brakewood@gatech.edu (Candace Brakewood), jm.fischer@gatech.edu (Jamie M. Fischer), alex.poznanski@gatech.edu (Alex Poznanski)

2. Methodology

2.1. Adapted Gravity Model

The gravity model has been in existence for over 100 years, and it is one of the most common models for trip distributions used in urban transportation planning (Meyer and Miller, 2000). The general gravity model has the following formula:

$$T_{ij} = \frac{P_i [A_j f_{ij} k_{ij}]}{\sum_{n=1}^m A_j f_{ij} k_{ij}} \quad (1)$$

where T_{ij} are the estimated trips between discrete zones i and j , P_i are the trips produced at i , A_j are the trips attracted to zone j , f is a function of the travel disutility between i and j , and k_{ij} is a *post-hoc* adjustment factor.

Based on the literature pertaining to trip distribution and public transportation origin destination estimation, the f function in the basic gravity model depends on travel time and a friction factor. Therefore, an adapted gravity model is formulated for this analysis, as follows:

$$T_{ij} = \frac{P_i A_j t_{ij}^{-b}}{\sum_{j=1}^N A_j t_{ij}^{-b}} \quad (2)$$

In the adapted model, f is a function of the travel cost,¹

$$t_{ij} = \frac{d_{ij}}{s} + \lambda W_{ij} \quad (3)$$

where d_{ij} is the distance between i and j , s is the effective network average operating speed, λ is a wait multiplier to the transfer time W_{ij} ($W = 0$ if no transfer is required between i and j). This time is raised to the negative of a friction factor b ; specifying the friction factor in this way ensures that the marginal cost of a minute of travel diminishes as absolute travel time increases. In the adapted model, the adjustment factor k is replaced with an iterative process that incrementally balances the model's predicted attractions with the starting values.

Since the friction factor and the travel time between stations are not known for the given system, we generate them using assumptions based in the trip distribution literature, as discussed in the following sections.

For this analysis, the travel cost t is calculated primarily in terms of travel time. Travel time is calculated as an aggregate function of the distance between stations, the speed of the rail system, and the potential transfer wait time. The travel cost function is then modified based on the perceived disutility of transfers wait time as shown in the following equation:

[insert equation]

¹Travel "cost" is expressed in perceived travel minutes. Proper unit conversion factors are suppressed for simplicity.

Monetary and other potential travel costs are not included in this analysis, since they are unlikely to affect transfer behavior. The included parameters are generated through Monte Carlo simulations, as discussed in the sections below. Probability density functions of our assumed distributions are given in Figure 1.

2.2. Distance between Stations

Vuchic (2005) ranks the average stop spacing of urban metro systems in the world. Of these, Athens has the shortest average stop spacing of 595m (0.37 miles) and Mexico City has the longest of 1222m (0.76 miles) (?). The arithmetic mean of average stop spacing in urban networks is 0.594 miles.

For this analysis, a random station spacing is generated based on the empirical information collected by Vuchic (2005). This analysis assumes that stations are spaced according to a lognormal distribution with a mode of 0.6 miles, a mean of 0 and a variance of 1. The lognormal distribution was chosen in order to exclude negative distances.

2.2.1. Travel Speed

Speeds are normally distributed with a mean of 21.4461 and a standard deviation of 5.2957 miles per hour. These assumptions are built on values published in the National Transit Database (?). The NTD was queried for the annual train revenue miles (distance) and annual train revenue hours (time) in all heavy rail systems in the United States (?). The distance values were normalized by the time in order to calculate average speed for each heavy rail system. To determine a representative of all heavy rail systems, a weighted average was calculated based on revenue miles per system; this reduced the impact of small systems on the overall average.

2.3. Transfer Time

Transfer time depends on the timing of train arrivals on intersecting rail lines, which is in turn dependent upon train headways. Rail networks with short headways (less than 10 minutes) do not tend to coordinate arrival times on intersecting lines, since transfer time will always be short. Rail networks with longer headways (greater than 10 minutes) tend to coordinate their train arrival times in order to create convenient transfer times. (?) Therefore, this study assumes that transfer times vary between 0 and 10 minutes according to a truncated normal distribution, with a mean of 5 minutes and a standard deviation of 2 minutes.

2.4. Transfer Time Multiplier

Transit passengers perceive out-of-vehicle time to be more onerous than in-vehicle time. Reasons for this include out-of-vehicle transfers, at grade transfer, being exposed to the elements, and uncertainty, among others (Guo, 2008). According to the 2nd edition of Transit Capacity Quality Service Manual (TCQSM),

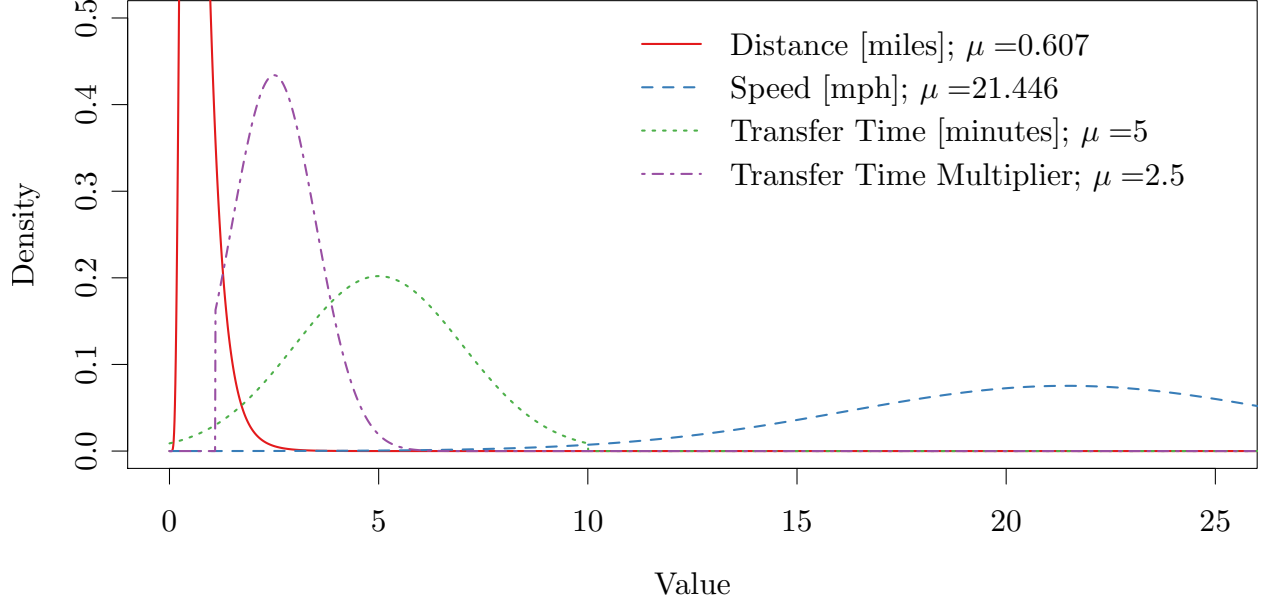


Figure 1: Probability density functions of random parameters used in the analysis.

passengers perceive one minute of wait time during transfers to be 2.5 times more onerous than one minute of in-vehicle travel time, on average, for work trips. The range for this transfer time multiplier is from 1.1 to 4.4 (?).

This analysis assumes a randomly distributed transfer time multiplier based on the details provided by the TCQSM: we randomly draw from a normal distribution with a mean of 2.5 and a standard deviation of 1.

3. Results

Given the distributions for the three variables discussed above, they were then combined into a gravity model and implemented in the open source software R. XXX draws were conducted for each probabilistic variable.

The final product of our analysis the shown in Figure 2 and Table XXX. As can be seen below, XX is the most likely number of transfers.

4. Discussion

We presented a probabilistic approach to determining the number of transfers from boardings and alightings in a given transit network. This approach is grounded in values found in practice for heavy rail station spacing, travel speeds, transfer time, and transfer penalty. 2 depicts the probability densities of the predicted

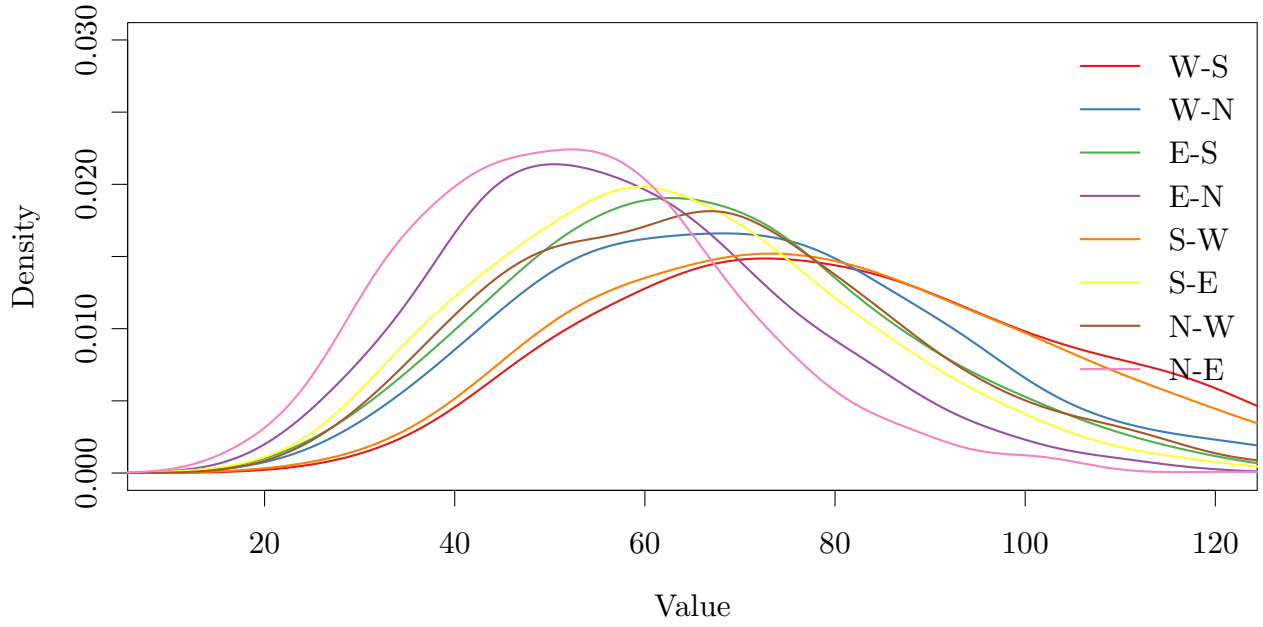


Figure 2: Probability densities of predicted transfer volume.

Table 1: Transfers by Direction

	W-S	W-N	E-S	E-N	S-W	S-E	N-W	N-E
Minimum	25.04	20.56	18.70	14.55	22.22	16.90	22.60	14.70
Maximum	183.21	163.87	150.70	121.34	181.54	148.44	144.95	121.51
Std. Dev.	25.78	22.60	20.36	18.18	25.45	19.63	21.26	16.65
Mean	82.41	70.86	66.37	57.75	80.10	63.26	66.87	52.59
25 th percentile	49.76	43.40	40.00	30.80	47.43	38.50	40.97	31.83
75 th percentile	158.48	141.03	129.40	105.09	156.34	126.84	126.58	104.37
Most Likely	72.60	68.30	62.80	50.49	73.02	59.70	67.08	52.14

transfer volume for each of the eight possible transfer situations. A fifty percent confidence interval was also presented, as the analysis was based on many assumptions due to the limited information provided.

While this overall approach utilizes a well-known model grounded in real-world values, there are noteworthy limitations to this approach. First, gravity models are often criticized for their lack of theoretical basis in a transportation planning context, as well as high degrees of error (Meyer and Miller, 2000). Another drawback is the assumed value of one for the adjustment factor, k , and the friction factor parameter, b . Because these parameters are calculated purely empirically, there is insufficient information to determine alternative values with any substance. Last, the travel cost function, which is based on travel time, was calculated using assumed values of station spacing and vehicle speed. In relatively, these values would exist for a given transit network and would therefore not need to be calculated using a probabilistic approach. Subsequently, a more robust analysis could be performed for an actual transit system using the methods presented with empirical data.

A word on execution

This project was executed as a training exercise on literate programming using R (R Development Core Team, 2011), `knitr` (Xie, 2012), and \LaTeX . The source code is available on GitHub as the `GT.TranspoComp` project.

References

- M. D. Meyer, E. J. Miller, Urban Transportation Planning: a decision-oriented approach, McGraw-Hill Science/Engineering/Math, 2000.
- Z. Guo, Transfers and Path Choice in Urban Public Transport Systems Transfers and Path Choice in Urban Public Transport Systems, Ph.D. thesis, Massachusetts Institute of Technology, 2008.
- R Development Core Team, R: A Language and Environment for Statistical Computing, URL <http://www.r-project.org>, 2011.
- Y. Xie, knitr: A general-purpose package for dynamic report generation in R, URL <http://yihui.name/knitr/>, 2012.