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. 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. We introduce a travel cost function comprised of variables grounded in the transit planning and operations literature. These variables are analyzed using Monte Carlo simulation. The results show that the most likely transfer volumes rane from 50 to 73 riders per direction per hour.

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 radial transit network with nine stations where hourly boradings and alightings are given. The approach uses a gravity model, which has been adapted using Monte Carlo simulation.

2. Methodology

2.1. Adapted Gravity Model

The gravity model was chosen for this analysis because it is simple and most commonly used among trip distribution methodologies (Meyer and Miller, 2000). The general gravity model has the following formula:

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

*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)

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:

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

In the adapted model, f has been replaced with a function of travel cost, t, raised to the negative of a friction factor b. Specifying the friction factor in this way ensures that the marginal utility of travel decreases non-linearly as travel time increases. For this analysis, a value of $b = 1$ is assumed for the idealized transit network, since no additional network characteristics are available to generate another value. 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.

2.2. Travel Cost Function

For this analysis, travel "cost" is expressed in perceived travel minutes, according to the following equation:

$$t_{ij} = \frac{d_{ij}}{s} + \lambda W_{ij} \tag{3}$$

where d_{ij} is the distance between i and j, s is the effective network average operating speed, W_{ij} is transfer time ($W = 0$ if no transfer is required between i and j), and λ is a transfer time multiplier.

Monetary and other potential travel costs are excluded from this analysis since they are unlikely to affect transfer behavior. Travel distance, operating speed, transfer time, and the transfer time multiplier 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.

2.3. Distance between Stations

Vuchic (2005) has tabulated 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 exclude negative distances.

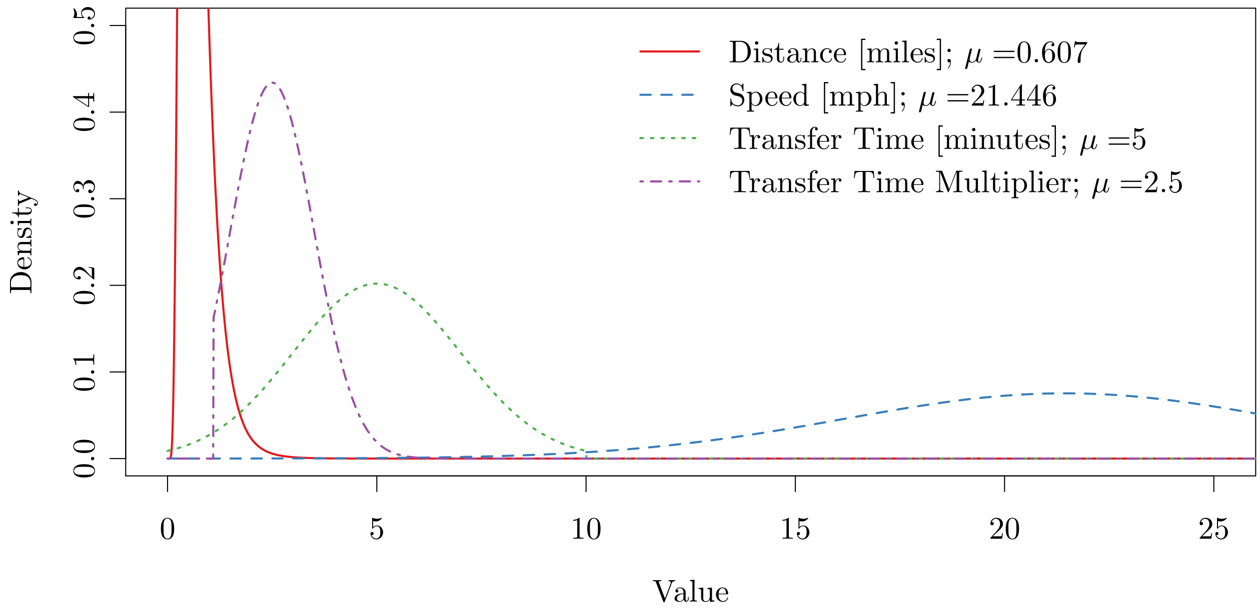

Figure 1: Probability density functions of random parameters used in the analysis.

2.3.1. Travel Speed

Speeds are assumed to be 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 by the Federal Transit Administration (**?**). The National Transit Database 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 number for all heavy rail systems, an average speed was calculated, weighted by the revenue miles per system in order to reduce the impact of small systems on the overall average.

2.4. 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 according to a truncated normal distribution, with a minimum of 0, maximum of 10, mean of 5 minutes and a standard deviation of 2 minutes.

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

2.5. Transfer Time Multiplier

Transit passengers perceive out-of-vehicle time to be more onerous than in-vehicle time. This modification to the travel cost is often called the "transfer penalty." According to Guo (2011), "the transfer penalty is affected by a broad range of factors, including safety and security, ease of way-finding during transfers, availability of escalators, weather protection, seating availability, lighting, air conditioning and ventilation, and concessions on the platforms." (Guo and Wilson, 2011) The 2nd edition of Transit Capacity Quality Service Manual (TCQSM) states that passengers perceive one minute of tranfer time 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

The adapted gravity model sumulation was run for 1000 draws, using the open source software R. The final results of this simulation are shown in Figure 2 and Table 1. The "most likely" values for each transfer direction, as shown in Table 1 refer to the most expected (peak) values shown in Figure 2. Because these are probabilisitic values, the 25th and 75th percentiles are also presented in Table 1, demonstrating a 50% confidence interval.

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

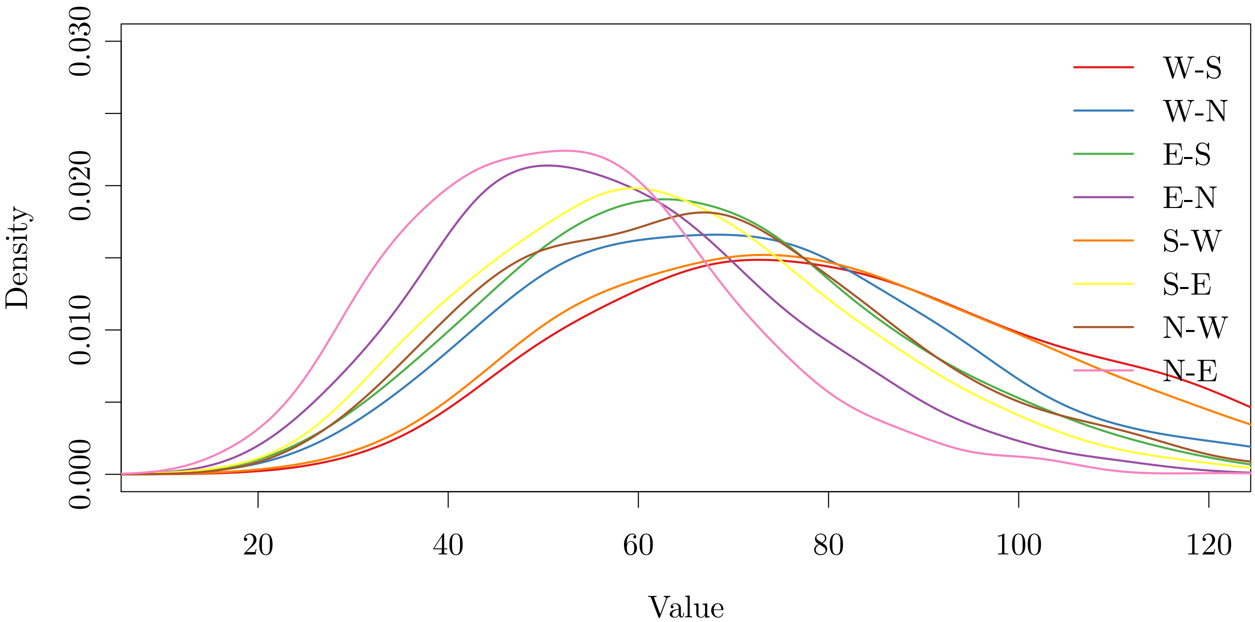

Figure 2: Probability densities of predicted transfer volume.

spacing, travel speeds, transfer time, and transfer penalty. 2 depicts the probability densities of the predicted 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, N. H. Wilson, Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground, Transportation Research Part A: Policy and Practice 45 (2) (2011) 91–104, ISSN 09658564.

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.