The Thesis Committee for Tyler James Beduhn
Certifies that this is the approved version of the following thesis:

Reliable Routing in Schedule-Based Transit Networks

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor: ______________________________

Stephen D. Boyles

____

Randy B. Machemehl

Reliable Routing in Schedule-Based Transit Networks

by

Tyler James Beduhn, BSE

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

December 2014

Acknowledgements

I would like to thank my advisor, Dr. Stephen Boyles, for his invaluable guidance, support and encouragement during my studies at The University of Texas at Austin. I would also like to extend by appreciation to Dr. Machemehl for his valuable feedback on this thesis. I also thank Alireza Khani, who I have enjoyed working with greatly, for spurring my interest in public transit research and the insightful discussion and brainstorming we had on related topics. I am also grateful to my friends and colleagues for their friendship and support during my time at UT, particularly Jake, Ankita, Amit and Sudesh. I am greatly indebted to my parents and family for their unbounded love and support throughout my entire life and education. Lastly, I would like to express my deepest appreciation to my fiancé, Audrey, for her endless support, patience, encouragement and sacrifice, and for never letting me forget the big picture.

Abstract

Reliable Routing in Schedule-Based Transit Networks

Tyler James Beduhn, MSE

The University of Texas at Austin, 2014

Supervisor: Stephen D. Boyles

A framework is proposed for determining the least expected cost path in a schedule-based time-expanded public transit network where travel times, and thus bus arrival and departure times at stops, are stochastic. Transfer reliability is incorporated in a label-correcting algorithm with a penalty function for the expected waiting time when transferring that reflects the likelihood of making a successful transfer. The algorithm is implemented in transit assignment on an Austin, Texas test network, using actual bus arrival and departure time distributions from vehicle location data. Assignment results are compared with those of a deterministic shortest path based on the schedule and from a calibrated transit assignment model. Simulations of the network and passenger paths are also conducted to evaluate the overall path reliability. The reliable shortest path algorithm is found to penalize transferring and provide paths with improved transfer and overall reliability. The proposed model is realistic, incorporating reliability measures from vehicle location data, and practical, given the efficient shortest path approach and application to transit assignment.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Background

Urban transportation systems are facing numerous challenges including increasing demand and congestion, limited funding and inadequate public transit facilities. The effects are felt by households and businesses alike, putting strain on economic and social development. Public transportation is a vital component of an urban transportation system. It has the capability of addressing many transport problems with more efficient use of limited right-of-way and increased vehicle occupancy, and it enhances personal mobility for a diverse group of the population. Public transit, however, has its own share of challenges. Service reliability continues to be a topic of growing interest and is regularly cited by users as an important quality of service measure. Reliability has significant implications on users' choice of transit routes, departure time, or even to use the mode at all. Uncertainty when using public transit can result from a variety of sources as outlined in Table 1 (TCQSM, 2013). Improved reliability can benefit passengers with more certain travel times and service operators with lower costs and increased ridership (Van Oort, 2011).

Transfers between routes is an integral part of transit service due to scattered origin and destination patterns and the high costs of supplying direct service between all areas of a city. Considering this, uncertainty in travel times becomes even more problematic as missing a transfer can result in costly delays. With the increasing adoption of automated data collection systems like automated vehicle location (AVL) and automated passenger counting (APC) systems that track vehicle travel times, schedule adherence and passenger activity, reliability measures can be easily obtained. These data sources and reliability measures are useful in evaluating operational performance and in developing more advanced planning models.

Table 1: Causes of unreliability in transit service.

Traffic condition	Differences in operator driving skills
Vehicle and maintenance quality	Wheelchair lift usage
Vehicle and staff availability	Route length and number of stops
Transit preferential treatments	Weather
Schedule achievability	Incidents and construction
Evenness of passenger demand	Operations control strategies

From a planning perspective, there is a desire for models that realistically capture user behavior and transportation system performance. Advanced transit assignment models, which assign passengers for given origin-destination (OD) pairs to a specific path, are useful in predicting the utilization of a public transit system. Route choice is a core component to these models and can take several forms including *a priori* routing, adaptive routing policies or utility maximization. Incorporating reliability into the transit routing problem adds more realism to assignment; for example, a transfer that should be made on the basis of the schedule is not always made in reality, and experienced passengers may choose their routes with their perception how the system operates. In short, an assignment tool that reflects the importance of reliability in a user's trip planning process can be a more realistic model and is useful in evaluating the impacts of service improvements.

1.2 MOTIVATION

Optimal routing in a transit network can be quite complex. There are several elements of transit network that make this challenging:

1. Routing involves waiting to board a vehicle at the origin and at transfer locations.
2. Transit service is time-dependent and often follows a schedule.
3. Passengers have the opportunity to transfer between routes.
4. Actual transit service can deviate from the schedule.

Given these challenges there is a need for a framework for determining an optimal reliable path in a schedule-based transit network that considers the stochastic nature of the service.

This thesis is motivated by the importance placed on reliability by transit users and the impacts it has on their route choice. While reliability can have many definitions, it can be generally thought of as the variability of a service attribute. In public transit, it is often associated with the timeliness of vehicles and the difference between a passenger's scheduled and actual travel time. Unreliability of transit can have a compounding effect on travel time with missed boardings leading to additional waiting costs to the user and potential disruptions in later planned segments of the trip. Balcombe et al. (2004) review studies that indicate that the 'excess' waiting time due to unreliable service has a much higher disutility to passengers than ordinary waiting time. This value is typically 2 to 3 times the valuation for normal waiting (Bly, 1976). Several others have emphasized the significant impacts unreliability at transfer points has on overall trip reliability and attractiveness of the mode (Turnquist and Bowman, 1980; Mai et al., 2012; Cedar et al., 2013).

As important as it is to recognize the value placed on a reliable transit service by users, it is equally important to incorporate such realizations into planning models. A realistic and practical transit model has many benefits. It can be the answer to questions such as "What will be the ridership of the new transit line? How will improved reliability benefit transferring passengers? How will passengers adjust their trip making to service changes? How much will ridership increase if the timeliness of a bus route is improved?" The usefulness of knowing the path which provides good reliability at transfer points is not limited to transit assignment modeling, though this is the main application in this thesis. It can also be useful for trip planning tools.

1.3 CONTRIBUTIONS

In this thesis a framework is presented for determining the *a priori* least expected cost path with reliable transfers in a stochastic schedule-based transit network. The term *reliable transfer* is used because the variability of transfer timing is considered with actual distributions of vehicle arrival and departure times at stops from AVL data. Unreliable transfers have a high probability of being missed, resulting in additional travel cost to the passenger. By minimizing the expected cost, unreliable transfers are penalized. The primary contributions are (1) modeling network stochasticity and its impact on transfers, (2) a computationally efficient and practical algorithm for finding the reliable path and (3) the use of empirical transit data for model development and testing. The algorithm is applied in passenger assignment, and network and passenger simulation is conducted to evaluate the paths of three assignment approaches. The results provide insight into the tradeoffs that exists relating to reliability.

1.4 Organization of Thesis

The thesis is organized as follows. Chapter 2 reviews related literature on routing in stochastic networks, routing with reliability and transit assignment. A background on the specific transit assignment model used for comparison and common public transit data sources is also included. Chapter 3 presents the proposed shortest path approach and solution algorithm. Numerical results of passenger assignment and simulation in the Austin, Texas transit network are included in Chapter 4, and concluding remarks are provided in Chapter 5.

Chapter 2: Literature Review and Background

2.1 INTRODUCTION

This chapter reviews relevant literature on routing in stochastic networks and on transit assignment, and provides a background on the tools and data sources used in the application of the developed routing framework. This research is based on the shortest path concept in transportation networks. Shortest path problems have long been considered in transportation applications for several reasons; they generally describe how users choose their travel routes, they can be adapted to optimize different objectives, and algorithms exist to efficiently solve them. The routing problem in a transportation network considers how an individual user would behave, specifically the path taken from a trip origin to the destination. Collectively knowing how people choose their routes leads to assignment, an important planning tool. The following sections review routing in a stochastic network, routing with reliability and transit routing. Literature that has considered transfer reliability in contexts other than assignment is also reviewed, along with the assignment model used in comparing the developed framework and useful transit data sources.

2.2 ROUTING IN A STOCHASTIC NETWORK

Optimal routing in transportation networks is a problem that has been studied from many approaches. The routing problem has many variants based on the nature of the network. One broad classification is static versus dynamic; often transportation networks are better viewed as dynamic as conditions (e.g. travel time) are dependent on time of day. Routing in dynamic networks, commonly referred to as time-varying or time-dependent, is also subject to uncertainty given that stochasticity in travel times arises from many causes. Stochastic routing can furthermore be classified into either *a priori* or *adaptive* problems. Hall (1986) was the first to study the shortest path problem in time-dependent networks with discrete stochastic link travel times and proved that link travel times cannot be simply replaced with their expected value at each time interval to solve an equivalent deterministic shortest path problem. Instead, a dynamic programming approach is used to determine an *a priori* least expected time (LET) path. Other related research includes a heuristic algorithm for a similar problem with continuous stochastic travel time proposed by Fu and Rilett (1998) and a modified label-correcting algorithm for generating LET paths and method for determining the lower bound on expected times of LET paths by Miller-Hooks and Mahmassani (2000).

The second class of stochastic shortest path problems involves an *adaptive* routing policy, also known as routing with recourse, where a user may update his or her route at any point. This problem has been considered in both a time invariant stochastic network (Waller and Ziliaskopoulos, 2002; Fan et al., 2005) and its time-dependent counterpart (Miller-Hooks and Mahmassani, 2000; Gao and Chabini, 2006). With the exception of a few studies (Waller and Ziliaskopoulos, 2002; Fan et al., 2005), the majority of the stochastic routing problems in literature assume link travel times (or costs) to be *independent* random variables (Fu and Rillete, 1998; Miller-Hooks and Mahmassani, 2000; Gao and Chabini, 2006; Frank, 1969; Fan et al. 2005). This thesis builds upon several of these past studies by determining the *a priori* least expected time path in a realistic public transit network, where the time the bus arrives/departs from a stop is an independent random variable. Thus, the reliability of transfers between bus routes is also inherently considered.

2.3 Reliability in Routing

Numerous researchers have incorporated reliability into the stochastic routing problem in one form or another. Frank (1969) defines an optimal path in a stochastic network as one which maximizes the probability of the travel time being less than a threshold, but the solution approach consists of an inefficient pairwise comparison of enumerated paths. An adaptive policy algorithm based on dynamic programming is used by Fan et al. (2005) in the problem of maximizing the probability of arriving on time in a static network; the convergence properties of the algorithm is investigated by Fan and Nie (2006). A corresponding *a priori* shortest path problem guaranteeing a given probability of arriving on-time while minimizing the time budget is developed by Nie and Wu (2009); the authors solve the problem with an exact label-correcting algorithm and extend the formulation to the time-dependent case. This problem is also considered with correlations in link travel times and solved with a simulation-based algorithm (Zockaie et al., 2013). Reliability has been incorporated in other forms as well. For example, Sivakumar and Batta (1994) introduce a variance constraint into the shortest path problem while Sen et al. (2001) use a linear combination of mean and variance in the objective function of the stochastic routing problem.

2.4 Transit Assignment

There have been several studies that consider transit service explicitly. Tong and Richardson (1984) develop algorithms for time-dependent (schedule-based) shortest path in a transit network based on either travel time or cost, however uncertainty in travel time is not considered. Adaptive routing strategy or hyperpath approaches have received greater attention in literature on transit networks. The passenger’s choice of bus line to board from an attractive subset of lines was first characterized by Chirque and Robillard (1975) in a probabilistic framework and was extended to strategies, hyperpaths and assignment in frequency-based transit networks (Spiess and Florian, 1989; Nguyen and Pallottino, 1988). Variations of schedule-based models which use more detailed arrival and departure time information for each vehicle have also been developed (Tong and Wong, 1999; Wilson and Nuzzolo, 2004; Hamdouch and Lawphongpanich, 2008).

Though the transit assignment problem has been studied from many different approaches over the last decades, very little consideration has been given to the impacts of reliability. Yang and Lam (2006) develop a probit-type reliability-based transit assignment model in which in-vehicle times are stochastic and assumed to follow a normal distribution, and the behavior of risk averse travelers is captured in a disutility function of travel time and its variation. However, their model relies on a simulation-based solution algorithm. Recently, reliability has also been incorporated into frequency-based assignment with capacity constraints (Szeto et al., 2011; 2013) and into schedule-based assignment with strategies using a mean variance approach (Hamdouch et al., 2014).

Aside from assignment, reliability in public transit networks has been considered at transfer locations in a scheduling, control and transit system performance measurement context (Knoppers and Muller, 1995; Muller and Furth, 2009; Lee et al., 2014). From a user perspective, transfers have a very high associated penalty due to their inconvenience (Guo and Wilson, 2011), and this penalty can range from 5 to 50 minutes of equivalent in-vehicle time depending on the mode and location (Currie, 2005). Since transfers depend on the arrival of two vehicles at a transfer point, they also undoubtedly contribute significantly to overall path unreliability. A framework is needed that links the time-dependent nature of the transit schedule, transfer timing and realistic performance characteristics (i.e. reliability) to passenger routing.

2.5 FAST-TrIPs Assignment Model

In Chapter 4, the shortest path framework developed in this thesis for determining the least expect cost path in a realistic transit network is applied to transit assignment. The passenger paths obtained from this assignment are compared with the deterministic shortest path based on the posted schedule and the paths from a previously developed assignment model. FAST-TrIPs[1] is a disaggregate passenger assignment and simulation model for schedule-based transit systems. It was developed as part of the SHRP 2 C10(B) project and has been tested in applications in Sacramento, CA, San Francisco, CA, Portland, OR and Austin, TX (Khani et al., 2013; 2014a; 2014b).

FAST-TrIPs models the transit network in a schedule based format, so each vehicle within a route is modeled separately according to the schedule. Several options are built into the model, one of which is the path choice model. Assignment can be done using either a deterministic shortest path (or least cost) or a stochastic multiple-path assignment. In the stochastic assignment a set attractive paths, or hyperpath, is generated for each passenger, who is stochastically assigned to an elementary path using utility values. The model requires a utility function, calibrated for the study area, that places weights on different components of a transit trip, for example in-vehicle time, waiting time, walking distance and number of transfers. Passengers are then assigned to a path in the choice set using the logit probability function

[1] FAST-TrIPs: Flexible Assignment and Simulation Tool for Transit and Intermodal Passengers

$$P(j)=\frac{e^{-\theta U_j}}{\sum_{i\in\Pi} e^{-\theta U_i}} \tag{1}$$

where P(j) is the probability of selecting path *j* among all paths in the attractive set *Π*, U_j is the utility of path *j* and θ is a dispersion parameter. The dispersion parameter reflects the sensitivity of users to cost differences across paths; a very large value may result in few alternative paths in the hyperpath whereas a small positive value results in more alternatives. Another feature of FAST-TrIPs is its ability to enforce vehicle capacity constraints. A simulation module captures the interaction between passengers and vehicles, and if a passenger is unable to board a crowded vehicle a penalty can be applied to its path. This could result in passengers adjusting their path choice with an iterative assignment.

Assignment using the stochastic path choice model of FAST-TrIPs is used as a comparison to assignment with the shortest path framework established in this thesis, using the bus system in Austin, Texas as the study area. In previous work, the logit-based route choice model for FAST-TrIPs was calibrated for transit passengers in Austin, TX (Khani et al., 2014b). The route choice model was estimated using data from an on-board survey conducted by the Capital Metropolitan Transit Authority (Capital Metro). Using reported origin, destination, boarding location, and bus route from the survey respondents, their observed path choices were inferred. The choice set of paths was then generated for each passenger using origin and destination locations and the approximate time of the intercept interview, and finally the logit model was estimated. FAST-TrIPs is a flexible tool that is calibrated to reflect user behavior and models individual vehicles based on the schedule, making it a good grounds for comparison to the methodology presented in this thesis.

2.6 TRANSIT DATA SOURCES

Many transit related data sources exist that are not only important to service operations, but can also be used in generating and testing planning models. Two sources in particular, General Transit Feed Specification (GTFS) files and automatic passenger counting/automatic vehicle location (APC/AVL) data, are critical to the implementation of the modeling framework in this work. In this section, a brief description of these data sources as well as how they are used is provided.

2.6.1 GTFS

Google's GTFS is a standard format for public transportation schedules published by transit agencies and made publicly available (Google Developers, 2012). GTFS is series of text files that describe the transit service in a trip-based format; the files are linked together with common attributes (see Figure 1). Six files are required to publicly post the feed, while seven additional files are optional. A description of each of the required files is provided in Table 2. In general, each route is made up of unique vehicle trips, and the scheduled arrival and departure time of each trip is listed for each stop it visits. The exact location of stops can be referenced in a separate file, and a calendar file specifies which days of the week a trip is in service. Though not required, a shape file can provide the geometry of each trip, and fares can be associated with each route.

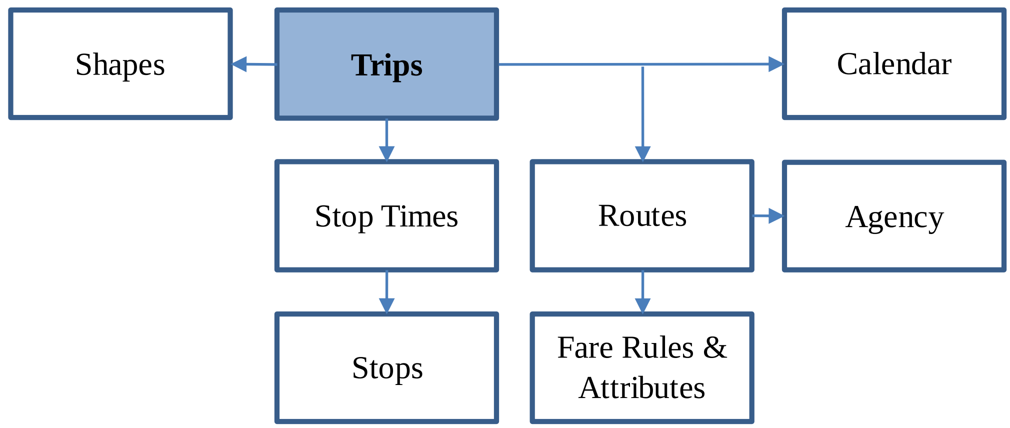

Figure 1: GTFS file relationships.

While the primary use of GTFS is for trip planner and time table publishing applications, it is also a powerful data source as the network input to planning models. It is an appropriate input to models with high temporal resolution, given its representation of individual vehicle trips. For example, it is used in generating many of the input files for the transit assignment model, FAST-TrIPs, described in the previous section. GTFS published by Capital Metro is used to generate the Austin transit network used for testing in Chapter 4.

Table 2: Required GTFS files.

File Name	Description
Trips.txt	List of all vehicle trips and their stops (service provided for a sequence of two more stops at a specific time)
Stop_times.txt	List of times when a vehicle (trip) arrives and departs from individual stops (schedule)
Stops.txt	Locations where vehicles pick up or drop off passengers (latitude and longitude)
Routes.txt	Transit routes (a group of trips provided as a single service to passengers)
Calendar.txt	Services IDs associated with when days of the week where service is available and the dates service starts and ends
Agency.txt	Information on the transit agency(ies) providing the data feed

2.6.2 APC/AVL

Automated data collection (ADC) systems are being widely adopted by transit agencies to provide both real-time and offline data. Two types of ADC systems are automated passenger counting and automated vehicle location systems. These systems are capable of capturing and storing enormous amounts of temporal and spatial data of different types that can be used to characterize a transit system's utilization and performance. AVL systems consist of GPS receivers positioned on vehicles which, when paired with other onboard sensors, provide a full picture of a route's spatial and temporal performance. In most AVL systems, a central computer does round-robin polling to identify bus locations in real-time, but additional "time at location" records can be created at designated stops or time points. Automatic passenger counting (APC) systems can use a variety of technologies for counting passengers; this includes pressure-sensitive mats and horizontal or overhead infrared sensing (Furth et al. 2006). Typically APC systems are installed on a fraction in a fleet due to cost, and buses equipped with the sensors are rotated around routes so that data on all routes can be collected. Capital Metro, for instance, has equipment installed on approximately 22% of their buses. The accuracy of the counting can fluctuate on the technology used and the algorithms used to convert sensing into passenger counts. In addition to passenger counts, the system usually includes location measurement and stop matching. When the bus leaves a stop an on-board computer creates a record with its on-off counts.

The data set used in this study is APC data with both spatial and temporal information from Capital Metro. The data set contains observed arrival and departure times at stops for a sample of vehicle trips during January – June 2013. The primary use of this data is to obtain measures of reliability, or timeliness, for bus routes in the PM peak period. During data processing each observed vehicle trip was matched to a GTFS trip ID to obtain the scheduled time and thus schedule deviation using the observed arrival and departure times for each stop in the trip. For each route and direction (e.g. northbound), data from all trips within the PM period was aggregated to get a mean and standard deviation for both arrival and departure time deviation from the schedule for each stop. This aggregation was used to ensure a large enough sample size at each stop as not every scheduled vehicle trip had a large number of observations during the data collection period. The final sample size across routes and stops ranges from 37 to 368 observations. The result is a mean and standard deviation of schedule deviation for both arrival time and departure time for each stop in each route and direction, reflective of service in the PM period. These measures are an integral part to the methodology described in the following chapter.

APC data can also reveal travel patterns that are needed when preparing disaggregate passenger demand input to a transit assignment model. Planning agencies typically have aggregate transit demand during various time periods for use in their models, for example, person trips demanded between two traffic analysis zones in the AM peak period. However when modeling passenger trips at a finer temporal resolution such as in the application of the routing framework of this thesis, demand needs to be disaggregate with passengers assigned a preferred arrival time (PAT)/preferred departure time (PDT). In order to disaggregate demand, a distribution of passenger boarding and alighting based on APC data can be used. Using the APC data from Capital Metro, this distribution has been estimated for Austin transit users (Figure 2). The distribution has been adjusted to consider that some passengers' trips included transfers, using an average number of boardings per trip of 1.25 and average unlinked trip time of 30 minutes (estimated from on-board survey and APC). The profiles show the expected peaks in the AM and PM with the lag between the boardings and alightings corresponding to in-vehicle time. Applying these profiles to an aggregate transit demand assigns a PAT (from alighting profile) or PDT (from departure profile) that is reflective of actual passenger behavior in Austin. The resulting disaggregate demand is used in the transit assignment applications in Chapter 4. Though not utilized in this thesis, the ridership estimates from APC can also be used to validate the outputs of transit assignment models.

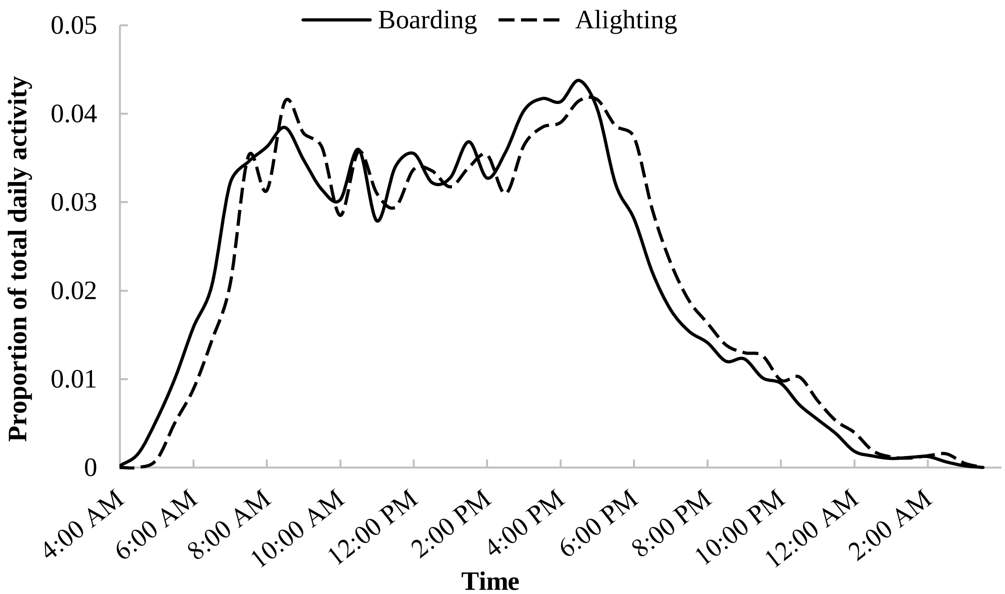

Figure 2: Distribution of boarding and alighting passengers based on APC

Chapter 3: Methodology

In this chapter, the framework for determining the optimal shortest path in a stochastic public transit network is provided. The transit network is represented with a time expanded graph with in-vehicle, walking and waiting links. The problem formulation, including the transfer waiting time model, reliable shortest path model and assumptions, is presented along with a label-correcting solution algorithm. Finally, an exercise to evaluate the primary assumption on the distribution of bus schedule deviations is discussed. The notation used in the problem is presented in Table 3 and is discussed in further detail throughout the following sections.

3.1 TIME EXPANDED NETWORK

In order to model the time-dependent schedule of the transit system, the network is represented in a time-expanded fashion that contains all possible movements passengers can take to their destinations at any point in time. A transit route is defined as a fixed set of stops that a vehicle visits to serve passengers; a route consists of individual vehicle trips that start at different times, and the time between successive trips is the headway. Throughout this thesis the problem is approached from the bus mode, though the methodology is generalizable to other fixed-route transit submodes such as rail. The network consists of nodes and links, represented by sets N and A respectively. A subset of nodes, N_V, represents the scheduled stop times of individual vehicles at each bus stop. These nodes are expanded further to represent separate arrival and departure times; subsets $N_{V\text{-}A}$ and $N_{V\text{-}D}$ ($N_{V\text{-}A} \cup N_{V\text{-}D} = N_V$) represent the scheduled arrival and departure times for the stop, respectively. Additionally, a subset of nodes, N_W ($N_V \cup N_W = N$),

Table 3: List of notation.

Notation	Description
N_V	Set of vehicle nodes representing the scheduled stop time at a stop; subsets $N_{V\text{-}A}$ and $N_{V\text{-}D}$ represent scheduled arrival and departure times, respectively
N_W	Set of walking arrival nodes representing the time passenger arrives at a physical stop after walking from another node
A_V	Set of in-vehicle links
A_W	Set of walking links
A_T	Set of waiting links (initial or transfer waiting time)
$\hat{t}_i$	Scheduled time associated with node $i \in N$, either the schedule arrival time if $i \in N_{V\text{-}A}$, the schedule departure time if $i \in N_{V\text{-}D}$ or the arrival time after walking if $i \in N_W$
$\tilde{t}_i$	Random variable of time associated with node $i \in N$
Y_{ij}	Random variable of the difference in time between two nodes i and j, equivalently $\tilde{t}_j - \tilde{t}_i$
VT_{ij}	Equivalent to Y_{ij} if link $(i, j) \in A_V$; in-vehicle time between nodes i and j
WT_{ij}	Equivalent to Y_{ij} if link $(i, j) \in A_W$; walking time between nodes i and j
TT_{ij}	Equivalent to Y_{ij} if link $(i, j) \in A_T$; (transfer) waiting time between nodes i and j
P_{ij}	Probability of a successful transfer from node i to node j
H_j	Headway of vehicle serving node $j \in N_{V\text{-}D}$
L_j	Subset of $N_{V\text{-}D}$; all nodes in the same route as, at the same physical stop as and scheduled later in time than node $j \in N_{V\text{-}D}$ representing subsequent potential connections if the transfer to j is missed
r	Origin node
s	Destination node
X_{ij}	Binary decision variable indicating if link (i, j) is on the optimal path from origin to destination
μ_i	Mean of $\tilde{t}_i$
σ_i	Standard deviation of $\tilde{t}_i$
$\Phi(\cdot)$	Standard normal cumulative distribution function
$E[\cdot]$	Expected value
$P(\cdot)$	Probability

represents a passenger's arrival time at a bus stop after walking from another node (i.e. from the origin or another bus stop if transferring). In the small network shown in Figure 3, node 9 is in set N_W, representing arrival at Stop B after walking from node 8, while all others are in N_V. Each node $i \in N$ has an associated time, $\hat{t}_i$, either the scheduled arrival or departure time of the bus if $i \in N_V$ or the arrival time at a stop after a deterministic walking time from a previous vehicle node at a different stop if $i \in N_W$. Nodes in the set N_V are unique to bus trips, that is, two buses scheduled to serve the same stop at the same time are modeled as two separate pairs of nodes. For the first stop in a bus route only a departure vehicle node ($N_{V\text{-}D}$) is needed, while only an arrival vehicle nodes ($N_{V\text{-}A}$) is needed for the last stop in a route.

Links include in-vehicle (A_V), walking (A_W) and waiting (A_T) ($A_V \cup A_W \cup A_T = A$). In-vehicle links connect each vehicle node associated with a given bus trip. The in-vehicle link connecting a given arrival time node with a corresponding departure time node is considered as dwell time. Walking links are created from each node in $N_{V\text{-}A}$ with an incoming in-vehicle link (i.e. not the start of a route) to corresponding nodes in N_W at nearby stops within a walking distance threshold. Waiting links are created from each node in N_W and $N_{V\text{-}A}$ to each other node in $N_{V\text{-}D}$ at the same bus stop that have a scheduled time of at most 30 minutes later, representing a transfer between vehicles or waiting to initially board a vehicle. Links to nodes scheduled earlier in time are not considered; while such transfers may be possible depending on the distributions of stop times, it is argued that they would be viewed as unreliable by the passenger. This constrained time window keeps the network size manageable, avoids the need to consider very unattractive transfers and should be chosen with typical bus headways in mind so that reasonable transfers are considered. Table 4 summarizes the allowable adjacent link types and quantity (if restricted) for each node type. The network can loosely be thought of as a space-time graph where the horizontal axis is space and the vertical axis is time. In this manner, a collection of vertical nodes represents a physical bus stop (see

Figure 3). This network representation explicitly represents all possible passenger movements, while prohibiting unrealistic movements such as transferring twice in a row (i.e. use of two consecutive waiting links) or walking after waiting to transfer (i.e. use of a walking link after a waiting link).

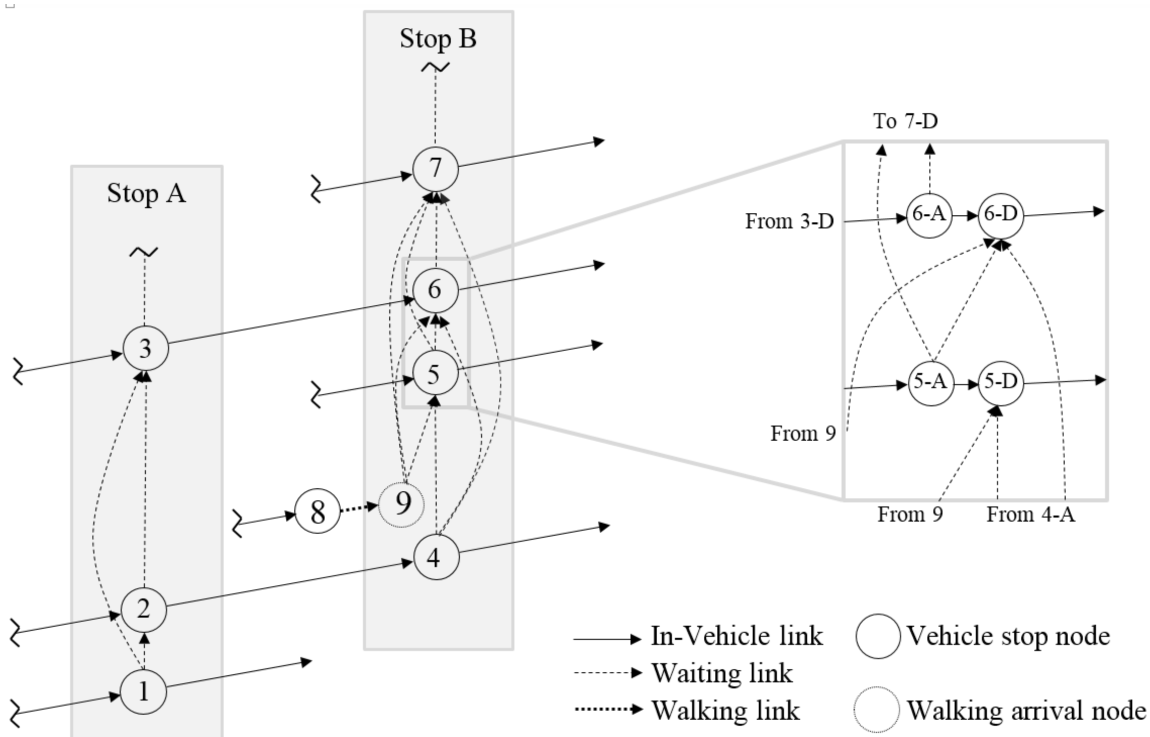

Figure 3: Sample network representation.

Table 4: Allowable adjacent links.

	Incoming Links	Outgoing Links
$N_{V\text{-}A}$	A_V (1)	A_V (1*), A_W, A_T
$N_{V\text{-}D}$	A_V (1**), A_T	A_V (1)
N_W	A_W (1)	A_T

*Zero if last stop in route, **Zero if first stop in route

3.2 Problem Formulation and Assumptions

3.2.1 Transfer Waiting Time Model

A transfer between routes depends both on passengers' arrival time at the boarding stop of the connecting route (either by another vehicle at the stop or by walking from another stop) and the departure time of the connecting vehicle. The actual arrival and departure times of vehicles are considered random variables, thus incorporating transit supply uncertainty. The notation used for the random variable of arrival or departure time of a vehicle, or arrival time of a walking passenger in the case of nodes in N_W, at a node i is $\tilde{t}_i$. The distributions of these random variables can be easily obtained from AVL data collected by transit operators. Generally, let $Y_{ij} = \tilde{t}_j - \tilde{t}_i$ be the random variable of the difference in time of two nodes i and j. Depending on the type of link connecting the two nodes, Y_{ij} is equivalently denoted as VT_{ij}, WT_{ij} or TT_{ij} if link (i, j) is in sets A_V, A_W or A_T, respectively. For in-vehicle links VT_{ij} is the travel time or dwell time of a bus between two nodes; similarly WT_{ij} is walking time. In the case of waiting links, TT_{ij} is the transfer waiting time. A transfer from i to j can be made successfully if $Y_{ij} \geq 0$, so the probability of making and missing the transfer is $P(Y_{ij} \geq 0)$ and $P(Y_{ij} < 0)$, respectively. Thus the expected waiting time is formulated in Equation 2 (reduced to

Equation 3) where H_j is the headway of the route serving node *j*. This formulation is made on the assumption that passengers attempt to board the next vehicle in the same route if a transfer to that route is missed. If the transfer is made, the expected waiting time is the expected time between service of nodes *i* and *j* given that this value is nonnegative (i.e. the transfer can be made). If the transfer is missed, the expected waiting time is the initial expected time between service of nodes *i* and *j* *and* the expected headway of the route. This additional wait time for a missed transfer serves as a penalty for unreliable transfers.

$$E[TT_{ij}] = P(Y_{ij} \geq 0)\, E[Y_{ij} \mid Y_{ij} \geq 0] + P(Y_{ij} < 0)[E[H_j] + E[Y_{ij} \mid Y_{ij} < 0] \tag{2}$$

$$E[TT_{ij}] = E[Y_{ij}] + P(Y_{ij} < 0)\, E[H_j] \tag{3}$$

Figure 4 shows example stop time distributions from node 1-A to node 2-D in the network in Figure 3. The first bus is scheduled to arrive at Stop A at $\hat{t}_1 = 8$ while the second bus is scheduled to depart at $\hat{t}_2 = 14$. However the actual stop times follow a distribution due to stochasticity in transit supply; there is a probability that Bus 1 will arrive after Bus 2 departs, and thus the transfer will be missed.

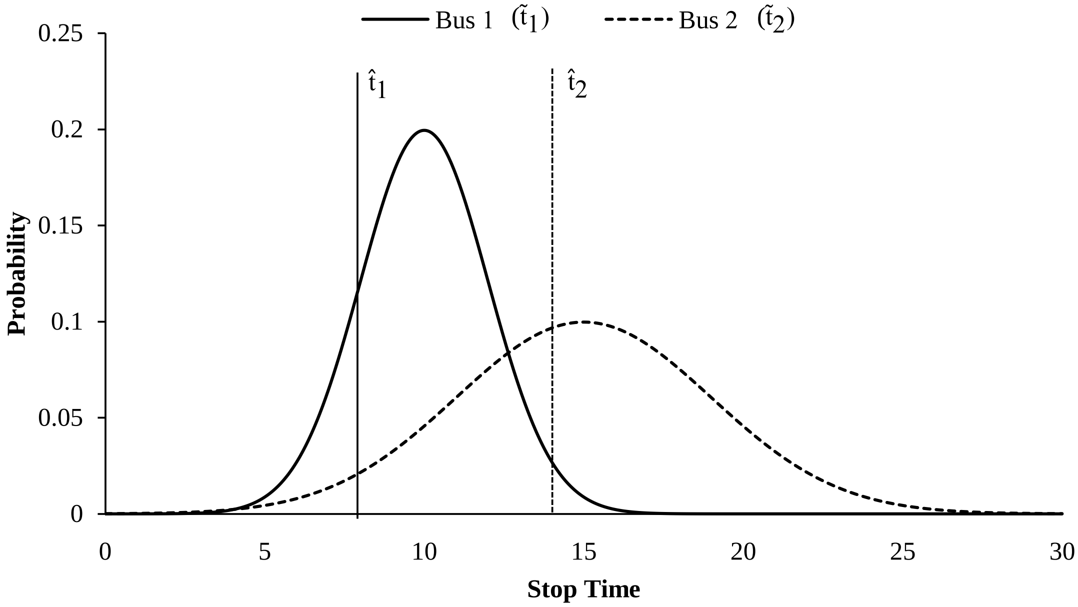

Figure 4: Example of stop time distributions at a transfer point.

3.2.2 Reliable Shortest Path Model

The objective of the routing problem in this paper is to determine the least expected cost path from an origin to a destination. A passenger's path is defined as a single, connected sequence of nonrepeating links in the underlying scheduled network that a passenger aims to take from an origin, r, to a destination, s. From a real-world perspective this is an *a priori* path that a passenger would decide on prior to the start of a journey and not necessarily the specific path the passenger would traverse due to realized instances of random vehicle stop times. The model is formulated in Equations 4-6 where X_{ij} is a binary decision variable indicating if link (i, j) is on the least expected cost path. The objective function (Equation 4) is the sum of the expected costs of used links by link type. Equation 5 is the conservation constraint, that is, for every node that is not the origin or destination, exactly one incoming and one outgoing link can be on the shortest path or else the node is not used.

$$\min Z = \sum_{(i,j)\in A_V} E[VT_{ij}]X_{ij} + \sum_{(i,j)\in A_W} E[WT_{ij}]X_{ij} + \sum_{(i,j)\in A_T} E[TT_{ij}]X_{ij} \tag{4}$$

$$\text{s.t.} \quad \sum_{j|\,(i,j)\in A} X_{ij} - \sum_{k|\,(k,i)\in A} X_{ki} = \begin{cases} 1 & \forall\, i = r \\ 0 & \forall\, i \neq r,s \\ -1 & \forall\, i = s \end{cases} \tag{5}$$

$$X_{ij} \in \{0,1\} \tag{6}$$

While transfer reliability is not an explicit constraint in the model, it is implicitly considered in the calculation of $E[TT_{ij}]$ (recall Equations 2-3). The model does not enforce that only reliable transfers are used but it does penalize unreliable transfers with the additional waiting time the passenger should expect if the transfer is missed. It is possible for an optimal path to include a transfer with high probability of being missed, perhaps if the route is high frequency and the passenger would only have to wait a few minutes to board the next vehicle. In other cases, there may be simply no other options and an unreliable transfer is required.

3.2.3 Assumptions

The routing model is based on the following assumptions:

Assumption 1: Bus stop times of different vehicles are independent random variables that are normally distributed.

Assumption 2: Passengers behave such that if a transfer is missed, they wait to board the next bus in the same route.

Assumption 3: Headways are sufficiently long so that vehicles within the same route do not overtake one another

The first assumption is a reasonable approximation that has been used in the literature (Knoppers and Muller, 1995; Muller and Furth, 2009). While the underlying vehicle stop time distributions may not be exactly normally distributed, this distribution generally captures the idea that transit vehicles operate around a schedule but are usually ahead or behind. It is also a reasonable estimate of how passengers interpret recurring uncertainty in transit service. It is hypothesized that the exact shape of the distribution will not have a substantial impact on the optimal path and that a normal distribution is reasonable approximation; this assumption is evaluated at the end of this chapter. This first assumption allows the use of closed form expressions when calculating expected transfer waiting times. The probability distribution of Y_{ij} is the convolution of the individual distributions of the arrival time of node i and departure time of node j, thus $Y_{ij} \sim N(\mu_{ij} = \mu_j - \mu_i, \sigma_{ij}^2 = \sigma_j^2 + \sigma_i^2)$ where μ and σ is the mean and standard deviation, respectively, of a bus's arrival or departure time at a node. The probability of making and missing a transfer is then given in Equations 7-8 where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

$$\text{P(make)} = P_{ij} = P(Y_{ij} \geq 0) = 1 - \Phi(-\mu_{ij}/\sigma_{ij}) \tag{7}$$

$$\text{P(miss)} = \overline{P}_{ij} = P(Y_{ij} < 0) = \Phi(-\mu_{ij}/\sigma_{ij}) \tag{8}$$

The second and third assumptions are needed to penalize for the risk of missing a transfer. When passengers behave according to Assumption 2, the expected headway of the connecting vehicle serves as a penalty (see Equation 3). The expected headway, $E[H_j]$, of the route serving node j is calculated by considering all later nodes in $N_{V\text{-}D}$ at the same stop that have an outgoing vehicle link in the same route as j. Let L_j be this set of nodes that share the stop and route of node j but have a scheduled stop time later than that of j ($|L_j| = n$), and let l index this set. The expected headway is then defined by Equation 9. This is interpreted as the expected headway that the *user* experiences.

$$E[H_j] = \sum_{l=1}^{n} P_{iL_j(l)} \left[\prod_{l'<l} \overline{P}_{iL_j(l')}\right] E[Y_{jL_j(l)}] \qquad (9)$$

The expected headway is the sum over all nodes in L_j of the probability that the transfer to a node in L_j from j is made and all other prior transfers are missed multiplied by the expected difference in time from j to the connecting node. Alternatively the expected headway can solely be the expected time until the next bus in the route; however this implies that the transfer to the next bus is guaranteed to be made. This may not always be possible, for example, for a high frequency route.

Together Equations 3, 7, 8 and 9, with $E[Y_{ij}] = \overline{Y}_{ij} = \mu_{ij}$, determine the expected waiting time at a transfer point under the stated assumptions while penalizing for the possibility of missing the transfer. The following subsection presents a solution algorithm proposed to solve the routing problem in Equations 4-6.

3.3 Reliable Shortest Path Algorithm

The solution approach proposed is a modified label-correcting algorithm for determining the *a priori* least expected cost path in a realistic, time-expanded transit network, denoted Reliable Shortest Path (RSP). A one-to-all labeling is implemented, that is, the optimal path to all other nodes is found from a single origin. The commonality between shortest path algorithms is the concept of distance labels. At any point in an algorithm, a distance label is associated with each node in the network that represents the distance (or time, cost, etc.) from the origin node to that node on a given path. In a modified label-correcting algorithm a scan eligible list (SEL) of nodes is maintained; the list represents nodes with outgoing links that, if used, might decrease the label of downstream nodes. Nodes in the list are removed one by one and outgoing links scanned and considered for updating the tail nodes. The node labels are an upper bound on the shortest path distance, and at termination these labels are the shortest path distance. More details on shortest-path algorithms can be found in Ahuja et al. (1993).

In the proposed RSP algorithm (see Figure 5), two labels are maintained for each node, a time label and a cost label, as well as the predecessor node that is used if a label is updated. The time label reflects the time incurred to reach a node from the origin based on the schedule. The cost label considers the actual distributions of vehicles arrival and departure times and transfer reliability and is based on the expected value of travel time. For example, the cost label of node j is determined based on the expected value of the random variable Y_{ij} associated with the link from the previous node i. The cost label need not be interpreted as a time at specific location in the network but instead a measure of the reliability of a path when compared against other labels. A SEL is used, and if the cost label of a node can be improved then the node is added to the list because labels on adjacent nodes can potentially also be improved. When a node is removed from the SEL the tail node of each outgoing link is scanned and considered for updating. When a label is improved the predecessor node used is also noted; this ensures the continuity in a path from the origin (i.e. Equation 5 constraint holds). The algorithm terminates when the SEL is empty, and the optimal path is revealed by following the predecessor pointers from the destination node back to the origin. The following notation is used in the algorithm, in addition to that already defined,

t_i : time label of node i

c_i : cost label of node i

c_i' : temporary cost label of node i

$L_i = [t_i, c_i]$: label of node i

q_i : predecessor node used on shortest path to node i from the origin

As with many shortest path algorithms, the proposed algorithm relies on Bellman's Principle of Optimality (Bellman, 1958). Simply stated, the shortest path can be found by breaking the problem into smaller subproblems (e.g. scanning one node at a time from the SEL) because every subpath in the shortest path from the origin to the destination is itself a shortest path. The RSP algorithm utilizes the following optimality conditions ensuring the cost labels represent the least expected cost from the origin:

$$c_j \leq c_i + E[VT_{ij}] \ \forall \ (i, j) \in A_V \tag{10}$$

$$c_j \leq c_i + E[WT_{ij}] \ \forall \ (i, j) \in A_W \tag{11}$$

$$c_j \leq c_i + E[TT_{ij}] \ \forall \ (i, j) \in A_T \tag{12}$$

They state that the for every link (i, j) the least expected cost path to j is no greater than the least expected cost path to i plus the expected cost of using link (i, j). Under the stated assumptions, $E[VT_{ij}]$ and $E[WT_{ij}]$ is simply the difference in the mean times of nodes i and j, or $\bar{Y}_{ij}$, and $E[TT_{ij}]$ is calculated with Equation 3.

```
1   algorithm Reliable Shortest Path
2   begin
3       L_r ← [0, 0] and q_r ← 0;
4       L_i ← [∞, ∞] and q_i ← Ø for each node i ≠ r;
5       SEL = {r};
6       while SEL ≠ Ø do
7           Remove the first element from SEL, i ← SEL(1);
8           for each link (i, j) ∈ A do
9               \bar{Y}_ij = μ_j - μ_i;
10              σ_ij = sqrt(σ_j^2 + σ_i^2);
11              if (i, j) ∈ A_T then
12                  t_j = t_i + (\hat{t}_j – \hat{t}_j);
13                  c_j' = c_i + E[TT_ij];
14              else
15                  t_j = t_i + (\hat{t}_j – \hat{t}_j);
16                  c_j' = c_i + \bar{Y}_ij;
17              end if;
18              if c_j > c_j' then
19                  L_j ← [t_j, c_j'];
20                  q_j ← i;
21                  if j ∉ SEL then SEL ∪{j}, append j to the end of SEL;
22                  end if;
23              end if;
24          end for;
25      end while;
26  end;
```

Figure 5: Reliable shortest path (RSP) solution algorithm

In the algorithm, cost labels are set considering the expected time of using a given link in the time expanded network (i.e. $E[VT_{ij}]$, $E[WT_{ij}]$ or $E[TT_{ij}]$), so there is consistency with the objective function and the optimality conditions. Any time the optimality conditions are violated, the label is updated to remove the violation (i.e. lines 17-18). At termination the label on a given node c_j is the least expected cost from the origin to that node; no other path exists that has a shorter expected travel time. If one did exist then c_j would have been replaced with the shorter path's label in the algorithm. Given this, the optimal path is provided at the termination of the RSP algorithm; an adaptation of a proof from Ahuja et. al (1993) is provided.

Proposition: *The RSP algorithm terminates with cost labels on every node $j \in N$ that represent the least expected cost of the optimal path from origin, r, to j.*

Proof: Consider any solution c_j satisfying the conditions in Equations 10-12. Let the path from r to j consist of nodes $[i_1 = r, i_2, \ldots, i_{k-1}, i_k = j]$. The conditions in Equations 10-12 imply

$$c_{i_k} \leq c_{i_{k-1}} + E[Y_{i_{k-1} i_k}]$$

$$c_{i_{k-1}} \leq c_{i_{k-2}} + E[Y_{i_{k-2} i_{k-1}}]$$

$$\vdots$$

$$c_{i_2} \leq c_{i_1} + E[Y_{i_1 i_2}] = E[Y_{i_1 i_2}]$$

since $c_{i_1} = c_r = 0$. Adding the inequalities yields

$$c_{i_k} \leq E[Y_{i_{k-1} i_k}] + E[Y_{i_{k-2} i_{k-1}}] + \ldots + E[Y_{i_1 i_2}] = \sum_{(i,j) \in A} E[Y_{ij}]$$

which shows that c_j is a lower bound on the expected cost of any path from the origin to node j. The cost label is also an upper bound on the expected cost of the path, therefore c_j is the least expected cost. □

A small example of the RSP algorithm applied to a portion of the network (nodes 1-4) in Figure 3 is included. A detailed representation of the subnetwork is again shown in Figure 6, and Table 5 contains scheduled as well as the mean and standard deviation of vehicle arrival/departure times. The expected headways of vehicles serving nodes 2 and 3 are also given (an additional vehicle departure node at Stop A in the same route as links *(2-D, 4-A)* and *(3-D, 6-A)* is assumed to have a scheduled and mean time of 44 and standard deviation of 0; the reader can confirm the expected headways given are consistent with Equation 9).

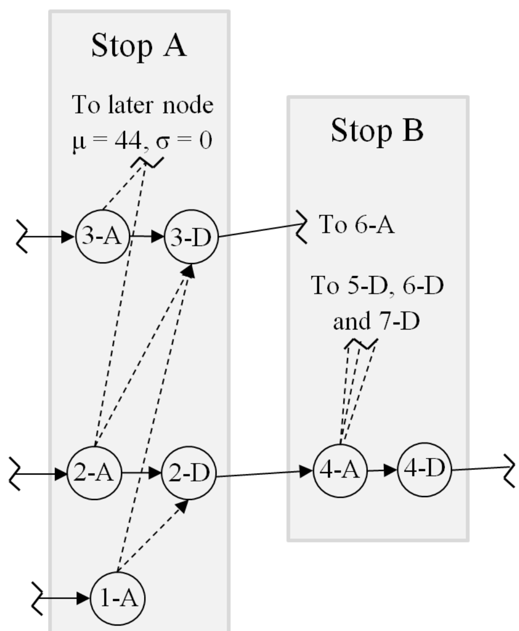

Figure 6: Subnetwork for algorithm example.

Suppose node 1-A is the origin; the algorithm proceeds as follows:

Iteration 0, Initialization:

$L_{1\text{-}A} = [0, 0]$; $q_{1\text{-}A} = 0$

$L_{2\text{-}D} = L_{3\text{-}D} = L_{4\text{-}A} = [\infty, \infty]$; $q_{2\text{-}D} = q_{3\text{-}D} = q_{4\text{-}A} = \emptyset$

SEL = {1-A}

Iteration 1, i = 1-A:

Remove node 1-A from SEL

Consider link (1-A, 2-D)

$j = 2\text{-}D$

$\bar{Y}_{1\text{-}A2\text{-}D} = \mu_{2\text{-}D} - \mu_{1\text{-}A} = 5$

$\sigma_{1\text{-}A2\text{-}D} = \sqrt{\sigma^2_{2\text{-}D} + \sigma^2_{1\text{-}A}} = 4.47$

Link (1-A, 2-D) $\in A_T$, consider the transfer

$t_{2\text{-}D} = t_{1\text{-}A} + (\hat{t}_{2\text{-}D} - \hat{t}_{1\text{-}A}) = 0 + (14 - 8) = 6$

$c_{2\text{-}D}' = c_{1\text{-}A} + E[TT_{1\text{-}A2\text{-}D}]$ where $E[TT_{1\text{-}A2\text{-}D}] = \bar{Y}_{1\text{-}A2\text{-}D} + \bar{P}_{1\text{-}A2\text{-}D}\, E(H_{2\text{-}D})$

$\bar{P}_{1\text{-}A2\text{-}D} = \Phi(-5/4.47) = 0.13$

$E[TT_{1\text{-}A2\text{-}D}] = 5 + (0.13)(11.03) = 6.43$

$c_{2\text{-}D}' = 0 + 6.43 = 6.43$

$c_{2\text{-}D}' = 6.43 < c_{2\text{-}D} = \infty$, so $L_{2\text{-}D} = [6, 6.43]$; $q_{2\text{-}D} = 1\text{-}A$

SEL = {2-D}

Consider link (1-A, 3-D)

$j = 3\text{-}D$

$\bar{Y}_{1\text{-}A3\text{-}D} = 16$

$\sigma_{1\text{-}A3\text{-}D} = 5.39$

Link (1-A, 3-D) $\in A_T$, consider the transfer

$t_{3\text{-}D} = 21$

$c_{3\text{-}D}' = 16.01$; $\bar{P}_{1\text{-}A3\text{-}D} = 1.5E\text{-}3$; $E[TT_{1\text{-}A3\text{-}D}] = 16 + (1.5E\text{-}3)(8.00) = 16.01$

$c_{3\text{-}D}' = 16.01 < c_{3\text{-}D} = \infty$, so $L_{3\text{-}D} = [21, 16.01]$; $q_{3\text{-}D} = 1\text{-}A$

SEL = {2-D, 3-D}

Iteration 2, i = 2-D:

Remove node 2-D from SEL

Consider link (2-D, 4-A)

$j = 4\text{-}A$

$\bar{Y}_{2\text{-}D4\text{-}A} = 5$

$\sigma_{2\text{-}D4\text{-}A} = 5$

Link (2-D, 4-A) $\notin A_T$

$t_{4\text{-}A} = 6 + (19 - 14) = 11$

$c_{4\text{-}A}' = 6.43 + 5 = 11.43$

$c_{4\text{-}A}' = 11.43 < c_{4\text{-}A} = \infty$, so $L_{4\text{-}A} = [11, 11.43]$; $q_{4\text{-}A} = 2\text{-}D$

SEL = {3-D, 4-A}

Table 5: Sample network information.

Node	$\hat{t}$	**μ**	**σ**	**E[H], (*link*)**	**stop**
1-A	8	10	2	--	A
2-D	14	15	4	11.03, *(1-A, 2-D)*	A
3-D	29	26	5	8.00, *(1-A, 3-D)*	A
4-A	19	20	3	--	B

The algorithm described suggests the SEL be maintained as a list, applying the FIFO (first-in, first-out) rule to select the next node to be scanned, however a deque structure can also be used. Additionally, expected waiting times for transfer links can be calculated during preprocessing and stored with the associated link to reduce computational time. The solution approach can easily be extended to also consider generalized costs. Often in transit routing passengers have different perceived weights associated with the various attributes of a trip (e.g. walking, waiting). This concept can be used in the RSP algorithm by applying the respective weight to the cost of using a link based on the link type. An additional transfer penalty could also be applied for the inconvenience or discomfort incurred which may not be captured in the penalty already incorporated for missing a transfer.

3.4 EVALUATION OF SCHEDULE DEVIATION DISTRIBUTION ASSUMPTION

A primary assumption in the methodology is that the bus stop time of different vehicles are independent random variables that are normally distributed. It was previously suggested that this is a reasonable assumption because it generally represents transit service operating around a schedule and a passenger's perception of service. However, in reality this distribution may not be normal, especially if control policies are in place to discourage vehicle operators from departing from stops ahead of schedule. In this section, the expected transfer time under the normal assumption is compared with the expected transfer time using the actual schedule deviation distributions observed from APC/AVL data in Austin, TX using simulation.

The test network that is used in this exercise is a subset of 7 bi-directional bus routes from the Austin transit network, specifically looking at service in the PM peak period. The reader can refer to the following chapter for more detail on the network. As described in Section 2.6.2, APC/AVL data was used to obtain a mean and standard deviation of schedule deviation for each stop within a route/direction. By applying this mean deviation to the scheduled times of vehicle nodes in the same route/direction in the time-expanded network, the mean arrival/departure times are obtained. This mean time and standard deviation of schedule deviation are used in calculating the expected transfer time, $E[TT_{ij}]$, according to Equation 3. This yields the expected transfer time under the normal assumption.

Alternatively, the APC/AVL data can be used to estimate the expected transfer time with simulation, using the observed stop time distributions. This is done by iteratively generating instances of the time-expanded network and determining the transfer time that would be experienced by a passenger. An instance of the network is generated by, for each vehicle trip, pulling a random observation of a bus in the same route/direction and in the PM peak period from APC/AVL. The schedule deviations observed at each stop are applied to scheduled time of the corresponding vehicle nodes in the time-expanded network; this is designated as the simulated time. Nodes in A_W are also assigned a simulated time by adding the deterministic walking time of the associated walking link to the simulated time of the previous vehicle node. For each possible transfer in the network (i.e. every link in A_T) the simulated transfer time is equal to the difference in simulated times of the two nodes making up the transfer, unless this difference is negative. A negative difference signifies the transfer would not be made in the instance of the network; later vehicle nodes in the same route, direction and stop are then considered until the transfer is possible, consistent with the assumed passenger behavior. This process of network and transfer simulation is repeated 10,000 iterations. The expected transfer time is simply determined by then averaging the simulated transfer times, since each instance of the transfer has an equal probability of occurrence in the simulation.

The expected transfer wait time under the normal assumption is compared with the simulated transfer wait time in the scatter plot shown in Figure 7. The slope of the best fit line, 0.98, indicates the expected wait time using the normal assumption marginally underestimates the actual expected wait time using the real distributions. This is seen particularly when the expected transfer time is long (i.e. over 35 minutes), however it is unlikely such long transfers would be used in an optimal path anyways. Two lines are offset 5 minutes from the best fit line to indicate nearly all of the transfers fall within this range. This suggests the normal assumption provides an estimate within 5 minutes of the actual expected transfer wait time. The mean difference between these two quantities over all possible transfers is only -0.35 minutes, or 21 seconds. This leads to the conclusion that the use of the normal assumption for vehicle arrival/departure time distributions is a reasonable approximation that will not likely lead the misidentification of optimal paths using the proposed framework.

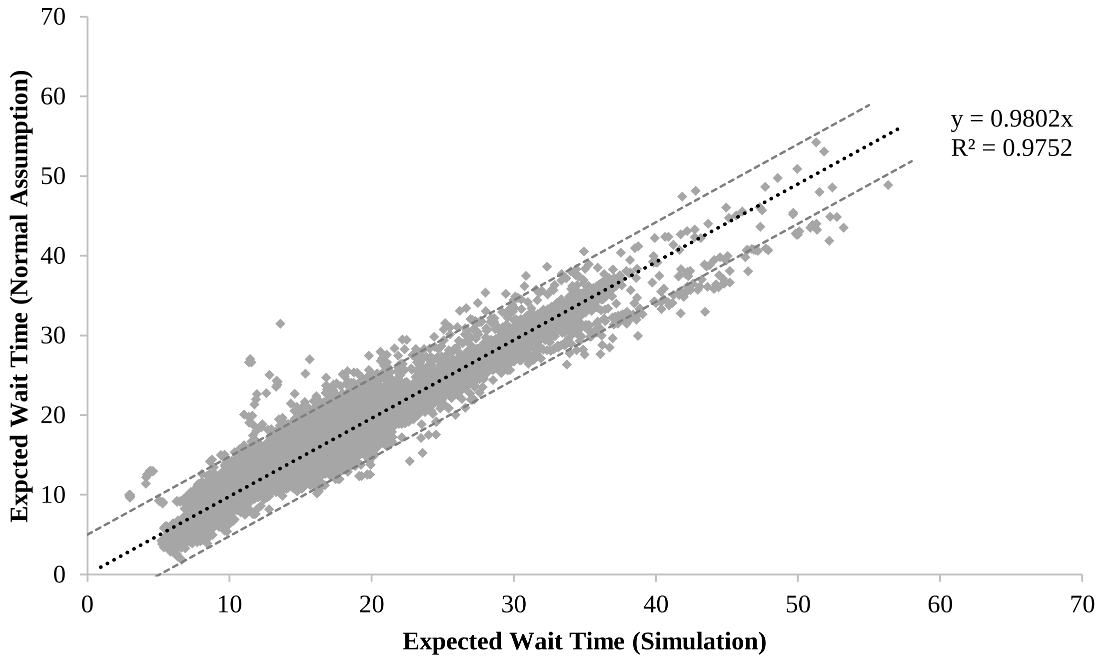

Figure 7: Comparison of expected transfer waiting times.

Chapter 4: Application to Transit Assignment Model

The RSP algorithm is applied in passenger assignment using a subset of routes from the Austin, Texas public transit network and PM peak demand. Assignment is done using both the RSP approach and a deterministic, scheduled-based, shortest path (DSP) approach. The RSP considers transfer timing and uncertainty whereas the DSP is only based on the schedule. The goal is to show how passenger routing can be improved with RSP over routing based on the timetable. Additionally, the results of the RSP assignment are compared with those of FAST-TrIPs, a schedule-based transit assignment model that has been calibrated for the behavior of Austin transit users (Khani, et al., 2014b). Passenger paths from all three approaches are simulated to demonstrate how each would perform under representative instances of the network and to determine the overall path reliability.

4.1 Network

A subset of seven bi-directional bus routes from the Austin, TX transit network is used to implement passenger assignment (see Figure 8). A time-expanded network is generated for vehicles trips that start between 2:30 PM and 7:30 PM using General Transit Feed Specification (GTFS) files (Google Developers, 2012). Vehicle nodes (set N_V) are created for stops in each vehicle trip, and are connected with in-vehicle links. Transfer walking links are created from every vehicle arrival node to walking arrival nodes (set N_W) at every stop within 0.25 miles. Waiting links are created from each arrival node (sets $N_{V\text{-}A}$ and N_W) to vehicle departure nodes scheduled no more than 30 minutes later at the same stop.

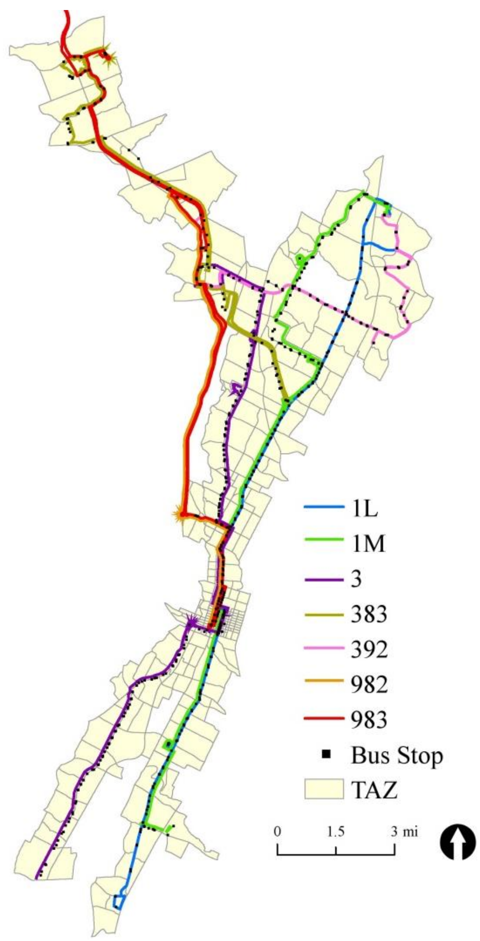

Figure 8: Austin, TX test network.

As discussed in Chapter 2, APC/AVL data was processed to obtain means and standard deviations of the deviations from the scheduled arrival/departure times for each stop in the routes. The mean deviations are applied to scheduled times of the vehicle nodes (N_V) and standard deviations are associated with the nodes for use in the solution algorithm. For walking arrival nodes (N_W), the mean and standard deviation is carried from the previous vehicle arrival node since walking time is assumed to be deterministic. The aggregated measures shown in Table 6 provide an idea of the overall reliability of each route. The mean schedule deviation across all routes and stops is 3.74 ± 6.23 minutes indicating service is usually behind schedule, as expected.

Table 6: Aggregate route reliability measures.

Route	**Direction**	**Schedule Deviation**	
		Mean	**Std. Dev.**
3	NORTHBOUND	+2.19	4.90
	SOUTHBOUND	+4.04	6.92
383	NORTHBOUND	+2.58	4.38
	SOUTHBOUND	+3.15	4.69
392	EASTBOUND	+4.99	7.09
	WESTBOUND	+5.11	6.75
982	NORTHBOUND	+2.10	5.84
	SOUTHBOUND	+3.24	8.22
983	NORTHBOUND	+2.03	5.75
	SOUTHBOUND	+1.79	6.95
1L	NORTHBOUND	+6.43	7.60
	SOUTHBOUND	+4.63	6.19
1M	NORTHBOUND	+4.35	6.04
	SOUTHBOUND	+2.90	5.76
All Routes		**+3.74**	**6.23**

Assignment models require a means of loading users onto the network. This is typically done using traffic analysis zones (TAZs), with their centroids connected to the network with centroid connector links. In a transit network, these connectors represent access from the origin to the boarding stop and egress from the alighting stop to the destination. TAZs with centroids within 0.5 miles from a stop in the included routes are used in this analysis. Each of the 343 centroids are split into two nodes, one for departure (origin) and the other for arrival (destination). Access links are connected from every TAZ departure node to a newly generated walking arrival node at every accessible stop. To accommodate passengers with different departure times within the analysis period, waiting links are created from these walking arrival nodes to every vehicle departure at the corresponding stop. For a given passenger, only initial waiting links to nodes with a scheduled time of at most 30 minutes past the passenger's arrival time at the boarding stop are considered. This setup allows for the same time-expanded network to be used for all passengers regardless of their PDT. Egress links are also created from every vehicle arrival node at accessible stops to TAZ arrival nodes. A walking speed of 4 mph is assumed for access and egress. The network representation with the inclusion of TAZs and access and egress links is shown in Figure 9. The complete time-expanded network contains 60,922 nodes (16,041 vehicle, 44,195 walking arrival and 686 TAZ nodes) and 331,937 links (15,913 in-vehicle, 39,771 transfer walking, 4,424 access, 91,025 egress, 91,858 initial waiting and 88,946 transfer waiting links).

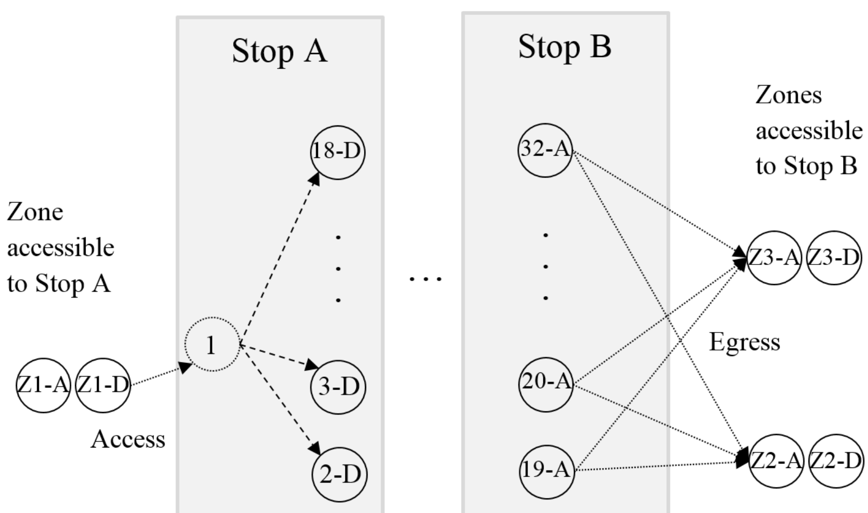

Figure 9: Time-expanded network with access/egress from TAZs.

The aggregate transit O-D trip table from the local metropolitan planning organization is used to generate disaggregate passenger demand. PM peak trips are assigned a specific preferred departure time by applying a departure time profile obtained from APC/AVL data (i.e. see Figure 2 in Chapter 2). The total passenger demand for the included TAZs with a PDT in the PM peak period (3:30 – 6:30 PM) used in assignment is 7,260.

4.2 PASSENGER ASSIGNMENT

Passengers are assigned paths in the time-expanded network using DSP, RSP and FAST-TrIPS, and passenger trajectories from each approach are recorded. The hyperpath assignment of FAST-TrIPs is used, and capacity constraints are not enforced because not every route serving the included TAZs is included in the test subnetwork. For the RSP assignment, an addition to the solution algorithm is included for determining the optimal departure time within a 30 minute window after the preferred departure time. The algorithm in Figure 10 is called after line 8 in the original RSP algorithm (Figure 5) if the link being scanned, (i, j), is an initial waiting link after an access link to a bus stop. Essentially, the departure time is incremented from the PDT to determine which departure time yields the least cost label on j. The resulting time and cost labels, t_j and c_j', are returned to determine if the label on j should be updated (i.e. line 18 of Figure 5). At the termination of the RSP algorithm, the departure time for the optimal path of the passenger can be determined as the difference between scheduled time of the initial boarding node and time label of this node because the time label includes both walking and initial waiting time.

```
algorithm getOptimalDepartureTime
begin
    Access walk time (time label of i):
    w = t_i
    if t̂_j ≥ PDT + w and t̂_j < PDT + w + 30 then
        offset = 0;
        optLabel = ∞;
        optOffset = 0;
        while offset < 30:
            arrTime = PDT + offset + w;
            if arrTime > t̂_j then break;
            end if;
            Ȳ_ij = μ_j - arrTime;
            σ_ij = σ_j;
            t_j = t_i + (t̂_j – arrTime);
            c_j' = c_i + E[TT_ij];
            if optLabel > c_j' then
                optLabel = c_j';
                optOffset = offset;
            end if;
            offset = offset + 1;
        end while;
        t_j = t_i + [t̂_j – (PDT + optOffset + w)];
        c_j' = optLabel;
    else continue;
    end if;
end;
```

Figure 10: Algorithm for determining optimal departure time for RSP.

Initial waiting time is not included in the DSP and FAST-TrIPs paths by nature; stochasticity in the transit service is not considered in these approaches so the resulting paths suggest passengers arrive at the stop at the scheduled bus departure time (i.e. passenger departure time is the scheduled time of the initial boarding node minus the access walking time). In reality, passengers may plan to arrive at the boarding stop earlier than the vehicle departure time to minimize the chance of missing the bus. To enhance the realism of these paths and to make them compatible with the simulation tests (i.e. so that initial boardings are not frequently missed), an estimate of initial waiting time from the literature is used to determine passenger departure times. A past study, conducted by Fan and Machemehl (2009), developed a predictive linear model for passenger waiting time using video data of passenger waiting in Austin, TX. A model using bus line headway as the only independent variable was estimated for transit planning purposes. This model, as described by Equation 13 where W is waiting time and H is headway in minutes, is used to estimate initial passenger waiting time and establish a departure time for DSP and FAST-TrIPs paths. Departure time is constrained to be at or later than PDT, as done with RSP.

$$W = 2.28 + 0.29H \tag{13}$$

4.3 Simulation

The simulation experiment involves generating instances of the network using APC/AVL data and simulating passenger's assigned paths to determine if they can be made successfully. An instance of the network is created by, for each vehicle trip, drawing a random observation of a vehicle trip in the same route and direction in the PM period from APC/AVL. The observed schedule deviations at each stop are applied to each corresponding vehicle node's scheduled time to get a simulated time. When working with the data, an issue was observed where an observation of a trip does not include records for every stop in the trip. This could be a result of faulty equipment or errors when recording the data. To resolve this problem, deviations are inferred from neighboring stops in the observed trip *with* data available. Four cases exist: (1) record exists for stop, (2) records exist for both upstream and downstream stops, (3) records exist for upstream stops only and (4) records exist for downstream stops only. This process is summarized in Figure 11. The simulated time of walking arrival nodes is the deterministic walking time added to the simulated time of the previous alighting vehicle node.

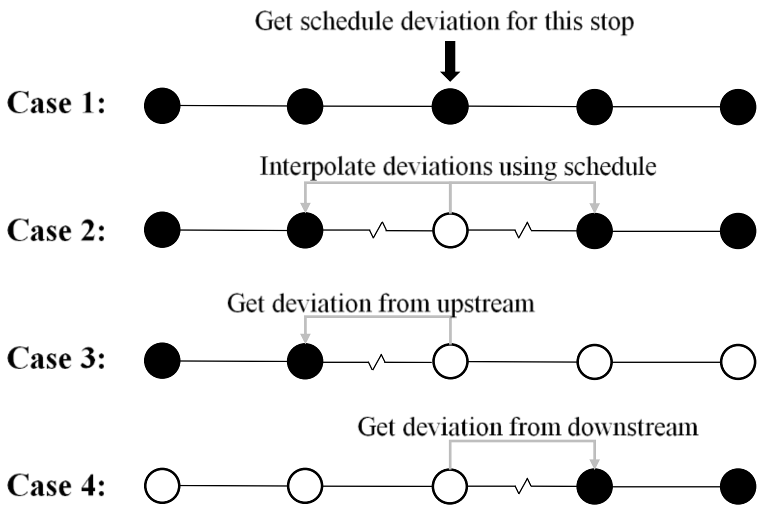

Figure 11: Generating an instance of the network from APC/AVL data.

Passengers' paths from assignment are then simulated using the instance of the network to determine if the suggested paths can be made successfully. A passenger may fail to proceed on the assigned path if the initial or transfer, if any, buses are missed. Passengers failing to board a bus are loaded onto the next bus in the same route as the missed bus, consistent with the assumed user behavior. This process of network and passenger simulation is repeated for 5,000 iterations to get path failure rates and arrival time at the destination for each passenger. The difference between the simulated arrival time and the scheduled arrival time, δ, is calculated and a distinction is made if the passenger has path failure or success (δ_{fail} and $\delta_{success}$, respectively). The travel time index, or the ratio of the simulated travel time to the scheduled travel time from assignment, is also determined.

4.4 COMPARISON OF RESULTS

The comparison of the DSP, RSP and FAST-TrIPs assignment results is shown in Table 7. Under both runs a small fraction of total demand is unassigned; these passengers largely have a PDT near the end of the analysis period so a trip cannot be made with the buses included in the generated network. While roughly 27% of passengers are assigned the same path under both DSP and RSP approaches, approximately 54% of passengers have the same final alighting nodes. This difference indicates that over half of the passengers are scheduled to arrive at the destination at the same time under both approaches, but their intermediate paths may differ. When comparing RSP and FAST-TrIPs paths, these portions are 15% and 37%, respectively. Path travel times are compared in two ways, either as the difference from the arrival time at the destination and the actual departure time (ADT) (i.e. including initial waiting time estimate) or as the difference from the arrival time and the PDT. The latter includes schedule delay, that is, the difference between preferred and actual departure time. When including this schedule delay, RSP paths are about 5% and 16% longer than the DSP and FAST-TrIPs paths, respectively.

Table 7: Comparison of DSP, RSP and FAST-TrIPs assignment results.

Assignment Results			
Measure	**DSP**	**RSP**	**FAST-TrIPs**
Assigned passengers	7252	7210	6954
Passengers with paths in common:			
RSP & DSP		— 1929 —	
RSP & FAST-TrIPs		— 864 —	
FAST-TrIPs & DSP		— 1088 —	
Average travel time (min.)*	45.8, 52.5	39.8, 52.7	45.4, 47.7
Average travel time ratio*:			
RSP:DSP		— 0.85, 1.05 —	
RSP:FAST-TrIPs		— 0.86, 1.16 —	
FAST-TrIPs:DSP		— 1.02, 0.93 —	
Total number of transfers	2379	1433	1141
Percent of passengers with transfer(s)	24.8%	17.2%	15.8%
Average boardings per passenger	1.33	1.20	1.16
Average transfer reliability	0.68	0.84	0.84
Average transfer offset time:			
Scheduled	4.69 min.	9.01 min.	11.1 min.
Experienced from AVL	4.96 min.	9.53 min.	11.2 min.

*First number uses ADT, second number uses PDT

The number of transfers is reduced considerably with RSP due to unreliable transfers being penalized, however FAST-TrIPS is slightly more stringent. Figure 12 compares the breakdown of number of transfers assigned to passengers in all approaches with the transfer rate of passengers from an on-board survey conducted by the local transit authority in spring 2010. Survey respondents traveling between the same zones used in assignment and during the PM period are considered in the chart. RSP appears to resemble the observed transfer rate the closest. The reliability of transfers suggested by the assignment (i.e. probability of making the transfer) is approximately the same for RSP and FAST-TrIPs, which is greater than DSP, as expected. The offset time between the two buses at a transfer in RSP paths in about 9 minutes; the implications of this are discussed later.

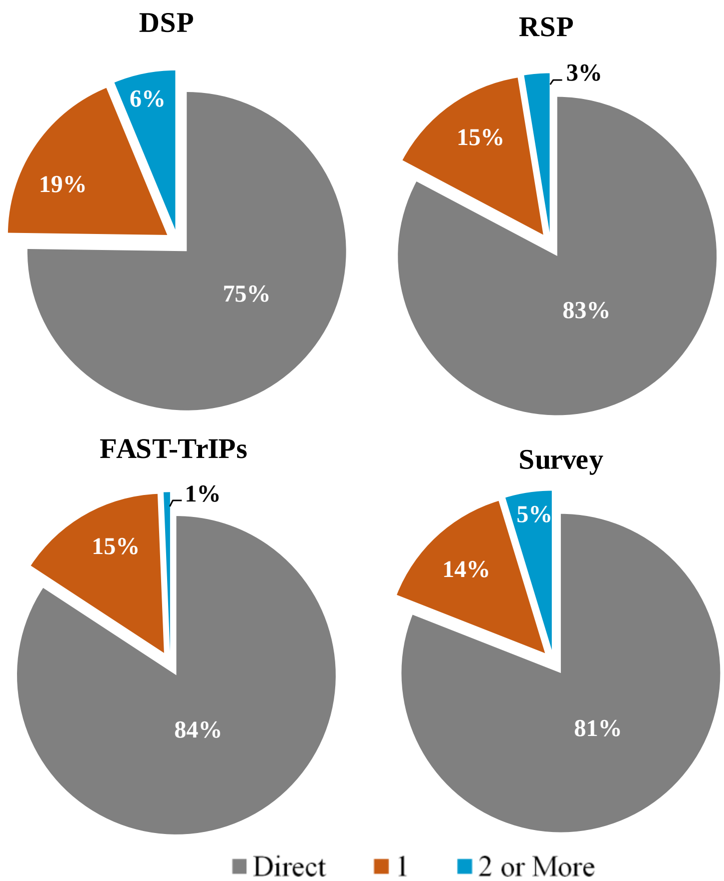

Figure 12: Comparison of transfer rate with on-board survey.

In addition to transfer rate, route ridership and load profile are other aggregate comparisons of the assignment methods. Figure 13 shows the total ridership for the PM peak predicted for each route. Given that RSP and FAST-TrIPs result in less transfers and will therefore have lower ridership predictions collectively, all three generally exhibit the same magnitude. Routes 1 and 3 are cross-town routes that serve more TAZs, and are generally known to be among the heavily used routes in the system. RSP assigns fewer to route 983, likely because this route has the longest headway (60 minutes) and the penalty for missing a boarding is extremely high. Route 982 has much higher ridership in RSP than FAST-TrIPs, possibly because it is used by more transferring passengers in RSP (i.e. FAST-TrIPs has more direct trips). To see finer differences in route ridership between RSP and FAST-TrIPs, a visualization tool is used to visualize load profiles (see Figure 14). Only minor differences are observed, for example, greater ridership at extremities of routes 1 and 3 in FAST-TrIPs and a localized "hotspot" mid-route 392 in RSP.

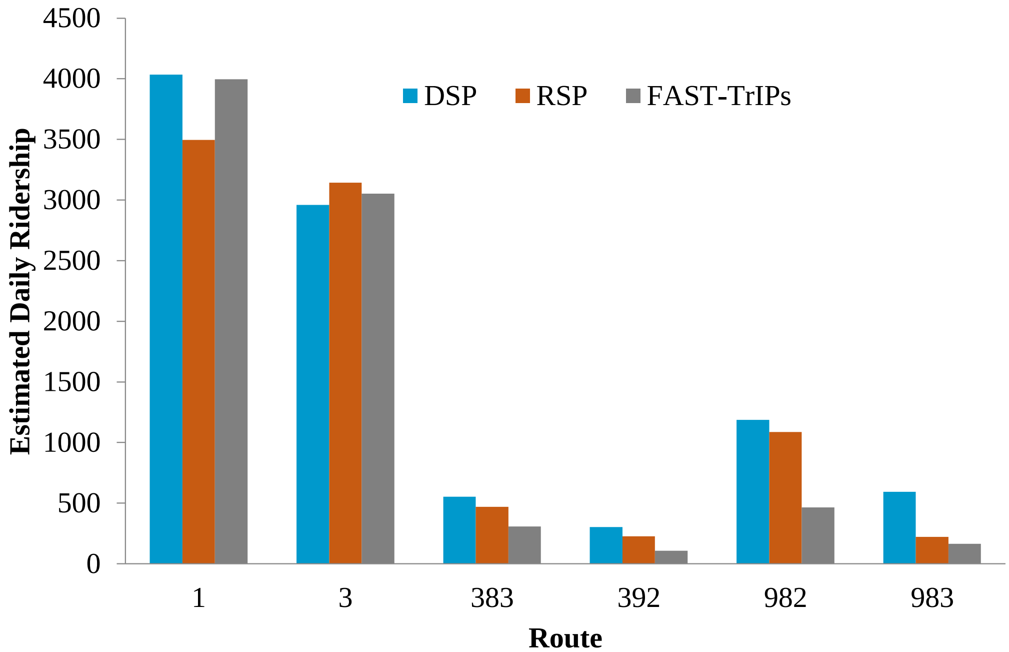

Figure 13: Estimated daily ridership.

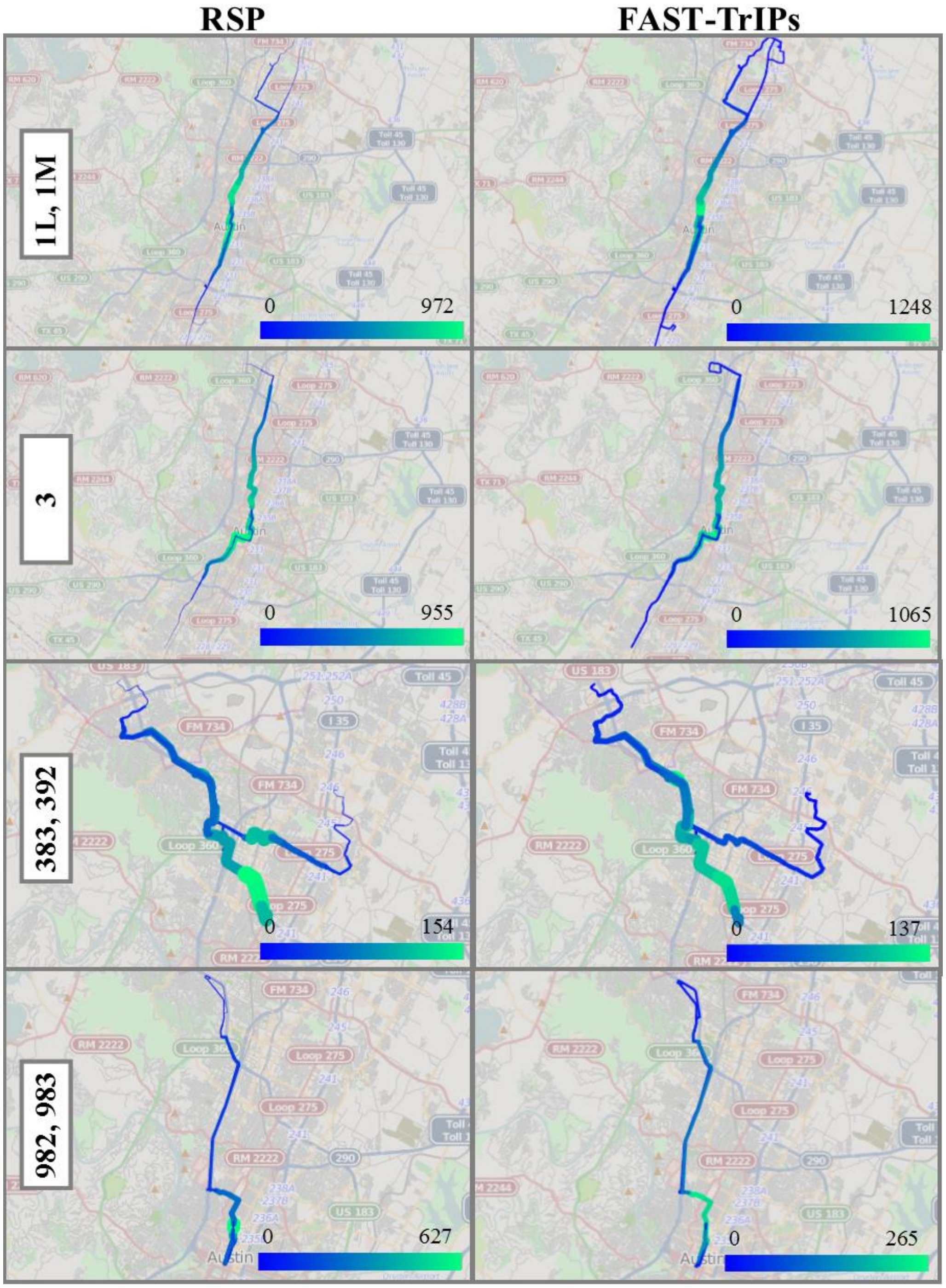

Figure 14: Visualized load profiles for RSP and FAST-TrIPs routes.

Ideally, these comparisons would also be done using field data such as APC or travel survey to see which approach is most representative of observed usage. However, since the tests have been done using only a subset of routes and zones, this becomes difficult. APC data are just counts and are unable to provide origins and destinations of passengers, so it most useful in assignment on the full network. While survey data provides these locations, other routes not included in the subnetwork can serve the included zones. This issue was observed when using the on-board survey data to compare route level ridership with the three assignments.

Simulation results using the ADT of passengers are shown in Table 8. The path failure rates can be interpreted as how frequently a passenger would miss boarding a vehicle in their assigned path if he/she were to repeatedly take that trip over time. It also represents the overall path reliability. RSP shows to be an improvement over DSP with an over 18% lower failure rate for transferring passengers. RSP is also comparable to FAST-TrIPs; the slight improvement FAST-TrIPs has in the failure rates is most likely to an improved initial boarding failure rate. This is because the process used for estimating initial waiting time (i.e. Equation 13) for DSP and FAST-TrIPs is more conservative than the process used in RSP. Average scheduled initial waiting times are 9.06, 2.48 and 6.64 minutes for DSP, RSP and FAST-TrIPs, respectively. FAST-TrIPs also has fewer transfers overall and therefore fewer opportunities to miss a boarding. Still, RSP is shown to result in less lengthy delays. The difference between passengers' simulated arrival times and scheduled arrival times (δ) are marginally less, suggesting actual arrival time at the destination is closer to, although still later than, the scheduled arrival time. The travel time indices are nearly identical, however recall that the average scheduled travel time using ADT is lowest for RSP (39.8 minutes).

Table 8: Comparison of simulation results using actual departure time.

Simulation Results (ADT)			
Measure	**DSP**	**RSP**	**FAST-TrIPs**
Average path failure rate:			
Initial boarding failure	1.43%	3.69%	2.71%
Path failure (all passengers)	10.9%	6.66%	4.97%
Path failure (transfer passengers)	39.5%	20.9%	17.3%
Average difference in simulated and scheduled arrival time:			
δ	8.07 min.	6.07 min.	6.46 min.
δ_{fail}	35.2. min.	30.4 min	32.9 min.
$\delta_{success}$	5.03 min.	4.46 min.	5.11 min.
Average travel time index	1.19	1.19	1.18

Since the initial waiting time is observed to influence the simulation results and a main focus of this work is on transfer reliability, simulation is repeated using the PDT of passengers. This assumes passengers depart their origins at the PDT and the schedule delay is built into their travel time. These results are given in Table 9. Now, the failure rate of RSP paths is marginally improved over FAST-TrIPs and still results in less delay. On average, passengers can expect their actual travel time to be about 11% longer than scheduled when taking RSP paths compared to 17% longer for DSP and FAST-TrIPs paths.

Table 9: Comparison of simulation results using preferred departure time.

Simulation Results (PDT)			
Measure	**DSP**	**RSP**	**FAST-TrIPs**
Average path failure rate:			
Initial boarding failure	1.23%	0.70%	2.68%
Path failure (all passengers)	10.7%	3.69%	4.93%
Path failure (transfer passengers)	39.3%	18.4%	17.2%
Average difference in simulated and scheduled arrival time:			
δ	8.01 min.	5.32 min.	6.44 min.
δ_{fail}	36.3 min.	32.3 min.	32.8 min.
$\delta_{success}$	5.00 min	4.27 min.	5.10 min.
Average travel time index	1.17	1.11	1.17

The assignment and simulation results can also give insight into designing more reliable transfers. The average scheduled offset or buffer between buses on the optimal paths with RSP is 9 minutes. In actual conditions (i.e. from APC/AVL), this difference is slightly longer, or 9.5 minutes. From Figure 15, the distribution of scheduled offset time is seen to have a large spread; this is most likely because the routes included have a range of headways (e.g. 15, 26, 30, 47 minutes). A general preference towards shorter offset times is still observed, however there comes a tradeoff with transfer reliability. Figure 16 shows this tradeoff. At short offset times transfer reliability fluctuates considerably based on the operation of the two buses involved, and additional offset time provides a buffer if vehicles are off schedule, improving the reliability of the transfer. The RSP approach can be used by planners to determine how different transfer offsets, control strategies, and improved reliability can impact passengers.

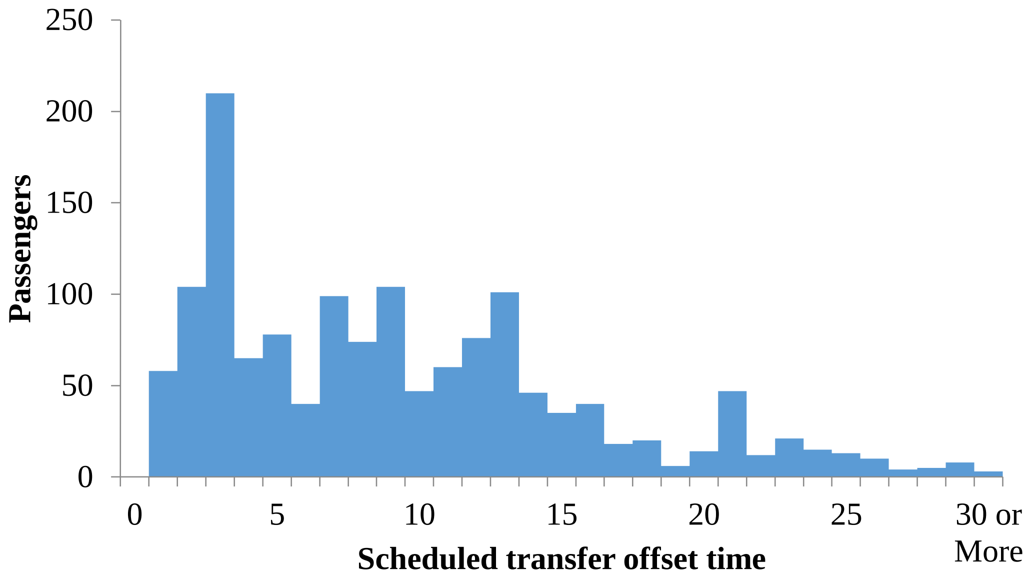

Figure 15: Distribution of scheduled transfer offset time in RSP paths.

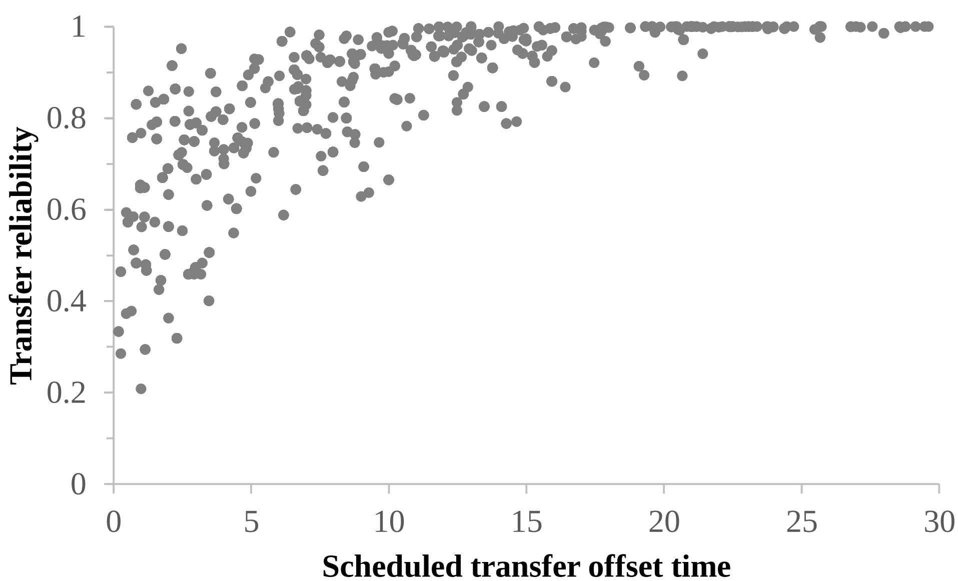

Figure 16: Tradeoff between transfer offset time and reliability.

Chapter 5: Conclusions

5.1 Implications of Work

This thesis has presented a framework, model and solution algorithm for determining the *a priori* least expected cost path in a schedule-based, time-expended public transit network. The framework overcomes many of the complexities of transit routing, incorporating time dependent service, all possible passenger movements with a careful network setup, and stochasticity in bus arrival and departure times at stops. The variability of transfer timing is considered by using actual distributions of bus arrival and departure times from AVL data. The probability of missing a transfer is penalized with the additional expected waiting time until the next bus in the route. A label-correcting algorithm, RSP, is proposed for solving the problem. Such shortest-path algorithms are generally regarded as efficient means of solving the routing problem, though the efficiency can vary with the approach used.

The RSP algorithm is integrated into transit assignment using a subset of routes in the Austin, TX network and vehicle location data from APC/AVL. Passenger paths are compared with those suggested from a deterministic, schedule-based, shortest path assignment (DSP) and a calibrated assignment model, FAST-TrIPs, that utilizes a logit-based route choice model. The network and passengers paths are simulated to determine overall reliability of paths from each assignment approach and the resulting delays that a passenger is likely to incur. RSP is found to assign passengers to paths with fewer transfers, at a rate consistent with a passenger survey, and improved transfer and overall path reliability. There is not a significant difference observed between RSP and FAST-TrIPs results, other than the stricter transfer penalty of FAST-TrIPs resulting in fewer boardings and thus route ridership and localized differences in load profile. In fact, while FAST-TrIPs does not explicitly consider reliability, the reliability of suggested transfers is nearly identical to that of RSP. This is worthy of future investigation in a complete or different transit network.

5.2 Future Work

This study lays the groundwork for future research extensions and consideration for reliable routing in transit networks. The proposed framework has some limitations that should be explored in the future. The utilized time expanded network setup is able to consider all possible movements and is compatible with all passenger departure times, but the network size will become very large for complete transit systems. The seven route Austin subnetwork used as a test is already very large, so assignment will the full network and demand will be computationally expensive. Alternative network setups or ways of reducing size should be considered. The main contributor to the size is not necessarily the number of routes included, or at least not directly. It is the transfer opportunities (e.g. walking links and nodes and resulting transfer waiting links) that make up the majority of the network, along with the means by which passengers enter and exit the network. For larger networks, the latter could be addressed by creating access and initial waiting links each time the algorithm is run for an origin and PDT, then removing them afterwards. To make application on a full system feasible, the large network size will need to be addressed. Tests on a complete network and comparison with transit data like APC and travel surveys will be beneficial in evaluating the difference in the RSP and FAST-TrIPs assignment approaches.

Future research can also be done to evaluate different user behaviors and objectives. The framework relies on the assumption that passengers board the next bus in the same route whenever a transfer is missed. In reality, passengers may seek out alternative routes to get them to their destination. How or if such strategies can be incorporated is something to consider. Alternative objectives such as determining the least expected time path that meets a given probability of arrival at the destination on time can also be investigated. This would require tracking the probability of getting to each node at a stated travel time in the labeling algorithm. At the termination of the algorithm, labels at the destination and their corresponding probabilities would be compared against a minimum acceptable probability of arriving on time. Finally, travel time or transfer reliabilities should be tested as independent variables in the FAST-TrIPs route choice model to determine their significance. Work on this has already started using Austin on-board survey and AVL data.

Overall, this thesis provides a methodological framework to begin addressing reliability in transit assignment context. The result is a realistic and practical routing model that can be used in transit assignment or in quantifying passenger benefits of improved reliability. The outcomes can also be useful to transit agencies in designing transit schedules and operating strategies to improve transfer reliability.

References

Ahuja, R. K., T. L. Magnanti and J. B. Orlin (1993). Network flows: theory, algorithms, and applications. Prentice-Hall, New Jersey, 133-165.

Balcombe, R., N. Mackett, N. Paulley, J. Preston, J. Shires, H. Titheridge, M. Wardman and P. White (2004). The demand for public transport: a practical guide. *TRL Report TRL593*.

Bellman, R. (1958). On a routing problem. *Quarterly of Applied Mathematics 16*, 87-90.

Bly, P. H. (1976). Depleted bus services: the effect of rescheduling. Report LR699. Transport and Road Research Laboratory, Crowthorne, U.K.

Cedar A., S. Chowdhury, N. Taghipouran, and J. Olsen (2013). Modelling public-transport users' behaviour at connection point. *Transport Policy 27*, 112-122.

Chriqui, C. and P. Robillard. Common Bus Lines (1975). *Transportation Science 9*(2), 115-121.

Currie, G (2005). The demand of bus rapid transit. *Journal of Public Transportation 6*(1), 41-55.

Fan, W. and R. B. Machemehl (2009). Do transit users just wait for buses or wait with strategies: Some numerical results that transit planners should see. *Transportation Research Record 2111*, 169-176.

Fan, Y. Y, R. E. Kalaba, and J. E. Moore (2005). Arriving on Time. *Journal of Optimization Theory and Applications 127*(3), 497-513.

Fan, Y. Y. and Y. Nie (2006). Optimal Routing for Maximizing Travel Time Reliability. *Networks and Spatial Economics 6*(3-4), 333-334.

Fan, Y. Y., R. E. Kalaba, and J. E. Moore (2005). Shortest Paths in Stochastic Networks with Correlated Link Costs. *Computers and Mathematics with Applications 49*(9), 1549-1564.

Frank, H. Shortest Paths in Probabilistic Graphs (1969). *Operations Research 17*(4), 583-599.

Fu, L. and L. R. Rilett (1998). Expected Shortest Paths in Dynamic and Stochastic Traffic Networks. *Transportation Research Part B 32*(7), 499-516.

Furth, P. G., B. Hemily, T.H.J. Muller, and J. G. Strathman (2006). TCRP Report 113: Using archived AVL-APC data to improve transit performance and management. Transportation Research Board, Washington, D.C.

Gao, S. and I. Chabini (2006). Optimal Routing Policy Problems in Stochastic Time-Dependent Networks. *Transportation Research Part B 40*(2), 93-122.

Google Developers (2012). *What is GTFS? – Transit.* Website: https://developers.google.com/transit/gtfs. Accessed May 21, 2014.

Guo, Z. and N. H. M. Wilson (2011). Assessing the Cost of Transfer Inconvenience in Public Transport Systems: A Case Study of the London Underground. *Transportation Research Part A 45*, 91-104.

Hall, R. W (1986). The Fastest Path through a Network with Random Time-Dependent Travel Times. *Transportation Science 20*(3), 182-188.

Hamdouch, Y. and S. Lawphongpanich (2008). Schedule-Based Transit Assignment Model with Travel Strategies and Capacity Constraints. *Transportation Research Part B 24*, 663-684.

Hamdouch, Y., W. Y. Szeto, and Y. Jiang (2014). A New Schedule-Based Transit Assignment Model with Travel Strategies and Supply Uncertainties. *Transportation Research Part B 67*, 35-67.

Khani, A., B. Bustillos, H. Noh, Y.C. Chiu and M. Hickman (2014a). Modeling transit and intermodal tours in a dynamic multimodal network. *Proceedings of the 93rd Annual Meeting of Transportation Research Board*, Washington, D.C.

Khani, A., E. Sall, L. Zorn and M. Hickman (2013). Integration of the FAST-TrIPs person-based dynamic transit assignment model, the SF-CHAMP regional activity-based travel demand model, and San Francisco's citywide dynamic traffic assignment model. *Proceedings of the 92nd Annual Meeting of Transportation Research Board*, Washington, D.C.

Khani, A., T. Beduhn, J. Duthie, S. Boyles and E. Jafari (2014b). A transit route choice model for application in dynamic transit assignment. Presented at the 5th Transportation Research Board Conference on Innovations in Travel Modeling, Baltimore, MD.

Knoppers, P. and T. Muller (1995). Optimized Transfer Opportunities in Public Transport. *Transportation Science 29*(1), 101-105.

Lee, A., N. Van Oort, and R. Van Nes (2014). Service Reliability in a Network Context: Impacts of Synchronizing Schedule in Long Headway Services. Presented at 93rd Annual Meeting of the Transportation Research Board, Washington, D.C.

Mai, E., G. List, and R. Hranac (2012). Simulating the travel time impact of missed transit connections. *Transportation Research Record 2274*, 69-76.

Miller-Hooks, E. D. and H. S. Mahmassani (2000). Least Expected Time Paths in Stochastic, Time-Varying Transportation Networks. *Transportation Science 32*(2), 198-215.

Muller, T. H. J. and P. G. Furth (2009). Transfer Scheduling and Control to Reduce Passenger Waiting Time. *Transportation Research Record 2112*, 111-118.

Nguyen, S. and S. Pallottino (1988). Equilibrium Traffic Assignment for Large Scale Transit Networks. *European Journal of Operations Research 37*, 176-186.

Nie, Y. and X. Wu (2009). Shortest Path Problem Considering On-Time Arrival Probability. *Transportation Research Part B 43*, 597-613.

Sen, S., R. Pillai, S. Joshi, and A. K. Rathi (2001). A Mean-Variance Model for Route Guidance in Advanced Traveler Information Systems. *Transportation Science 35*(1), 37-49.

SHRP 2 C10(B) (2014). Partnership to develop an integrated advanced travel demand model with mode choice capability and fine-grained, time-sensitive networks. Transportation Research Board. Available at http://apps.trb.org/cmsfeed/TRBNetProjectDisplay.asp?ProjectID=2828. Acesseed August 2014.

Sivakumar, R. A. and B. Rajan (1994). The Variance-Constrained Shortest Path Problem. *Transportation Science 28*(4), 309-316.

Spiess, H. and M. Florian (1989). Optimal Strategies: A New Assignment Model for Transit Networks. *Transportation Research Part B 23*(2), 83-102.

Szeto, W. Y, M. Solayappan, and Y. Jiang (2011). Reliability-Based Transit Assignment for Congestion Stochastic Transit Networks. *Computer-Aided Civil and Infrastructure Engineering 26*, 311-326.

Szeto, W. Y., Y. Jiang, K. I. Wong, and M. Solayappan (2013). Reliability-Based Stochastic Transit Assignment with Capacity Constraints: Formulation and Solution Method. *Transportation Research Part C 35*, 286-304.

Tong, C. O. and A. J. Richardson (1984). A Computer Model for Finding the Time-Dependent Minimum Path in a Transit System with Fixed Schedules. *Journal of Advanced Transportation 18*(2), 145-161.

Tong, C. O. and S. C. Wong (1999). A Stochastic Transit Assignment Model Using a Dynamic Schedule-Based Network. *Transportation Research Part B 33*, 107-121.

Transit Capacity and Quality of Service Manual, Third Edition (2013). TCRP Report 165. Transportation Research Board, Washington, D.C.

Turnquist, M. A. and L. A. Bowman (1980). The effects of network structure on reliability of transit service. *Transportation Research Part B 14*, 79-86.

Van Oort, N (2011). *Service Reliability and Urban Public Transport Design.* TRAIL Thesis Series T2011/2. TRAIL Research School, Delft. Available: http://www.goudappel.nl/media/files/uploads/2011_Proefschrift_Niels_van_Oort.pdf

Waller, S. T. and A. K. Ziliaskopoulos (2002). On the Online Shortest Path Problem with Limited Arc Cost Dependencies. *Networks 40*(4), 216-227.

Wilson, N. H. M. and A. Nuzzolo (2004). *Schedule-Based Transit Modeling: Theory and Applications*. Springer, New York.

Yang, L. and W. H. K. Lam (2006). Probit-Type Reliability-Based Transit Network Assignment. T*ransportation Research Record 1977*, 154-163.

Zockaie, A., Y. Nie, and X. Wu. and H. S. Mahmassani (2013). Impacts of Correlation on Reliable Shortest Path Finding. *Transportation Research Record 2334*, 1-9.

Vita

Tyler James Beduhn was born in Madison, Wisconsin. After graduating from Sun Prairie High School in 2009, he began his studies at The University of Iowa in civil engineering with a minor in business administration and emphasis in transportation. He received his Bachelor of Science in Engineering in 2013, graduating with highest distinction. In August 2013, he began his graduate studies in transportation engineering at The University of Texas at Austin under supervision of Dr. Stephen Boyles. He received his Master of Science in Engineering in civil engineering in December 2014. After graduation, he moved to Arlington, Virginia to begin a career with Kimley-Horn and Associates, Inc.

Permanent email address: tyler.beduhn@gmail.com

This thesis was typed by the author.