A Methodology for Identifying Inconsistencies Between Scheduled and Observed Travel and Transfer Times using Transit AVL data: Framework and Case Study of Columbus, OH

A Thesis

Presented in Partial Fulfillment of the Requirements for the Degree
Master of City and Regional Planning and Master of Science in the
Graduate School of The Ohio State University

By

Yuxuan Wang, B.S.

Graduate Program in City and Regional Planning and
Graduate Program in Civil Engineering

The Ohio State University

2020

Master's Examination Committee:

Andre Carrel, Advisor

Zhenhua Chen, Advisor

Gulsah Akar

Rabi Mishalani

Abstract

Transit schedules are essential planning tools for many passengers – especially for low-frequency services, transfers, or travel decisions made ahead of the time. This underscores the continued importance of accurate transit schedules to passengers. Agencies, on the other hand, also count on reliable travel times and schedule adherence data for timetable and headway scheduling, as well as operational planning, to provide reliable services to passengers. The objective of this study is to develop a tool to identify potential discrepancies between scheduled and observed mean travel times at a stop-to-stop level after removing outliers, and to identify potential inconsistencies between scheduled and observed transfer times at the transfer point level. Demonstrations of the tool's application are provided using data collected from the Central Ohio Transit Authority (COTA).

In the COTA application, the study reveals that for certain route sections, the schedules do not accurately capture the observed temporal variation in mean travel times throughout the day, especially between peak hours and midday. In other cases, spatial patterns are visible in the discrepancies between observed and scheduled travel times, with segments near major intersections tending to have longer than scheduled travel times, which result in further delays. Stop-to-stop level travel time distributions are examined to support scheduling purposes. Several distributions are shown to provide a better fit to observed stop-to-stop travel times than commonly used distributions, namely, the Epsilon skewed normal, Generalised Extreme Value, and mixture normal distributions. Furthermore, this research examines correlations between travel times across stop-to-stop segments to understand the degree to which delays may propagate. The results show that most stop-to-stop travel times can be considered independent from the those of adjacent segments. Higher positive correlations can be observed at areas with lower traffic and passenger activities. In addition, potential driver behaviour changes when running ahead of schedule are pointed out by calculating the conditional travel times given no delays.

In a further analysis, discrepancies between scheduled and observed transfer times are examined. Like many other agencies, transfers are not coordinated nor guaranteed by COTA. Overall transfer reliability is examined at the system level, and conditional probabilities of passengers being able to make a connection after a delay on the first bus are derived. It is found that three minutes of scheduled transfer time correspond to an 85% probability of the transfer being successful. If the first bus arrives on or before the transfer bus is scheduled to arrive, there is also an 85% probability of the transfer being successful. A general comparison of observed and scheduled transfer times shows that around 20% of transfer points have shorter than scheduled transfer times and 30% have longer than scheduled transfer times. For the remaining 50% of transfer points, the scheduled transfer times roughly correspond to the observed transfer times. Such analyses allow the agency to inform passengers regarding the risks associated with a planned transfer, and it will allow agencies to assess risks of propagating delays to future trips associated with holding the transfer bus to guarantee transfers.

Vita

2018 .... B.S. Computer and Information Science

Publications

Research Publications

Chen Z., Wang Y. "Impacts of severe weather events on high-speed rail and aviation delays". *Transportation research part D: transport and environment*, 69:168–183, 2019.

Fields of Study

Major Field: City and Regional Planning

Major Field: Civil Engineering

Table of Contents

Page

Abstract ii

Vita iv

List of Tables vii

List of Figures viii

1. Introduction 1

1.1 Motivation 1
1.1.1 Objectives 4
1.2 COTA System Overview 5
1.2.1 Transit System Redesign (TSR) 6
1.2.2 Service Standards 6
1.2.3 Service Monitoring and Service Changes 7
1.3 Organization of this Thesis 8

2. Literature Review 10

2.1 Travel Time Variability and Adjustments 10
2.2 Travel Time Monitoring 12
2.3 Transfer Applications and Coordination 16
2.4 Current Transit Scheduling Methods 18
2.5 Discussion 21

3. Data and Data Processing 25

3.1 Introduction to GTFS 25
3.1.1 GTFS Static 26
3.1.2 GTFS Real Time 31

3.2 Data Collection . . . 34
3.3 Data Processing . . . 34
3.3.1 Matching GPS to Trips . . . 35
3.3.2 Calculate Stop Times . . . 36
3.3.3 Extract Travel Times . . . 37
3.3.4 Outlier Removal . . . 39
3.4 Data Aggregation for Travel Time Analysis . . . 40
3.4.1 Data Visualisation . . . 43

4. Travel Time Analysis . . . 44

4.1 Travel Time Inconsistencies Across Time and Space . . . 47
4.1.1 Stop to Stop Level . . . 47
4.1.2 Trip Level . . . 52
4.1.3 Line Level . . . 57
4.1.4 System Level . . . 60
4.2 Stop to Stop Travel Time Distribution . . . 65
4.3 Stop to Stop Travel Time Correlations . . . 76
4.4 Conditional Stop to Stop Travel Time Given Delay . . . 82

5. Transfer Time Analysis . . . 87

5.1 Missed Transfer Probabilities Given Scheduled Transfer Time . . . 89
5.2 Missed Transfer Probabilities Given Arrival Time of First Bus . . . 93
5.3 Identifying Transfer Time Inconsistencies Across Time . . . 98
5.3.1 Transfer Level . . . 98
5.3.2 System Level . . . 100
5.4 Holding buses for transfer passengers . . . 102

6. Discussions and Conclusions . . . 110

6.1 Research Summary . . . 110
6.1.1 Travel Time Analysis . . . 111
6.1.2 Transfer Time Analysis . . . 114
6.2 Preliminary Conclusions . . . 115
6.3 Potential Use of the System . . . 118
6.4 Limitations and Future Research . . . 119

Appendices 122

A. Appendix . . . 122

List of Tables

Table **Page**

4.1 Mean Travel time vs Scheduled Travel Time. 50

4.2 Sample Travel Data for Line 31 Trip 697167 53

4.3 Line 31 Northeast Weekday Stop Classification Counts 58

4.4 Number of stop pair classifications trend for weekday services 61

4.5 Travel Time Distribution Counts for Trip 697167. 68

4.6 Travel Time Distribution Counts by Hour. 74

4.7 System-wide Correlation Classification by Hour 80

4.8 On Time Performance at Stop Level for Trip 697167. 84

5.1 Observed Missed Transfer Probability Given a Scheduled Transfer Time in Minutes. 91

5.2 Missed Transfer Probabilities Given Scheduled Transfer Bus Arrival and Actual Arrival of First Bus. 95

5.3 Transfer Point Classifications by Hour. 101

5.4 Transfer Holding Strategies Compared for 2019.01.23. 105

A.1 Best Fit Distributions for each Segment in Trip 697167 by AIC. . . . 122

A.2 Complete Trip Travel Time Table for Trip 697167. 124

List of Figures

Figure **Page**

3.1 Example of Stop Time Calculation 37

3.2 Example of outlier Removal Results 41

4.1 Illustration of Travel Time Inconsistencies Between King & Delashmut and King & Olentangy River Road. 51

4.2 Trip Level Travel Time Difference Visualization for trip 697167. . . . 54

4.3 Line 31 Northeast Weekday Trip Classification Plot 57

4.4 Systemwide stop to stop level travel time classification count plot . . 62

4.5 Systemwide Stop to Stop Level Travel Time Visualization for Weekdays at Noon 63

4.6 Systemwide Stop to Stop Level Travel Time Visualization for Weekdays at 5 PM 64

4.7 Travel Time Histogram and Fitted PDF for Trip 697167 from North Star & Presidential to 1197 Kinnear. 69

4.8 Travel Time Histogram and Fitted PDF for Trip 697167 from King & Olentangy River Road to King & Olentangy Trail. 70

4.9 Stop to Stop Travel Time Distribution for Trip 697167. 72

4.10 Travel Time Distribution Plot By Hour. 73

4.11 Stop to Stop Travel Time Correlation Map for Trip 697167 78

4.12 System-wide Correlation Classification Plot by Hour 81

5.1 Missed Transfer Given Schedule Time. 92

5.2 Missed Transfer Given Difference Between Schedule Time and Delay. 96

5.3 Transfer Classification Plot by Hour. 102

5.4 Missed Transfer Counts using Different Holding Strategies 106

5.5 Delayed Next Terminal Departure Counts using Different Holding Strategies 107

5.6 Holding Counts using Different Holding Strategies 108

Chapter 1: Introduction

1.1 Motivation

A reliable public transit system is essential to attract passengers, and literature have pointed out the importance of transit reliability (Carrel et al., 2013) from passengers' perspective. However, despite the importance of service reliability to passengers, there has been relatively fewer attention to transit reliability at stop to stop level. Researchers and agencies have been analysing transit services at terminal to terminal level or timepoint to timepoint level. Yet, passengers may not travel from terminal to terminal or timepoint to timepoint (Halvorsen et al., 2019). As a result, most researchers have also cited the lack of data as a factor for not being able to examine transit systems at greater detail (Mazloumi et al., 2010). In addition, there is a mismatch between the passenger information software based on GTFS which provides stop arrival times and the common scheduling practice at timepoint or terminal level (Wessel et al., 2017). Furthermore, since not all transfer points are located at timepoints, transfer performances cannot be evaluated directly using timepoint level data. More comprehensive description of stop to stop travel times and transfer point performances is still needed.

This thesis aims to build a system that collects and analyses the vehicle location data. The system analyses travel times and transfer time at a finer grain level, i.e. at stop to stop level instead of timepoint to timepoint level commonly used in practice. Agencies already monitor their on-time performance at timepoint level and address any significant on-time performance issues using these data by adding time to trips or segments where vehicles consistently run late, since most AVL systems do not provide travel time analyses by default (Muller and Furth, 2001). The main goal of the system focuses on other aspects of the scheduling, which is to calculate whether there are inconsistencies between the scheduled and actual travel times and transfer times. These calculations aim to capture temporal and spatial variations of these inconsistencies if any. The analyses will in turn be done at stop to stop level, trip level, route level, and system level.

Other than the main goal, this system will also have the ability to aggregate travel time and arrival time data in different ways for other studies. The following studies are demonstrated using the system. If agencies wish to provide more accurate stop arrival or departure times, stop schedules generated by unknown software processes after the scheduling process at timepoint level are not particularly helpful for making improvements, thus more descriptive statistics are still needed. Agencies need to keep in mind variation and potential skewness of the travel time distribution when determining slack times for segments with larger variation. These skewness and distribution information could then be used in simulations to determine a good operation strategy. Agencies might also want to prioritise addressing issues on certain segments. If the stop to stop travel times are independent, meaning the variance can be summed up, agencies could prioritise reducing variation on segments with larger variances thus reducing the overall terminal to terminal variance. In addition, agencies might want to see if simple travel time assumptions could be used as starting point to reduce their workload for stop level scheduling. This thesis attempts to model travel time distributions and travel time correlations as well to address the above questions. In addition, since driver behaviours might also contribute to inconsistencies between scheduled and observed travel times, conditional travel times given delays is explored but not further examined by the thesis due to lack of data.

As for transfer studies, agencies may want to adopt a set of transfer related standard, similar to on time performance. Thus, probability of missing transfers is calculated as a reference. In order to help agencies potentially improve passenger transfer experience using of these aforementioned observations, simulation for holding strategies based on historical data is demonstrated using the system. The system can also be easily expanded to incorporate other data sources from other agencies such as weather, traffic, and land use data.

To illustrate the potential use of the system, it is applied to data collected from Central Ohio Transit Authority (COTA). COTA redesigned services in May 2017 to have brand new routes and schedules, and the shift to a grid like service network sometimes requires passengers to make a transfer to complete their trips. The agency also made its AVL data public in 2017. Comprehensive AVL datasets from COTA were collected for this thesis to offer a comprehensive view of the agency's new transit system operations for the past two years starting January 2018, when the data collection began, to March 2020, when the results were obtained for this thesis.

1.1.1 Objectives

Based on the motivations presented above, this subsection summarises the main objectives for this study. There are two main objectives, one related to stop-to-stop travel times and one related to transfer times. This drives the structure of the analyses, which will be referred to hereafter as the "travel time component" and the "transfer time component".

For the travel time component, the objective is to identify inconsistencies between the actual service experienced by passengers and the scheduled travel time expected by the passengers. This analysis will be done at stop to stop level, trip level, line level, and system level to capture the spatial and temporal differences. In addition, this thesis will look at the distributions of the stop to stop travel times, due to its importance on the theoretical analyses done by previous research (Kieu, 2015). Furthermore, since the sum of stop to stop travel time and stop dwelling times should be the terminal to terminal travel time, this thesis examines the correlations between the travel time on neighbouring stop to stop segments to provide more theoretical information when summing up stop to stop segments. Finally, this thesis explorers the potential behavioural change from operators when they run early and when they run late, to see if these stop to stop level travel times are consistent with previous research by Levinson (1991), which shows drivers tend to get further delayed if they leave the starting terminal late.

For the transfer analysis, the objective is to describe transfer times at system level from actual observed data in addition to the previous research efforts that focused more on deriving these transfer times from theoretical travel time and transfer time distributions or simulations (Bookbinder and Désilets, 1992). Then, the probabilities of missing a transfer given different amount of information is explored. Finally, this thesis considers the impact of a simple holding strategy that would improve the probability of successful transfers.

1.2 COTA System Overview

This thesis used data from Central Ohio Transit Authority, commonly referred to as COTA. COTA is the primary transit authority in Central Ohio. It provides services mainly in Franklin County, and portions of Delaware, Fairfield, Union, and Licking counties (McCann, 2016). COTA's fixed route services consist of 13 local lines runs through downtown Columbus; 10 crosstown lines serving outside downtown area; 13 rush-hour only commuter lines, 2 limited stop services, 4 special services.

COTA redesigned their transit system in 2017 through streamlining bus routes, adding more frequent services, and introducing several all-day limited stop services. A more detailed background on their service design and operation can be found in later sections in this section.

After long delays, they also launched real-time bus tracking through the Transit App and began publishing relevant data in General Transit Feed Specification (GTFS) online for the public. In addition, the GTFS feed describes all of the agency's stops, trips, routes, and schedules, and GTFS real time provides GPS tracking most of their vehicles, with a small portion of vehicles missing from the feed due to equipment or internet outages. This makes analysing services at more detailed level possible. Since their real-time bus tracking system and their redesigned system operation became more stabilised in January 2018, it is a more reasonable starting point to collect data and analyse their newly designed services in detail.

1.2.1 Transit System Redesign (TSR)

Over the past decade, central Ohio region experienced rapid growth. COTA was expecting 8.8% population growth and 4.4% job growth in its service area by 2025 (McCann, 2016). This growth will lead to more public transit users, either due to increased congestion or due to increased demand for quality transit services (McCann, 2016). The main goal for TSR is to respond to the increasing demand brought by the regional growth and changes in development patterns, as development patterns are shifting back to more urban designs. The resulting higher density also increased transit demand which requires more consistent, simplified, and frequent services (McCann, 2016). The redesigned system was implemented on May 1, 2017.

The new transit system follows a "grid like" approach, where lines run as linear as possible. This system depend more on connections with other linear lines to allow passengers to get to areas served by other lines. During the day, these transfers are generally inferred from the published schedule by passengers, i.e. these transfers are not specifically scheduled nor guaranteed by the agency. A timed-transfer system, commonly referred to as "line up" goes into effect at 10pm everyday to guarantee transfers between buses in downtown area due to reduced service level at night.

1.2.2 Service Standards

According to COTA service standards (McCann, 2016), TSR classified COTA services into three categories, Frequent, Standard, and Rush Hour.

Frequent lines operate every 15 minutes or better most of the day, seven days a week. The main service goal is aligned for best ridership outcomes. The service is designed as linear as possible on the same corridor with no deviations are allowed without justification (McCann, 2016).

Standard lines operate every 30 or 60 minutes, 7 days a week. For lines that operate every 30 minute, the service purpose is mixed with ridership and coverage, where the line have a mixture of ridership and coverage segments. While for the 60 minute services, the main goal is to provide coverage and basic level of access to people. These services provide as linear as possible on the same corridor, with limited deviations to serve major activity centers (McCann, 2016).

Rush Hour lines operates only during weekday peak periods. It offers non stop service mainly between suburban areas and downtown Columbus. The direction that goes with peak flow is considered as commute services, while the services going against the peak flow is considered as reverse commute services. These services are designed as direct as possible for the express segments (McCann, 2016).

1.2.3 Service Monitoring and Service Changes

Current service change data collection and service change process are based on several Intelligent Transportation Systems (ITS) implemented between 2015 and 2019 (McCann, 2016). The core of COTA's ITS system, which also related to this thesis, are the Computer Aided Dispatch (CAD), Automatic Vehicle Location (AVL) systems. ITS created opportunities for COTA to monitor their operations, responding proactively to changes in demand or traffic, and provide real-time travel information for passengers.

The CAD and AVL system provides operational metrics, such as whether the bus is running on time compared to the scheduled times using their on time performance metrics (McCann, 2016). Vehicle locations used in these systems are communicated via 4G cellular data system every 15 seconds. The system also achieves the vehicle locations which allows the agency to review incidents and evaluating route performances. Currently, the operational metrics are measured at timepoint levels.

To ensure their services are in accordance with the published schedule, COTA monitors its on-time performance based on its own service standards (McCann, 2016). "A vehicle is considered 'on-time' when its arrival is from zero to 4 minutes and 59 seconds after the scheduled time. A vehicle is considered 'late' when it arrives five minutes or more after the scheduled time" (McCann, 2016). Express lines can arrive at stops up to five minutes early after the express segment where the buses do not stop, whereas other services are not allowed to run early at timepoint stops. Their on-time performance standard is 80% of all buses should arrive at scheduled timepoints "on-time". These on-time performance data is available at least on a monthly basis, where the system should be reviewed on average over each month.

COTA's service changes happen every trimester, on the first Monday of January, May and September every year. These changes occur to improve their service metrics, such as on-time performance mentioned above, as well as to improve productivity of existing services (McCann, 2016).

1.3 Organization of this Thesis

The rest of this thesis document is organised in the following way.

Literature on previous transit reliability research is presented in Chapter 2. This chapter reviews past research on general use of AVL data, bus travel time variability, passenger metrics from agencies, and transfer coordination.

Chapter 3 gives an overview of the methodology used in this study, details include data sources, data processing, data storage, and data aggregation.

Chapter 4 begins the COTA study, with detailed discussion of spatial and temporal travel time inconsistencies between scheduled travel times and observed mean travel times, travel time distributions, conditional travel times given a certain bus delay, and travel time correlations from stop to stop.

Chapter 5 focuses on the temporal transfer time inconsistencies at different transfer points, and conditional probability of successful transfers given scheduled transfer time and given actual arrival time in relation to scheduled transfer bus departure time. In addition, this chapter applied a simple holding strategy on the observed transfer data to illustrate the effects of holding.

Chapter 6 discusses the findings and concludes this thesis. This chapter summarises and combines the findings in previous chapters with potentially useful operational insights for agencies. Limitations, potential improvements, and future expansions of this study are also proposed and explored.

Chapter 2: Literature Review

This chapter focuses on the previous research efforts relating to this thesis. The main topics discussed below are travel time variability, travel time monitoring, transfer coordination, and transit scheduling methods. These literature were selected with a specific focus on the larger dataset applications. After the detailed summaries of previous literature, there is a short discussion that ties previous research efforts to the contribution of this thesis.

2.1 Travel Time Variability and Adjustments

Travel time variability has limited prior studied due to the lack of comprehensive dataset (Mazloumi et al., 2010), and thus frameworks for analysing these data still needs to be developed. However, prior studies mainly focused on passenger experiences waiting at bus stops. With the implementation of GPS tracking systems on transit vehicles, Mazloumi et al. (2010) used bus GPS data from Melbourne suburbs to examine factors that caused complete trip (i.e. terminal to terminal) travel time variability. They found travel times in a given departure window are best characterised by normal distribution during peak hours while off peak windows are best fitted by log normal distribution. Then, they explored factors, such as route length, number of stops, traffic signals, delays, industrial land use, rain, and time of day, that impact complete trip travel times using ordinary least square linear regression. They found that route length is the most important route characteristic that contributes to travel time variability, while industrial land use had highest impact on travel time variability as an external factor. They conclude travel time distributions are important inputs to understand travel time variability, and that shortening route sections or overall route lengths, consolidating bus stops, multi-door boarding and alighting, could reduce travel time variability. They also found buses running ahead of schedule tend to have longer travel times than late buses. The rain factor is only found to be significant in AM peak.

El-Geneidy et al. (2011) mentioned the issue there are little effort to studied collected AVL and APC data in evaluating transit performance. They analysed travel time adherence and reliability for one cross-town route in Minneapolis, MN. at time point level and route level. Their result demonstrate potential ways of identifying service level decline, and they recommended schedule revision to increase the travel time and arrival time adherence. Due to low ridership at certain stops, they also recommended stop consolidation to decrease service variability. They also fitted linear regression line to model run times at timepoint level. The result indicate that run time is longer at the end of service pattern, and departure delays increase the run time.

Liao and Liu (2010) developed a data processing framework for transit performance analysis, using AVL, APC, automatic fare collection (AFC), and schedule data, to enable further analysis in transit performance and support operational decision making. They utilised one month of archived data obtained from Metro Transit in Minneapolis, MN. They analysed these data at the time point level, which includes time point schedule adherence, and link travel time between timepoints. They demonstrate their model can assist agencies in conducting further analysis in transit performance and factors causing delays at timepoint level, thus enables agencies to develop more robust transit schedules.

Another area related to transit scheduling is predicting vehicle arrival times, which has been thoroughly studied by scholars. However, O'Sullivan et al. (2016) pointed out the non-linear natures of existing bus arrival time prediction models, and potential drawbacks of these models due to variations in input data. They argue the prediction problems are complicated, due to complex interactions among determining factors such as passenger demand, weather, accidents, and service capacity. They developed a metamodel approach, which enhances existing models using quantile regression to place bounds on associated errors.

2.2 Travel Time Monitoring

This section focuses on real-time vehicle movement monitoring and their applications with a stronger attention on agency practices from New York City Transit and Tri-County Metropolitan Transportation District of Oregon (TriMet). The more generalised approaches not specific to any agencies proposed by academics are also discussed later in this section.

Transit agencies monitors their real-time vehicle movements during their daily operations. Agencies can obtain operational metrics and performance analyses from the archived vehicle location data to improve their service. To reiterate, COTA uses real-time GPS locations to monitor their vehicles and calculate bus delay times, which allows dispatchers to instruct bus drivers to run limited or express service to get back on schedule if running behind or instruct bus drivers to stop if running early (McCann, 2016). COTA archives information to calculate performance metrics and to assist incident resolutions (McCann, 2016).

New York City Transit has used traditional operation-focused performance metrics to evaluate their performance (Halvorsen et al., 2019). However, they mention that these metrics, like on time performance at terminals, are useful for agency to take actions to improve operations, but they do not reflect customer experiences. Halvorsen et al. (2019) and Graves et al. (2019) argue that passengers rarely travel a line from terminal to terminal, and customers tend to understand system performance in terms of additional time they have to spend waiting or riding. As a result, they developed passenger centric performance metrics for their subway and bus services to allow transit planners to review customer experience issues along a route. This allows New York City Transit to improve their transparency and public communications. They developed an online subway dashboard and two new passenger centric metrics, additional platform time and additional train time. Similarly, New York City Transit had also developed customer focused journey time metrics for their bus services (Graves et al., 2019), additional bus stop times and additional travel time. They are calculated using the origin destination matrix and vehicle movement data, either through their train supervision records or bus GPS records, which allow them to assign passengers to vehicles. Their new metrics received praise from many people, and are now used by NYCT for public communications. To illustrate the relationships between the passenger centric service metrics and traditional operational focused metrics, they calculated the correlations. These metrics are then shown to have moderate correlation between the existing operational focused metrics like on time performance and wait assessment (i.e. percent of passengers who waited longer than 5 minutes). They argue this is likely due to traditional metrics like on time performance do not take the extent of poor performance into consideration and are limited to 0% to 100%.

The archived AVL data is also used in schedule revision in agencies. Another typically monitored item is the adherence to the scheduled arrival time, which allows agencies to revise schedules and remedy delays along a route. As mentioned in Chapter 1, COTA modifies their schedules three times per year based on on-time performance data at the timepoint level.

Coleman et al. (2018) at New York City Transit developed a data-driven approach to prioritize schedule revisions for the most in need bus routes by evaluating service capacity, ridership data, and terminal to terminal travel time data. They cited the labour intensive schedule revising process and limited resources as reasons for limited schedule revisions at New York City Transit with most schedules revised once every two years. They utilise their automatic vehicle location data and their ridership data to identify routes with the greatest discrepancy between scheduled versus actual travel time as well as planned capacity versus the peak load point ridership for in depth reviews. They analysed these capacity and running times based on fifteen to sixty minute time intervals throughout the day, with shorter time interval during peak hours and longer time interval during off-peak hours. Then, they calculate scores for every route based on surplus or insufficient travel time or capacity, then these individual scores are multiplied to obtain the overall ranking of the route. Highest scores are assigned to routes that exceed their performance threshold, i.e. median running time deviates more than 2 minutes, and ridership either exceeds capacity or are below half of the provided capacity. This prioritisation process makes schedule revisions at New York City Transit more responsive to changes in travel conditions.

TriMet in Portland, Oregon also utilises their performance metrics to improve their transit schedules (Kimpel et al., 2008). They record time point to time point travel times along the route and calculates 20th, 50th, and 80th percentile travel times. They consider the 50th percentile run time is the time typically required for operators to operate the segment, while 20th and 80th percentiles are arbitrary upper and lower bounds as indicators for extreme insufficiencies and extreme surpluses in the schedule. They then applied cluster algorithms and nonparametric statistics on the median run times to obtain an optimal run time for their scheduling process.

Some scholars have also developed approaches that are not specific to any agencies. Basak et al. (2019) developed a data-driven approach to optimize transit schedules. They choose to optimize the probability of buses arriving at timepoints on time. They utilised genetic algorithm and particle swarm optimization algorithms for the optimisation process. Their result show particle swarm optimization requires less computing resources while also providing good results.

Kieu (2015) studied the travel time variability in route level by fitting various statistical distributions to model the observed terminal to terminal travel times. Kieu (2015) found public transit travel time is significantly different from private transportation, and log-normal describes the travel times the best.

Bates et al. (2001) discusses passengers' valuation of travel reliability. They conclude that travel time reliability is highly valued by passengers, travel time distributions generally do not follow "nice distributions", such as normal distributions, thus medians become more representative when distributions are not well behaved and medians should be used when evaluating travel times.

Previous literature have pointed out the importance of transit reliability in passengers' decision making process. Carrel et al. (2013) found that in-vehicle delays are more likely to drive people away from transit than having longer waiting times at the origin stop. They also pointed out that passengers may evaluate services reliability based on the real time information they received from mobile apps, instead of the service schedule provided by the agencies.

2.3 Transfer Applications and Coordination

The New York City Transit passenger centric metrics also considered the impacts from passenger transfers (Graves et al., 2019; Halvorsen et al., 2019). They utilised the origin destination matrices to generate linked trips, and match each individual trip segment to running vehicles. Both groups uses scheduled running time to model customer decisions to reflect the customers' lack of knowledge on future delays when deciding which train to board. However, when service interruptions occur, such as service changes, station closures, or the vehicle a passenger is more likely to board, considering overcrowding, are delayed by more than 15 minutes, their model will look for alternative routes that complete the trip faster. Due to major service changes or track works during late nights and weekends, their metrics are not applied to these time periods due to the change in overall trip patterns.

Kieu (2015) developed a simulation model to evaluate transfer times given different bus travel time variability and bus departure times at a transfer point. From the simulation, Kieu (2015) found there is no planned transfer time (i.e. transfer determined by planners from one vehicle to another built into the schedules) that minimizes the mean transfer time nor the probability of missing a transfer. Through online (holding based on real time bus arrival prediction) and offline (scheduling and operation changes) coordinated transfer strategies proposed by the author, the mean transfer time can be reduced by 20 percent and the probability of missing the transfer can be reduced by 80 percent.

Bookbinder and Désilets (1992) proposed a model for transfer optimization and coordination. The model takes into account bus travel time variations and different objective functions such as the expected transfer times or the variation in transfer times. They utilised bus travel time distributions to model the transfer times at the transfer stop. The authors show the expected transfer times are hard to improve, but the variations in transfer times can be reduced easily. They further conclude that "transfer optimization under a no-hold policy may produce more improvements when headways are long" (Bookbinder and Désilets, 1992).

Knoppers and Muller (1995) studied the positive and negative implications resulted from transfer coordinating strategy. Based on vehicle and passenger arrival time distribution the authors examined waiting times expectations and variations at transfer stops. Through derivations, they conclude schedule coordination becomes less effective when frequencies are higher on the connecting line. They also conclude that delayed departure at one transfer point may endanger subsequent transfer points.

Shafahi and Khani (2010) is a more recent attempt in optimizing transfer waiting times. Shafahi and Khani (2010) formulated the transfer problem into a mixed integer programming problem and a genetic algorithm problem to optimize transfers by minimizing waiting times. Their most significant assumptions are that the headways are uniform, travel times are given, and dwelling time at stops are constants. They conclude that mixed integer programming is good for smaller agencies, while the genetic algorithm is better for larger agencies.

Another improvements to the transfer optimisation is from Parbo et al. (2014). They incorporated passengers' modified route choice as results from schedule changes into their optimisation process by using the Danish national route choice model and national transit schedule. Their research results led to a yearly reduction in weighted waiting time equivalent to 45 million Danish Kroner. However, they suggest future research to look consider non-deterministic bus arrival times in addition to their scheduled based analyses.

2.4 Current Transit Scheduling Methods

Wessel et al. (2017) proposes a method to retroactively improve the accuracy of transit agencies' GTFS feed by using real time vehicle locations provided. Using these observed data, they created a routable bus timetable representing service delivered which can then be used in assessing transit performance, reliability, and accessibility issues. They argue that transit schedules are often done at timepoints, whereas arrival and departure times for stops in between timepoints are not yet defined in practice. GTFS, on the other hand, requires arrival and departure times for every stop served by a certain trip. There is a discrepancy between the general scheduling practices and what's shown to the passengers through GTFS, which calls for further investigation into stop level scheduling practices.

Ceder (1987) listed four bus headway setting methods. He mentions service frequencies are determined to maintain adequate service quality and minimize number of buses in operation and efficiently allocate resources to gather ridership data. He proposes three types of headways. Equal headway simply means headways are evenly spaced during each time period. Balanced headways, are not evenly spaced, but they are set so that observed passenger loads would be similar on all buses. Smooth headways are simply the average of equal and balanced headways. Then, headways can be set based on passenger loads. Ceder's first two headway setting methods are based on point checks, i.e. ridership obtained at one point along the route. The first method is based on data gathered at max load point during the whole day. The second method is to determine service frequency based on hourly or period max load point. The third and fourth method are based on ride checks, i.e. data obtained from a complete bus trip. The third method sets the frequency so that the average load along the route is at or under the desired capacity. The fourth method adds a constraint to the third method that limits the length of the route where the load exceeds a certain overcrowding limit. In addition, special requests can be applied to the headways, such as clocked faced headways. Once the headways are set, the average round trip times including holding and layover times are used as additional inputs for setting trip departure and arrival times.

Levinson (1991) synthesizes the agency practices on bus running times. He points out that good schedules should allow enough running time to operate the route, account for congestion and signal delays, and provide enough layover time at the terminal so that late buses can start the next trip on time. He also found that if drivers leave the terminal late, it is unlikely that they are able to catch up the lost time. Levinson (1991) also concludes that scheduled run times should be set at a value slightly less than the mean or median run time in order to ensure that the majority of operators do not have to kill time in order to maintain their schedule adherence.

In a later synthesis, Furth (2000) argued that whether scheduled running times are based on average running times, ideal running times, or a high percentile running times (e.g. 75 to 95 percentile) is a matter of agency policy. Ideal running times help prevent buses from running early, which is less desirable for passengers than trips running late. However, under ideal schedules, trips will tend to run late, and will need greater scheduled layovers at terminals. The high percentile running times help prevent trips from being late, but increases the chances of running early unless agencies implement strict policy to reduce early departures. The high percentile running times also mean less layover time is needed in the schedule. The average running time policy is a compromise between those extremes. Furth argues while average running time is sufficient to determine scheduled running time, scheduled layovers should be determined based on running time variability. As a result, most agencies choose to use a high percentile, typically 85th percentile, as their "half cycle" time, i.e. the terminal to terminal travel time plus layover time at the terminal. This allows them to set standards so that a small percentage of vehicles will not be able to start the next trip on time.

Muller and Furth (2001) argue that people mistakenly assume automatic vehicle location systems provides travel time analyses to agencies by default, simply because it tracks the vehicles. However, this is not the case. They further argued that schedules at timepoints must be set correctly to allow good schedule adherence. They pointed out that European systems consider every stop as time points to avoid large schedule deviations. They argue precise travel time scheduling under urban congestion requires larger data provided by trip time analysing system. They shows a scheduling strategy that sets travel time to be the 85th percentile from every timepoint to the terminal. The idea is that if drivers end up early and holding at a time point, the remainder of the trip should also satisfy the on time requirement with high probability. Then, the individual segment travel times are set using repeated subtractions.

Coleman et al. (2018) summarized the scheduling practice at New York City Transit. They review express routes yearly, weekday schedules at minimal every two years, while the weekend are required to be revised at least every four years. Their scheduling process mainly includes analyzing ridership and end to end running time data. During peak hours, they set their load standard to 100% of standing and seating rooms, while during off peak hours, the load standard is reduced to refer the desired maximum load by passengers. Their bus GPS data are matched to timepoints to determine the running times. If the actual ridership exceeds their designed capacity, service frequency is increased.

2.5 Discussion

This section points out or reiterates potential improvements and research needs from the existing literature listed in previous sections. The literature review shows that agencies commonly design their schedules at timepoint level or at trip level, and most studies reviewed in this chapter use time point level schedule adherence data or terminal to terminal run time data to conduct their analyses for bus systems. Yet, GTFS standard requires arrival and departure times at stop level, which are not yet clearly defined by most agencies (Wessel et al., 2017). There is an obvious mismatch between the agencies' service design at timepoint level and passenger experience at stop level. In addition, despite timepoints are commonly placed at higher ridership stops to evaluate service reliability, not all passengers travel from timepoint to timepoint (Halvorsen et al., 2019). This mismatch between what is provided by the agency and what is perceived by passengers calls for investigation on whether GTFS stop to stop travel times are consistent and representative with the actual travel times experienced by the passengers, and the stop to stop level analysis represent passenger's experience closer.

In addition, transit scheduling practices reviewed in this chapter also pointed out that scheduling normally picks 85th percentile as the "half cycle" time (Kimpel et al., 2008), as previous literature assume 85th percentile is the mean plus one standard deviation. This assumption means that if drivers start their trips on time, 15 percent of the drivers will be so late to get to the terminal that they will not be able to start their next trip on time (Furth, 2000). However, this practice may potentially violate the statistical definition when creating stop schedules or even timepoint schedules, which require summing up individual travel time for each segment to the terminal to terminal travel time. By definition, the sum of stop to stop travel times and dwelling times at each stop adds up to the total terminal to terminal travel time. Unfortunately, the sum of 85th percentiles stop to stop travel times does not equal to 85th percentile terminal to terminal travel time. Regardless, in practice, some agencies determine timepoint to timepoint travel times by subtracting the 85th percentile travel times from one timepoint to the terminal from the 85th percentile travel time from previous timepoint to the terminal (Muller and Furth, 2001). Since percentiles cannot be directly summed up or subtracted, this practice may have contributed to the inconsistencies passenger experience at stop to stop level as well. Based on these potential issues, this thesis will focus on the mean travel times. In addition, the practice of getting 85th percentile by adding one standard deviation to the mean may only be true when the data are normally distributed. Thus, the distributions of the data need to be examined further at stop to stop level if agencies decide to continue this practice at stop to stop level.

Furthermore, Carrel et al. (2013) also highlights the importance of transfer times which are more important compared to waiting time at the origin stop. Despite agencies use some transfer points as timepoints for their service evaluation, not all passengers transfer at timepoints. In addition, most of the transfer optimisation and coordination literature reviewed in this chapter (Bookbinder and Désilets, 1992; Kieu, 2015) are based on statistical bus travel and arrival time variations and expected transfer times are inferred from these distributions. There still lacks a comprehensive overview of what are the actual transfer times passengers experienced.

One reason for the lack of research is a lack of comprehensive datasets, which has been pointed out by other researchers (Mazloumi et al., 2010). It is not surprising that the selected studies on travel time variations and deviations only analysed a few transit lines, and none of them conducted systemwide analyses. One or two transit line might not be adequate to reflect systemwide travel time deviation pattern.

This thesis tries to look at the transit unreliability or inconsistencies experienced by passengers at stop to stop travel times and transfer times in addition to agencies' on time performance monitoring. In order to add more evidence to the previous theoretical derivation research efforts on travel time and transfer time distributions (Bookbinder and Désilets, 1992; Kieu, 2015), this thesis also examines stop to stop travel time distributions and correlations. This thesis will also examine the conditional travel times when drivers run late to see if the results are consistent with previous research done at terminal to terminal level by Levinson (1991).

This thesis also aims to look at transfer times from the observed data rather than the derived distributions or simulations (Bookbinder and Désilets, 1992; Kieu, 2015). The focus for this thesis is to examine transfer time inconsistencies, if any, at each individual transfer point, and probabilities of missing a transfer are calculated using the observed data.

Chapter 3: Data and Data Processing

This chapter describes the data and methodology used in the thesis. In general, the data used here are GTFS static and GTFS real time feeds provided by COTA. The system collects GTFS real time feeds every 10 seconds, and GTFS static feeds when updates become available. There are more detailed descriptions for the GTFS standards in this chapter. Then the data is stored in a relational database for processing and analysing travel time and transfer times. More detailed descriptions can be found later in this chapter.

3.1 Introduction to GTFS

General Transit Feed Specification (GTFS) was developed by Tri-County Metropolitan Transportation District of Oregon (TriMet) along with Google. The goal was to enable online trip planning through data that are shared with the general public. TriMet and Google formatted transit data from the agency into a format that could be imported into trip planning applications like Google Maps. It provides a universal specification for transit agencies to publish their data to the public. As a result, GTFS has been adopted by various agencies around the world and third-party software developers for numerous purposes, including trip planning and real-time information systems. COTA partnered with Transit App to provide trip planning and real-time data to passengers.

GTFS feeds published by COTA described in the following subsections, as well as most other agencies, contain two main components, GTFS Static and GTFS real time. GTFS Static contains transit information that are likely to stay the same for some period of time. These information include transit schedule, fare, and geographical shapes. On the other hand, GTFS real time provides real time information that contains predicted arrival times, vehicle location updates, as well as service alerts.

3.1.1 GTFS Static

A GTFS Static feed is similar to a relational database schema which consists of a series of tables and their relations, except GTFS is stored and published in a series of comma-delimited text files. A GTFS Static feed defines and describes the agency's common information, such as their lines and schedule, for the public. The required information by the GTFS Static standard are agency, stop, routes, scheduled trips, scheduled stop times for each trip, stops, and service dates information. This thesis extensively uses all these required information, along with the additional shape file provided by COTA's GTFS file. The following paragraphs describes some important fields related to this thesis and how COTA defines these fields.

The Agency dataset is a required dataset by GTFS. This dataset contains the agency or agencies whose services are described in the dataset. COTA's GTFS feed only contains one agency, thus the file contains only the header row and one data row.

Agency ids that help distinguish services run by different agencies. Agency Name is the full name of the agency. Agency URL is the website for the agency. Agency timezone specifies the timezone which the agency operates in, which also defines the time zones for the defined arrival and departure times. There are additional fields that further describes the agency, such as the customer service phone number. These additional fields are not described here as they are not used in this thesis.

Calendar and Calendar Dates datasets describe when will the service run, which is also required by GTFS standard. Calendar dataset allows agency to distinguish weekday and weekend services by specifying a weekly schedule. Calendar Dates, on the other hand, defines individual dates whether a service will or will not occur.

Calendar file contains a unique identifier, service ID, for a set of dates when services are available. Each service is defined by days of the week, from Monday to Sunday. Start date defines the start of the service interval defined in this GTFS feed. Similarly, end date specifies the end date for the service interval described in this GTFS feed. Each day of the week has an indicator whether the service operates on all corresponding days of the week in the date interval ranging from start date to end date. The field is 1 when service is available, 0 when service is not available. COTA uses this file to define most of its services, as most of its services are categorised into weekday, Saturday, Sunday services.

Calendar dates file only contains three fields, service id, which references the calendar file. It identifies a set of dates when a service exception occurs for a specific service id. Date field defines when the service exception occurs. Together, service id and date are used as compound keys in GTFS standards, meaning each service id and date pair can only show up once in calendar dates file. Finally, there an exception type indicator that shows whether the referenced service has been added or removed from the corresponding date. COTA mainly uses this file to specify service exceptions, such as major holidays when buses operate on Sunday schedules. COTA also uses this file to define service dates for services that do not follow a strictly follows day of the week schedule, such as additional OSU services that operates only when OSU is in session.

The Route dataset defines transit routes, which is a group of trips that are advertised as one single service. In COTA's case, there are 47 rows listed in the dataset.

Each lines is assigned a unique route ID, that identifies the lines. Each line is then associated to the agency operating the route by referencing the agency ID in the agency file. Route Short Name is for a short name for the riders to identify a route. This might be different from the Route ID field. For example, COTA uses 101 as internal route id for the CMAX line, where CMAX is the route short name known to passengers as well as displayed on the header sign. Route long name is the full name of the route. COTA uses this field to describe the main corridor a route serves, or the main destination a route serves. There are additional fields that allows passengers to distinguish the services, such as the route color. These additional fields are not described here as they are not used in this thesis.

Trip dataset lists the trips on each route. Each trip is a trip from the origin terminal to the destination terminal on a specific time of the day.

Each trip is assigned a unique trip ID, which help identifies a specific trip. A trip is also associated with three foreign keys, route id, shape id, and service id. These three fields references the route, geographical shape that describes the path vehicle follow, and service dates for the corresponding trip. Each trip also contains the header sign of the trip, which is the text identifying the trip's destination that appears on the displays for passengers to distinguish short turns and branches on the route. Direction id indicates the directions of travel. All of COTA's trip are linear and bi-directional, meaning a service operates from one terminal to another terminal, then back to the original terminal. Block ID identifies the block which the trip belongs. A block contains one or more sequential trips a vehicle will follow during the day. This is different from COTA's run number. COTA uses run numbers to specify one or more sequential trips a driver will follow during the shift. A run can be different from the block, as driver shift changes may occur in the middle of the route while the vehicle and passengers will continue the trip.

Stop dataset is a collection of all the stops where transit vehicles pick up and drop off passengers.

In the dataset, each line corresponds to one stop, distinguished by a unique stop ID. Stop Code field shows the stop number posted on the stop signs, which allows agencies to use phone-based information systems for passengers to get information. Stop Name is the full name of the stop, which allow people to understand the location of the stop. For COTA, it is the name of the road intersection where the bus stop is located. Stop latitude and longitude fields describes the geographical location of each stop. There are other optional fields not mentioned here, since those are not applicable for COTA system or not relevant to this thesis.

Stop Time dataset corresponds to each trip, and delineates specific times a transit vehicle arrives and departs from each trip. These stop level arrival or departure times are determined by proprietary softwares agencies use after the scheduling process at timepoint level. People generally use these software as black boxes not knowing the process used to generate these stop to stop arrival or departure times. For COTA data, it is reasonable to speculate that these stop level arrival and departure times are linearly interpolated, since stop to stop segments between two timepoints have the same travel speed.

Stop times data has two foreign keys, trip id and stop id. These two foreign keys reference a trip these stop time applies to, and stops being served in this particular trip. Then each stop, from beginning to the end of the trip, is assigned a stop sequence number. The stop sequence will monotonically increase as the bus runs along the trip. Each stop is also associated with an arrival time and a departure time. COTA does not model dwelling times at stops, except for the line ups. This means, the arrival time is the same as the departure time in this dataset. For services after midnight, these two times are set to values greater than 24:00:00. Each stop also has a corresponding shape distance traveled value showing the actual distance traveled along the associated geographical shape.

Shapes dataset details the paths a transit vehicle will follow for a collection of trips. COTA uses the road center lines to define these paths.

Shape id is required to identify a shape. Each shape has a collection of shape points that are used to define the paths a vehicle will follow. Shape points are stored as latitude and longitudes. Each shape point is assigned a shape point sequence, showing the order of shape points that connects to the path of the trip. Each shape point is also associated with shape distance traveled value, similar to the stop times dataset. This shows the actual distance from the beginning of the path to the shape point specified in this record.

Transfer file in the GTFS standard is optional. Most applications derive transfer points based on geographical proximity of stops, like most passengers would plan their transfers based on the published schedule. However, this file will help specify additional rules for trip planning applications when identifying transfer points. Especially when transfers are not possible between routes at specific locations. COTA does not provide a complete transfer rules file, and I had to manually add more transfer points into the transfer file following the specified GTFS format for the transfer analysis in Chapter 5.

Each transfer is identified by a from stop id and a to stop id. These two stop ids are foreign keys referencing the stops dataset. In conjunction, they identify the stop where transfer begins, and the stop where the transfer ends. Each pair of stop ids is also assigned a transfer type, specifying the types of transfers, such as untimed transfers which COTA operates mostly on. There are also an extended version for this file, where it further identifies from and to trips for a specific transfer point. My transfer files will follow the extended version, since the goal of the analysis is to show detailed arrival, departure, and transfer times at a given transfer point throughout a given day.

Other files, such as frequencies, are not modeled in COTA's GTFS feed, as they are optional and COTA does not operate headway based service. Therefore, this thesis omits their explanations.

3.1.2 GTFS Real Time

GTFS Real Time (GTFS-RT) is an extension to the GTFS that allows transit agencies to provide non-static updates, such as arrival time updates, vehicle locations, and service alerts, to its passengers. COTA utilises all three elements of GTFS-RT for passenger information, Trip Update, Vehicle Position, and Alert. Currently at

COTA, Trip Updates and Vehicle Locations updates are sent out every 15 seconds. Service Alerts are sent out once they are in effect. Most of the fields specified in these feeds are optional, due to different services status being involved. Thus, the following descriptions will focus on COTA's real time feed.

Trip Update provides real-time updates on a trip's status. It can either show future updates, such as estimated arrival time for future stops, or show past events, such as stop times for previous stops that allows passenger to determine whether the vehicle has passed. COTA's Trip Update feed consists of 5 fields, Trip Descriptor, Vehicle Descriptor, Stop Time Updates, and Timestamp. Each field contains one entity, except stop time update which can contain more than 1 stop times.

Trip Descriptor allows real time data to be matched to one trip in GTFS Static feed using the primary key Trip ID. Again, since COTA does not operate a headway based service, this is enough for our purposes. It also avoids collision with the future trips under the same Trip ID by specifying the trip's Start Date. Vehicle Descriptor gives additional information on the vehicle operating on the trip. COTA uses the bus label visible to passengers as the identifier for each vehicle. Stop Time Updates gives the updated stop times for both future stops and past stops for the trip. Each Stop Time Update shows the updated arrival and departure time for a given Stop ID on the trip. Timestamp gives the time at which the vehicle's progress was measured.

However, there are problems with COTA's Trip Updates feed. GTFS Trip Updates feed is mainly designed for passenger information, and it allows past stop times to be overwritten. Current COTA system estimate arrival times for both future and past stops are calculated as current delay time in minutes plus scheduled arrival time. Past stop times will be overwritten by COTA's system once the delay minutes changes.

Therefore, the recorded data has potentially large deviations from reality. Based on these facts, I chose to reconstruct these values using vehicle position data, which provides more certainty.

Vehicle Position shows real time information about a vehicle's geographical location. COTA's vehicle position feed has 5 fields, Trip Descriptor, Vehicle Descriptor, Vehicle Position, and Timestamp.

Similar to Trip Updates, Vehicle Positions gives the same information in Trip Descriptor and Vehicle Descriptor fields to allow data users to match vehicles to a specific trip in the GTFS Static feed. Timestamp is also the moment when the vehicle's location is measured. In addition, Vehicle Positions feed gives the geographical location of a vehicle, by using the latitude and longitude coordinates in WGS-84 system.

Alert gives passenger data about service disruptions and future planned disruptions. COTA's alert feed has 5 fields, active period, informed entity, effect, header and description texts.

Informed Entity specifies the passengers whom should be notified or impacted by this alert. This can be done at several levels, agency level, route level, route level by direction, trip level, or stop level, given the corresponding primary keys to the GTFS Static feed. Active Period defines a range of time when the alert should be shown to the passengers. Effect shows the impacts to the service, such as detour and significant delays. Header text is a quick description for the alert, similar to a title, whereas the description text describes the alert in detail.

3.2 Data Collection

The data collecting process was implemented using Java for programming language. The process started in January 2018 and is still on-going as of March 2020 when the results were obtained for this thesis. It is designed to save all GTFS static and real time data, and update the results once new data become available. Thus, data available to use in this thesis is an almost 100% sample from COTA's GTFS feed. It is not entirely 100% due to equipment or internet outages. The data collection send requests every 10 seconds, and COTA system will respond with locations of all vehicles that has drivers logged in. As of January 2020, more than 334 million distinct real-time bus GPS tracking records, over 1.6 million trips in the system were recorded, and more than 102 million distinct stop times are recorded. Due to the massive amount of data, raw GTFS Static and GTFS Real-Time data are converted to a MySQL relational database for storage using the structure outlined in previous section. This allows storage of the entire dataset while reducing overhead storage spaces.

Meanwhile, data processing process was implemented in late 2018, which detailed processing steps are provided in the next section. The data processing is scheduled to automatically run once a day, so that the research results, which are available on my website https://norman.cloud, will always reflect the latest data.

3.3 Data Processing

Generally speaking, the actual arrival time at a stop equals to scheduled departure time at first stop plus first stop delay time plus the sum of stop level travel times from the first stop plus any extra holding times. Therefore, this study examines travel time at stop level to reveal inconsistencies between scheduled and actual travel times under typical conditions, in order to address the concerns mentioned in the introduction. Again, studying travel times at stop level can help transit planners pinpoint which route sections are causing inconsistencies and allows transit agencies to strategically allocate their resources such as staff hours.

In addition, actual travel times can indicate when vehicles are catching up to schedule if the actual travel time is shorter than the scheduled travel time. On the other hand, if the actual travel time is longer than the scheduled time, the vehicle is falling behind. In the case of a trip being severely affected by a major disruption or break down, unlike the scheduled arrival time adherence data, travel time data from other sections of the route that are not affected can still be used in the analysis.

Currently, COTA does not report stop level travel time or schedule adherence data, thus stop level travel time data needs to be estimated first using the GPS coordinates reported by each bus, then the travel time data can be aggregated for the analyses.

3.3.1 Matching GPS to Trips

Since the dataset is very large, typical GIS software cannot process all the GPS points. An ad-hoc spatial join method was implemented to match the GPS data to the bus routes. Since a line can be represented as the limit of a series of points when the number approaches infinity, bus routes were converted to a series of points that are less than 1 meter apart for the estimation.

There are additional constraints to consider when matching GPS data to the shape file. Since the routes has a direction and vehicles operate from the beginning to the end of the line, GPS points cannot be spatially joined directly. GPS points had to be sorted by time and by scheduled trip, then the nearest route points were matched sequentially from the beginning of the line. Then, based on the nearest route points, the route distance can be calculated for each GPS point.

Due to the limitations of GPS precision, if a vehicle is off route by more than 50 meters it is considered to be in detour, and the corresponding data points are saved but not considered in this study, since this study only considers typical travel conditions.

3.3.2 Calculate Stop Times

Using the route distance and time spent at each point, the GPS points can be grouped into 3 categories, arriving at stop, at stop, and departing from stop. GPS points that are within the range from 30 meters before the stop to 15 meters after are considered to be within the stop boundary, since buses are required to stop at the designated sign and more than 1 bus can stop at the same stop at the same time.

For the at stop category, the recorded GPS points around stops are grouped based on above classification to calculate stop arrival and departure times. The arrival and departure time for each stop along the route are estimated using the earliest time and latest time at a location within the stop boundary where the vehicle spend the most time. Note, this is different from the GTFS schedule published by COTA, which doesn't model dwelling times at stops, since modeling dwell times can help identify and remove some outliers such as excessively long dwell times at stop. The reason for using the location where the vehicle spent most time is due to the case when buses report one or two GPS points within the stop boundary when arriving or departing the stop. However, when stationary, buses could send out multiple GPS points at the same location. This is illustrated in Figure 3.1 as an example, where the bus send out one GPS points while approaching and departing the stop respectively, and the bus send out multiple GPS points while stationary. If there are no recorded GPS points around a certain stop, the nearest recorded GPS point is used to determine stop arrival and departure times.

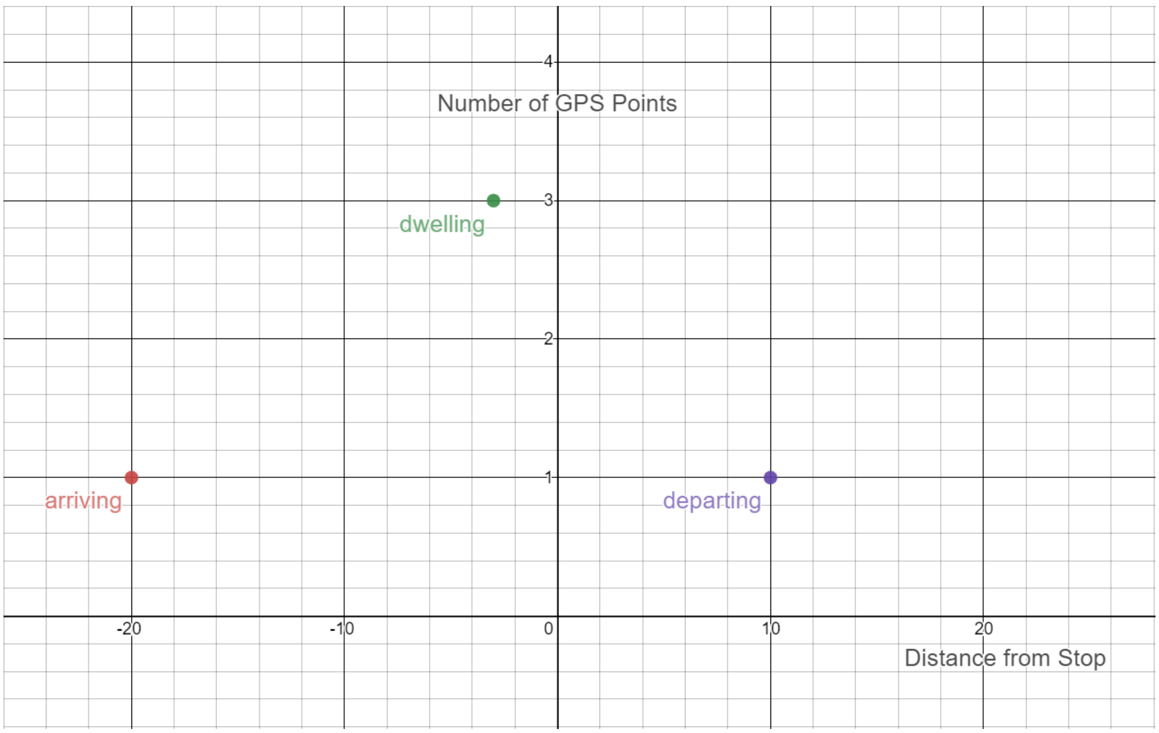

Figure 3.1: Example of Stop Time Calculation. X-axis represents the distance from the stop where negative numbers represent before stop and positive numbers represent after stop. Y-axis shows the number of GPS points at a given location.

3.3.3 Extract Travel Times

This step calculated the travel time between two stops using the estimated stop arrival times. Since times on COTA schedules are considered as arrival times and dwell times are added into the travel times, this thesis tries to be consistent with the agency's schedule when comparing the observed and the scheduled travel times. To be more precise, since COTA uses arrival based schedule, the scheduled travel time consists of the time between the arrival at stop 1 and the arrival at stop 2. The sequence of events for an arrival based system is that a bus arrives stop 1, dwells at the stop 1, departs from stop 1, and arrives at stop 2. Thus, travel time between two stops in this study is considered to be dwelling time at the first stop for boarding and alighting, plus run time to the second stop.

There are several potential ways for outliers to appear in the observed data. Since extreme outliers will skew the mean and this thesis is based on the sample means, this thesis performs an outlier removal process outlined in the next section. For COTA, buses are not allowed to leave earlier than the scheduled arrival time, and drivers are instructed to slow down or stop if they are running early to wait for the schedule to catch up. If drivers leave on time and purposefully avoid running early, the travel time would never be shorter than scheduled. Therefore, the arrival and dwelling times at each stop were calculated and compared using the outlier removal process outlined in the next subsection to remove data with excessive or unusual overall dwelling times. As it was not possible to determine how much of the observed dwell time pertained to the boarding and alighting process and how much pertained to potential holding after the completion of said process, the full dwell times were removed, including boarding and alighting times and potential holding times. Example of such cases are holding at certain stops when running early or assisting wheelchair users. Another case is the late-night line ups, where most bus lines meet downtown to provide transfer opportunities in order to accommodate for lower frequency. When calculating travel times, data with excessive dwelling times were also compared to scheduled holding times in order to reflect both travel times and scheduled holding times. In addition, cases like vehicle breakdowns or equipment outages will also show unusual stopping time at a location, and these records also contributes to the outliers and thus removed.

3.3.4 Outlier Removal

To reiterate, outliers might affect the mean, and this thesis is based on comparing observed sample means recorded from AVL data to the schedule. Due to major disruptions such as accidents, reroutes, and break downs, there are multiple outliers in the dataset. Major delays and unusually long travel times should also be removed, since the dispatcher might instruct the bus to run limited stops or express service in certain areas to recover from the delay. Since the goal of this study is to compare actual travel times under typical travel conditions against scheduled travel times, these outliers in travel times and dwelling times were removed from the analysis.

This study uses Density Based Spatial Clustering of Applications with Noise (DBSCAN) for this step. DBSCAN (Schubert et al., 2017) is a density-based clustering algorithm designed to identify clusters and outliers in the data. The algorithm identifies three types of points: core points, which have more than the minimum required points within a given distance; border points, which are within reach of core points but have less than minimum required points within the given distance; and outlier points, which are outliers that do not belong in any group. In short, given a set of inputs, the algorithm groups points that are close together and marks outliers in lower density areas where the nearest neighbour is too far away. The algorithm is set to keep at least 80% of the data as core and border points, which corresponds to having less than or equal to 20% outliers. The parameter choice and model sensitivity can be a subject of future research.

In this application, DBSCAN is used to remove outliers only; cases with multiple clusters are not considered. Pending further research, the meaning of more than 1 cluster should be examined. Figure 3.2 is an example plot of the DBSCAN algorithm applied to one stop pair. In the figure, x-axis is the recorded delay at first stop in seconds, and y-axis is the actual travel times in seconds. From the figure, we can observe that the algorithm marked data points that are not close to the dense core points as outliers, namely the points with relatively longer travel time or points with longer delays.

Although not shown here, excessive dwelling times can potentially be identified. For example, when holding happens, the bus would be running early where x-axis will be a negative number and spending excessive amount of time where y-axis will have a larger number, which make the data points closer to the top left corner further away from the dense core points. However, one potential issues is when holding at a certain stop is the norm for a given trip, in that case, holding cases will have more density, whereas the non-holding cases are considered as far away and removed. This case should be very rare, since most of operators avoid driving early since their bonuses depend on it. If this does happen, analysts can easily spot them by looking at the dwelling times directly, since they will be much longer than expected.

3.4 Data Aggregation for Travel Time Analysis

One advantage of using the relational database, mentioned in previous chapter, is that, the data can easily be aggregated in various ways, such as based on time of

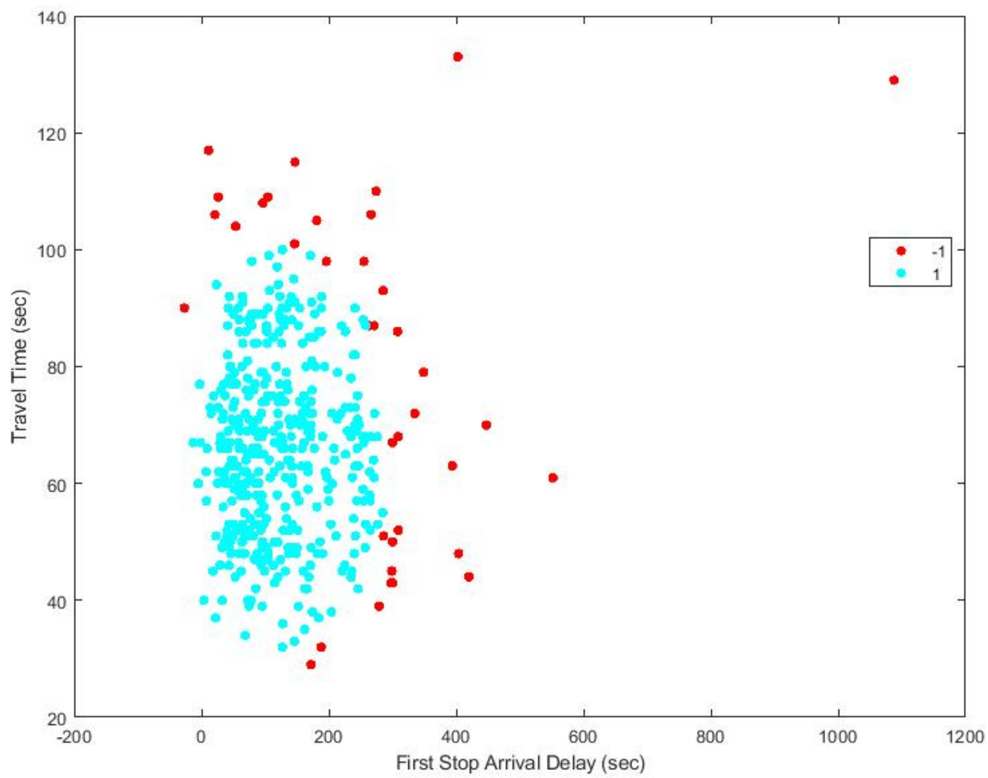

Figure 3.2: Example of outlier Removal Results, where class 1 represents the data used in the analyses and class -1 represent the outlier data not used in the analyses.

day (timestamp), service type (whether the service is a weekday, Saturday, or Sunday service), direction, or whether the line is an express service. In addition, a sequence of stops can also be aggregated to reflect timepoint to timepoint, or terminal to terminal travel times which are consistent with the service metrics at the agency.

This study analysed the data by conducting a two tailed T-test to check the hypothesis of whether the scheduled travel time is the actual travel time mean. To reiterate, mean travel times are used here for better theoretical justification, and outliers are removed so that the means will not be affected severely by outliers.

If the T-test fails to reject this hypothesis at 95% level, it is considered balanced travel time, meaning the statistical tests fails to reject the hypothesis that the scheduled mean is the actual population mean. If the T-test rejects the hypothesis and the scheduled travel time lie on the extreme ends of the sampling distribution, the scheduled travel time is considered surplus or insufficient based on whether the scheduled time is larger or smaller than the sample mean, respectively. Since a common on-time performance goal used by agencies is 85%, this study compares the scheduled travel time to an 85% upper bound in the cumulative distribution, which leaves 15% of relatively longer travel time observations on the tail. If the scheduled time is larger than the 85% upper bound, then the schedule is considered extreme surplus, since less than 15% of the actual travel time will be slower than scheduled. Similarly, for symmetry, 15% lower bound was chosen to reflect extreme schedule insufficiencies, since only 15% of the trips can adhere to the scheduled travel time.

3.4.1 Data Visualisation

To better illustrate the spatial component of the analysis, a data visualization tool was developed. The visualization is based on Google Map API and connected to the database server, which allows users to dynamically change the level of analysis and parameters of the data aggregation. There were four major steps to create this visualization: i) The trip and related stop pairs were selected from the database. ii) The actual travel times were aggregated and compared against the scheduled travel time. iii) Each stop pair was colour coded based on the statistics. iv) Each stop pair was then drawn on the maps. Currently, the website allows users to query for trip level analysis and system wide analysis. All stop pairs within the analysis scope will be colour coded and displayed on the website. The colours range from green to yellow to red, indicating surplus, adequate, and insufficient travel times respectively. If the actual travel times is significantly slower than scheduled travel time, the section between the stop pair is coloured as red. Similarly, if the typical travel time is significantly faster than scheduled travel time, the section is coloured as green. The section is coloured as yellow if the difference is not significant. For the screenshots of the website, please refer to the results and discussion section where detailed interpretations of the results are presented.

Chapter 4: Travel Time Analysis

There are several aspects relating to passenger experiences and evaluates transit services, whether the bus arrives on time is one of the aspects. Passengers will expect a bus to show up at a given stop when it is the scheduled arrival time. As of now, COTA measures and adjusts its service based on timepoint level on-time performance data. However, passengers don't necessarily wait for a bus at timepoints (Halvorsen et al., 2019). There are also other aspects that are crucial to passengers experiences and agency operations, such as travel time and transfer time. This chapter will focus on the travel times at stop to stop level in addition to their existing service metrics, i.e. on time performance at timepoints.

For agencies, on-time performance will not tell the whole picture on whether the current schedule is adequate, since it does not tell the extent of poor performances (Halvorsen et al., 2019). For example, operators might leave the terminal late for various reasons, such as maintenance issue or when the drivers with line knowledge are confident that they will catch up to the schedule. However, since COTA does not allow operators to run ahead of schedule, it is also not adequate to tell whether the schedule is sufficient if operators consistently have to slow down significantly between timepoints to kill time and wait for the schedule to catch up.

To passengers, whether in vehicle travel time is longer or shorter than scheduled travel time is an important indicator for service reliability (Carrel et al., 2013). Schedule is less relevant when headways are short, since there will be another bus to pick up the passenger shortly. However, since the most frequent headway for COTA is 10 minutes which happens during weekday rush hours only, bus lines from COTA cannot be considered as frequent enough for schedules to be less important. In addition, real time information published by COTA may also influence passengers' perceived travel time. However, since it requires additional inputs from riders this will not be discussed in this thesis.

For agencies, if the observed in vehicle travel time is longer, then the trip is further delayed. On the other hand, if the observed in vehicle travel time is shorter, buses will catch up the schedule. In the cases where travel time is extremely insufficient, the agency might need to add additional vehicles to the line to adhere to their schedule. On the other hand, if travel times are more than sufficient for drivers to complete the route, agencies might be able to remove vehicles from the line.

The following sections uses COTA Line 31 as an example to show different level of analysis, since the entire system consists of 43 lines and more than 3000 stops, writing them all in one report would be unrealistic. This line is chosen since it is one of the newly designed routes after TSR. It mostly combined sections from 4 pre-TSR crosstown routes mainly on the western half of the line, namely 80, 81, 82, and 84, whereas the eastern half of the line is brand new. Also, additional information were obtained from the drivers operating on this line that may help understand the result presented here. Similar analyses for other lines or stops can follow the frameworks and steps outlines below.

To reiterate the data section, the stop level arrival times are produced by proprietary software used by COTA after the scheduling process which is at timepoint level. In addition, it is reasonable to speculate that the stop arrival times are linearly interpolated, since stop to stop segments between two timepoints tend to have the same travel speed. It is worth to note that the agency measures its performance only at timepoints, and may not hold them to the stop to stop schedules. However, the schedules shown on the agency's website also use these stop level arrival times. Passengers can easily interpret these stop arrival times as schedules the agency adheres to. In addition, the AVL system used by dispatchers and the drivers will categorise a vehicle and react accordingly based on these generated stop level arrival times. Despite these stop arrival times are not determined at by schedulers, it is generally treated as a schedule by passengers and the agency. Thus, this thesis will refer to them as scheduled arrival times. Since these stop arrival times are generated using unknown algorithms from a commercial software, the results or interpretations may differ from the software's intentions.

This chapter first looks at the inconsistencies between the scheduled travel time and the actual travel times at stop to stop level, trip level, line level, and system level. This will indicate insufficiency patterns that agencies might want to pay closer attention. Second, the chapter examines how these stop to stop travel times are distributed. This could aid agencies in finding a more suitable statistical distribution that can be used in their scheduling process. Next, the chapter explorers how are these stop to stop travel times correlate. The results could help agencies in identifying priorities when planning for their services. Finally, the chapter analyses the conditional travel times to see what travel times can be reasonably achieved if drivers don't slow down when they run early.

4.1 Travel Time Inconsistencies Across Time and Space

4.1.1 Stop to Stop Level

The algorithm calculates and compare observed mean travel times to scheduled travel time for all sections using t-tests at 95% confidence level where the null hypothesis being tested is that the observed mean travel time for a given trip equals the scheduled travel time, whereas the alternative hypothesis is that the observed mean travel time does not equal to the scheduled travel time. Since COTA does not have seasonal schedules, the calculation uses all non-outlier data observed from past two years. Although analysts could look at seasonal variations using the system, seasonal or even monthly aggregation is not shown here. Here is an example stop pair to show the findings, from King Ave and Delashmut Ave to King Ave and Olentangy River Rd.

This stop pair is located near Ohio State University campus and runs through Lennox Town Center, which is a main shopping centre in the campus area. King Ave is classified as a major collector, Olentangy River Road is classified as a minor arterial with highway access to State Route 315, which is classified as other freeway and expressway, and Delashmut Ave is classified as local. The entrance to Lennox Town Center is located after the Olentangy River Road stop. There are no traffic signals or stop signs between the two stops. (ODOT, 2016)

To reiterate, generally speaking, transit terminal to terminal travel time equals the sum of stop to stop travel times and dwelling times at each stop. Thus, in theory, these random variables needs to add up to represent the terminal to terminal travel time as well. The means can be summed up regardless of its distribution, whereas the percentiles cannot be summed up to equal a given percentile. In addition, refer to the scheduling practices mentioned in literature section, current scheduling literature focus more on the terminal to terminal travel times (Kimpel et al., 2008). These literature also pointed out that 85th percentile is the half cycle time, which includes the terminal to terminal travel time and the layover or recovery time. The terminal to terminal travel time is generally picked from the mean or the median, and the layover time can be determined from the standard deviation, where adding one standard deviation is assumed to yield 85 percentiles, or a given percentiles specified by the labour contract or agency standards (Kimpel et al., 2008). This study chooses to focus on the mean, based on its additivity and percentiles are provided as a comparison to previous literature. To make these results generalisable and due to the lack of information, scheduling practices at COTA, such as the practice of using mean travel time and using high percentile to calculate layover time, are also assumed to be the same as mentioned in previous literature.

Weekday trips for this stop pair are selected and analysed by trip. Table 4.1 shows the results returned from the database for the analysis, including the scheduled arrival time, sample size, mean, standard deviation for actual travel times, scheduled travel time, and the t-statistic. Additional information is provided on the 15th, 50th, and 85th percentiles of the sample. Again, the percentiles are provided as a reference for current practices. Actual travel time sample mean and scheduled travel time for each trip are plotted in Figure 4.1. The x-axis illustrates the trip, sorted by time, and y-axis shows the travel time in seconds.

A more general classification scheme is also provided here in Table 4.1. The classification is determined based on the t-statistics and the commonly used 85th percentile. The extreme cases are determined by the 85th percentile, extremely sufficient if 85 percent of travel times are less than scheduled, extremely insufficient if 85 percent of travel times are longer than scheduled. If the travel times don't fall into the extreme category and t-statistics are able to reject the null hypothesis mentioned earlier, the segment falls into the sufficient or insufficient categories, sufficient where t-statistic is positive, insufficient where t-statistic is negative. Lastly, if t-statistic is unable to reject the null hypothesis, the segment is categorised as balanced.

Table 4.1 shows that for 33 out of 36 trips, the actual travel time sample mean is lower than the scheduled travel time. Also, scheduled travel time is longer than the actual travel time sample mean plus one standard deviation for 27 out of 36 trips. Based on the figure and t-statistics in the table, it is shown that actual travel time between the two stops are generally shorter than the scheduled travel time.

By comparing the sample standard deviations in the table, it shows that variation in observed travel time increases around AM PM peaks and lunch hours. These trips with larger standard deviations also coincide with larger sample means. The scheduled travel times are within one standard deviation from the actual travel times for these trips. Two trips have insufficient travel times, in which the sample mean is larger than the scheduled travel time. Especially for the 7:32 trip, the scheduled time is slightly less than the median, which indicates that less than half of the trips can operate within the given travel time. This indicates that the schedule becomes tighter, and thus may be harder to adhere to during peak time periods compared to off-peak hours.

Table 4.1: Mean Travel time vs Scheduled Travel Time.

Time	Sch	Ssize	Smean	SSD	Diff	Tstat	15PC.	50PC.	85PC.	Sch.PC.	Classification
05:06:53	37	389	27.2211	3.9031	-9.7789	-49.4142	24	27	31	100	ExSurplus
05:36:53	37	387	25.6460	4.8041	-11.3540	-46.4939	21	25	31	100	ExSurplus
06:06:04	45	402	26.0498	4.7708	-18.9502	-79.6409	22	26	29	100	ExSurplus
06:35:04	45	387	27.8992	5.1413	-17.1008	-65.4334	23	27	34	100	ExSurplus
07:03:15	52	407	31.9410	8.0431	-20.0590	-50.3134	24	31	41	99	ExSurplus
07:32:25	60	387	65.5866	30.1243	5.5866	3.6482	39	61	105	75	Insufficient
08:04:15	52	394	43.5381	19.5083	-8.4619	-8.6099	28	37	64	83	Surplus
08:35:04	45	397	34.8237	14.4266	-10.1763	-14.0548	22	31	46	90	Surplus
09:05:04	45	399	27.9749	7.0607	-17.0251	-48.1645	21	27	35	100	ExSurplus
09:34:04	45	408	30.1961	7.0871	-14.8039	-42.1929	23	29	38	100	ExSurplus
10:03:04	45	385	27.5247	6.2199	-17.4753	-55.1277	21	26	34	100	ExSurplus
10:33:04	45	398	29.3467	6.5002	-15.6533	-48.0420	23	29	36	100	ExSurplus
11:03:04	45	404	27.9480	7.0917	-17.0520	-48.3296	20	27	35	100	ExSurplus
11:33:04	45	389	27.4910	6.9859	-17.5090	-49.4330	20	27	35	100	ExSurplus
12:03:04	45	408	31.8554	10.1464	-13.1446	-26.1677	26	30	37	99	ExSurplus
12:33:04	45	398	45.9824	27.3642	0.9824	0.7162	22	31	42	96	Balanced
13:03:04	45	388	40.5026	20.0286	-4.4974	-4.4231	26	35	73	76	Surplus
13:32:04	45	408	32.4338	9.2138	-12.5662	-27.5482	26	33	62	81	ExSurplus
14:02:04	45	410	32.9146	8.7528	-12.0854	-27.9580	24	31	41	96	ExSurplus
14:31:15	52	395	31.6506	7.1570	-20.3494	-56.5090	25	31	41	96	ExSurplus
15:00:25	60	389	34.4422	9.9951	-25.5578	-50.4327	25	32	38	100	ExSurplus
15:29:25	60	395	40.4278	17.9622	-19.5722	-21.6559	26	33	43	92	ExSurplus
15:58:25	60	408	38.3186	13.6814	-21.6814	-32.0100	27	34	55	88	ExSurplus
16:28:25	60	387	38.3437	16.0534	-21.6563	-26.5382	25	29	40	92	ExSurplus
16:58:25	60	386	53.5259	32.8965	-6.4741	-3.8665	27	35	50	95	Surplus
17:31:15	52	393	74.0356	49.9244	22.0356	8.7500	26	34	51	93	Insufficient
18:01:15	52	391	45.1586	24.3439	-6.8414	-5.5571	30	39	92	81	Surplus
18:31:15	52	393	41.5954	19.8599	-10.4046	-10.3859	30	53	137	59	Surplus
19:02:15	52	406	32.0246	6.9987	-19.9754	-57.5094	27	36	64	87	ExSurplus
19:33:15	52	392	30.8724	7.1712	-21.1276	-58.3311	27	34	58	88	ExSurplus
20:04:04	45	389	29.5476	4.9975	-15.4524	-60.9842	26	31	40	100	ExSurplus
20:34:04	45	401	29.8579	5.6765	-15.1421	-53.4173	25	29	37	99	ExSurplus
21:04:04	45	403	30.8437	5.5001	-14.1563	-51.6689	25	29	34	100	ExSurplus
21:34:04	45	404	29.9554	5.2768	-15.0446	-57.3064	25	29	35	99	ExSurplus
22:04:04	45	399	31.1454	5.7724	-13.8546	-47.9431	26	30	37	100	ExSurplus
22:34:04	45	407	32.4079	10.6820	-12.5921	-23.7819	25	29	35	99	ExSurplus

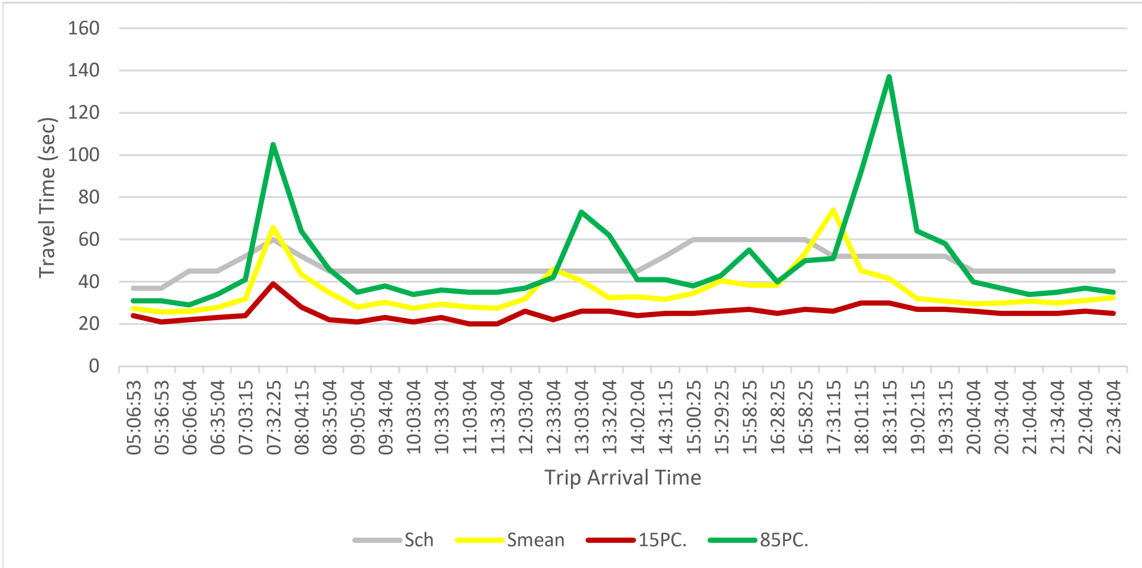

Figure 4.1: Illustration of Travel Time Inconsistencies Between King & Delashmut and King & Olentangy River Road.

Figure 4.1 shows another interesting pattern when comparing the sample mean to the scheduled travel time. For AM and PM peaks, the schedule generally reflects the increasing travel time during the time periods, although the schedule might not be increased enough in some cases. However, for the lunch hour, the schedule is not reflecting the increased travel time at all, although the t-statistics are generally less than 0. This makes it slightly harder for buses to adhere to the scheduled travel time.

Since most buses run faster than schedule, and the travel time distribution during the off-peak hours are relatively tight (with smaller standard deviations), the schedule can reallocate the surplus times to other sections of the route where the scheduled travel time tends to be insufficient. For the trips that might have tight schedules, the schedule can add more travel time to reflect the increase more accurately.

4.1.2 Trip Level

The trip 697167 is one of the scheduled trips on Line 31. It leaves the layover point at Grandview Yard at 6:46 on weekday mornings and heads northeast towards Easton Transit Center. It serves the city of Grandview Heights, The Ohio State University, MAPFRE Stadium, and Easton area.

By aggregating the actual travel times for each stop pair in this trip, and analysing their deviations from scheduled travel time, we can analyse trip level travel time deviation using Table 4.2. Here, only data between two timepoints are shown to illustrate the findings. For more complete trip level data, please refer to the appendix Table A.2. Again, based on the reasoning outlined in previous sections, this thesis focus on the mean and percentiles are provided as references.

The data visualization (Figure 4.2) shows inconsistencies between scheduled and actual travel times for most stop pairs along the route for this specific trip. For this specific trip, only 6 stop pairs out of 65 are classified as balanced. 14 stop pairs are classified as surplus, while 24 are classified as extreme surplus. 14 stop pairs are classified as insufficient while 7 are extremely insufficient.

Table 4.2: Sample Travel Data for Line 31 Trip 697167, where bold StopIDs are time points.

FromStop	ToStop	Sch.	S.Size	S.Mean	S.S.D.	Diff.	T-Score	15PC.	50PC.	85PC.	Sch.PC.
...											
GRA5THN	GRAKINGN	35	374	36.96	9.91	-1.96	3.83	27	35	48	50
GRAKINGN	KINSTAW	57	388	67.95	13.21	-10.95	16.33	53	66	84	21
KINSTAW	NORCHAN	41	390	49.59	13.44	-8.59	12.62	36	48	62	25
NORCHAN	NOSTARN	54	390	66.30	13.88	-12.30	17.50	50	66	84	23
NOSTARN	KIN1197E	84	403	61.02	10.56	22.98	-43.70	51	59	73	100
KIN1197E	KENKINS1	99	392	71.84	12.57	27.16	-42.78	58	71	86	98
KENKINS1	KENSTELS	39	393	50.40	12.27	-11.40	18.42	38	48	64	17
KENSTELS	KINKENE1	71	398	102.54	23.02	-31.54	27.33	79	104	127	7
KINKENE1	KINDELE	75	391	43.29	9.31	31.71	-67.34	34	41	53	101
...											
SUM		555		549.89	118.2			426	538	681	

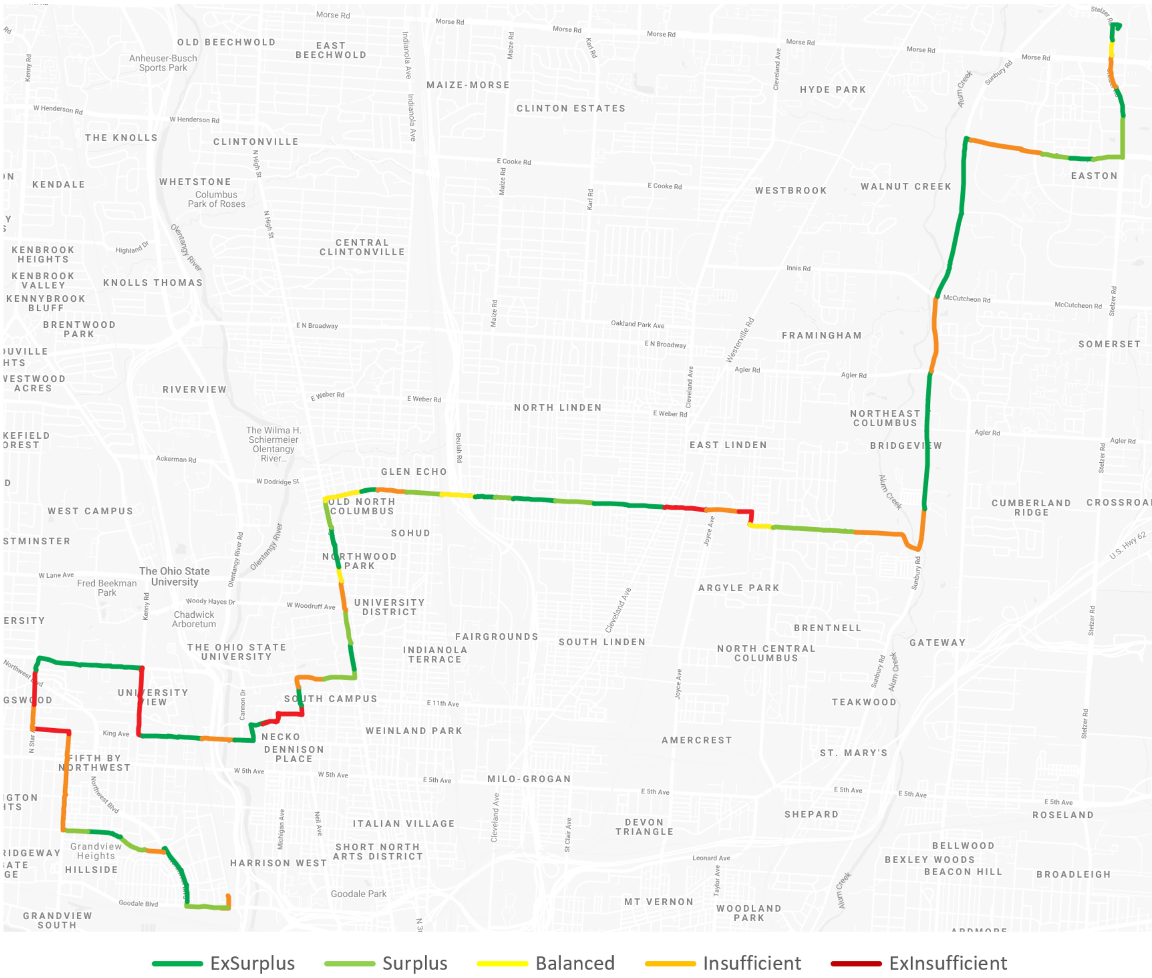

Figure 4.2: Trip Level Travel Time Difference Visualization for trip 697167.

In addition, by looking at the difference in seconds between scheduled travel time and sample mean in Table 4.2, the insufficiencies are quickly offset by surpluses right after for some sections. One example is the section between time point GRA5THN (Grandview Ave & W 5th Ave) to KINKENE1 (King Ave & Kenny Rd), the sum of the differences is just 5 seconds, but the deviations from scheduled travel times between these two timepoints are high. In several segments, the differences between scheduled and actual travel times are more than 20 seconds, which is a magnitude larger than the overall difference. Since current practices focus on terminal to terminal travel times or timepoint to timepoint travel times, the small overall deviation at timepoint to timepoint level and the larger differences at stop to stop level further imply that traditional scheduling done at route or timepoint level is not precise enough to reduce inconsistencies at stop to stop level. Analysts in transit agencies might have to focus on the stop to stop level data, if available, in order to reduce these stop to stop level inconsistencies.

Furthermore, the sum of 85th percentiles is greater than the sum of means plus one standard deviation, assuming these stop to stop segments are independently distributed. This shows again that percentiles cannot be added up to equal a given percentile, which highlights the potential issues when analysing stop to stop travel times using percentiles.

One important aspect to note is that number of segments in each classification alone cannot be used to describe overall travel time surplus or insufficiency for the whole trip. Overall loses could be smaller than the overall time gained in surplus sections, and vise versa. In the example used above, there are 6 insufficient segments where the trip loses time and 3 surplus segments where the trip catches up. However, the sum of gains is larger the sum of loses, making the overall mean travel time less than the scheduled time.

The above observations contribute to a high variation across segments in delay times, due to vehicles catching up in one section while lagging behind in another. Although the number of surplus stop pairs is larger than the insufficient stop pairs, this does not necessarily mean the vehicles run early since the sum of insufficiencies could be longer than the sum of surplus travel times for specific trips due to population standard deviations in actual travel times. Further research is required to produce a recommended stop arrival time for scheduling purposes.

There are clusters of sections where the scheduled travel times are insufficient, meaning scheduled travel time is less than actual travel times, near the areas of Fifth by Northwest, University View, Ohio State University South Campus area, and Arlington Park. The characteristics of these insufficient sections are in close proximity to major intersections through closer examination on the map.

Since the time points are often located at major intersections where there are also transfer points to other routes, vehicles often have longer travel times than scheduled near the timepoints. Actual travel times at time points have insufficient travel times, either before or after eight out of eleven time points.

Opposingly, the surplus sections often lie between two major intersections. This suggests that scheduled travel time is longer than the actual travel times between major intersections. This makes sense, since most of stops are located between major intersections, and there are fewer major intersections than the regular local intersections due to road hierarchy. These observations will be examined further in later sections to see if there is a systematic pattern.

Based on above information, schedules can be adjusted to reflect this travel time pattern by adding more time near major intersections and reducing time between major intersections. This also proves the drawbacks of analysing travel time at time point level which does not reflect the travel times between stops very well. Vehicles are left behind schedule at major intersections where timepoints are typically located , but they catch up in between two timepoints. Therefore, time point level analysis does not show travel time variation precisely and does not show underlying travel time variation pattern.

4.1.3 Line Level

Again, using the Line 31 as example, the trip level travel times analyses are aggregated into the line level analysis based on terminal departure time. Based on the classification obtained for each trip, the numbers of stop pairs in each class were plotted in Figure 4.3 and shown in Table 4.3. For example, the 6:46 trip used throughout this thesis has 24 extreme surplus segments, 14 surplus segments, 4 balanced segments, 17 insufficient segments, and 8 extreme insufficient segments.

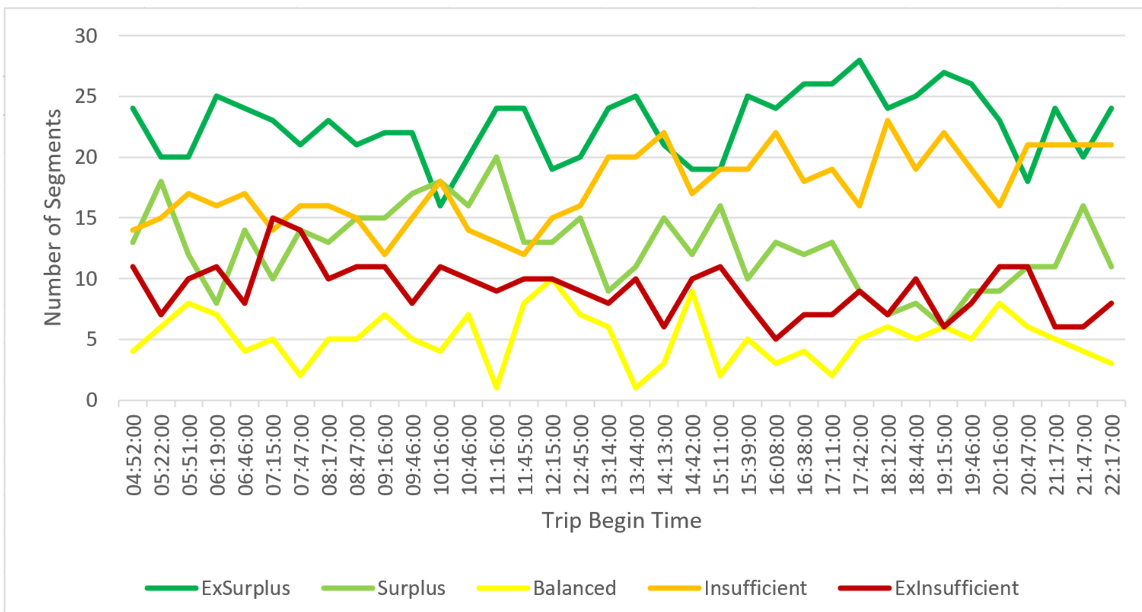

Figure 4.3: Line 31 Northeast Weekday Trip Classification Plot.

Just like the trip level analysis specifically for trip 697167, unsurprisingly, travel time for other trips in the day shows similar trend. The extreme surplus sections,

Table 4.3: Line 31 Northeast Weekday Stop Classification Counts.

Time	ExSurplus	Surplus	Balanced	Insufficient	ExInsufficient
04:52:00	24	13	4	14	11
05:22:00	20	18	6	15	7
05:51:00	20	12	8	17	10
06:19:00	25	8	7	16	11
06:46:00	24	14	4	17	8
07:15:00	23	10	5	14	15
07:47:00	21	14	2	16	14
08:17:00	23	13	5	16	10
08:47:00	21	15	5	15	11
09:16:00	22	15	7	12	11
09:46:00	22	17	5	15	8
10:16:00	16	18	4	18	11
10:46:00	20	16	7	14	10
11:16:00	24	20	1	13	9
11:45:00	24	13	8	12	10
12:15:00	19	13	10	15	10
12:45:00	20	15	7	16	9
13:14:00	24	9	6	20	8
13:44:00	25	11	1	20	10
14:13:00	21	15	3	22	6
14:42:00	19	12	9	17	10
15:11:00	19	16	2	19	11
15:39:00	25	10	5	19	8
16:08:00	24	13	3	22	5
16:38:00	26	12	4	18	7
17:11:00	26	13	2	19	7
17:42:00	28	9	5	16	9
18:12:00	24	7	6	23	7
18:44:00	25	8	5	19	10
19:15:00	27	6	6	22	6
19:46:00	26	9	5	19	8
20:16:00	23	9	8	16	11
20:47:00	18	11	6	21	11
21:17:00	24	11	5	21	6
21:47:00	20	16	4	21	6
22:17:00	24	11	3	21	8

surplus sections, and insufficient sections are of the same magnitude, while balanced sections and extreme insufficient sections are of the same magnitude.

Numbers of segments in each category vary quite quickly. Since there are 18 drivers operating the line on a weekday, there is a chance that driver behaviours might influence the travel time inconsistency. Due to the lack of data, the operators' preferences, such as departing the terminal late or driving unnecessarily slowly when running ahead of schedule, were ignored in this analysis. However, this table does show a general trend in travel time inconsistencies. Although operator behaviours were touched on by previous literature, operators' behaviour at stop to stop level should be further studied as an improvement to this study.

Another reason for the sudden variations on line 31 specifically might be due to class changes. Since this line runs through Ohio State University campus, and during class change, there are lots of pedestrian traffic. These large pedestrian traffic volumes would impede bus travel time significantly. Also, during class change, like rush hours, there are more riders catching the bus to different destinations, which would increase the dwelling time at each stop on campus.

Generally, this line has more extreme surplus sections than other categories. The number of extreme surplus sections decreased during the morning peak, while the numbers in other classes increased. Another decrease in the number of extreme surplus sections occurred during lunch hour, where the number of extreme insufficient and balanced sections increased. During the evening peak, the number of extreme surplus numbers decreased again. However, during the PM peak, the insufficient sections are the most common. This shift towards insufficiency during the morning and evening peak hours suggests that once again, like the stop level analysis, the scheduled travel times do not increase enough compared to the increase in actual travel times. This means that, during the peak hours, buses are less likely to recover from delays.

Another observation from the data, which is not shown here due to limited space, is that insufficient stop pairs during off peak hours tend to remain insufficient during peak hours. This is consistent with the analyses in previous sections, which also shows that schedules do not increase enough during peak hours, and schedules do not increase enough for insufficient stop pairs to become balanced or even surplus stop pairs.

4.1.4 System Level

For the system level analysis, all the lines and trips that were scheduled to operate were analysed similar to the trip level and stop level analysis, then aggregated by hour. Figure 4.4, shows the percent of stop pairs in each category by hour, where Table 4.4 shows the actual counts in each category by hour. For example, there are 516 segments classified as extreme insufficient as shown in the table, which corresponds to roughly 20% of segments in the system shown in the figure around 6 am.

The results show the observations in line level and trip level analyses still apply to the extreme surplus cases. The number of extreme surplus stop pairs are the most common type. The extreme surplus cases decrease, and the extreme insufficient cases increase during AM and PM peaks. Again, this suggests that during peak hours the travel time did not increase enough, and vehicles have less surplus travel times for delay recovery.

Another interesting fact shown here is that after the PM peak, the number of insufficient and extreme insufficient stop pairs decreased and the extreme surplus cases

Table 4.4: Number of stop pair classifications trend for weekday services, where each table cell represent the amount of segments classified in a given category at a given hour.

Hour	ExSurplus	Surplus	Balanced	Insufficient	ExInsufficient
4	294	87	22	64	51
5	1484	389	127	310	249
6	1216	395	135	372	516
7	1391	396	145	419	564
8	1282	428	152	414	622
9	1118	441	135	426	637
10	1501	423	154	418	487
11	1341	407	144	435	566
12	1243	432	134	472	685
13	1406	453	182	430	514
14	1197	474	137	440	655
15	1201	441	139	465	818
16	1286	416	143	448	789
17	1087	413	137	434	928
18	1532	432	124	431	590
19	1616	395	141	343	398
20	1526	440	131	358	443
21	1581	441	128	365	305
22	1472	453	131	319	263
23	1480	404	118	294	186
24	634	194	56	104	71

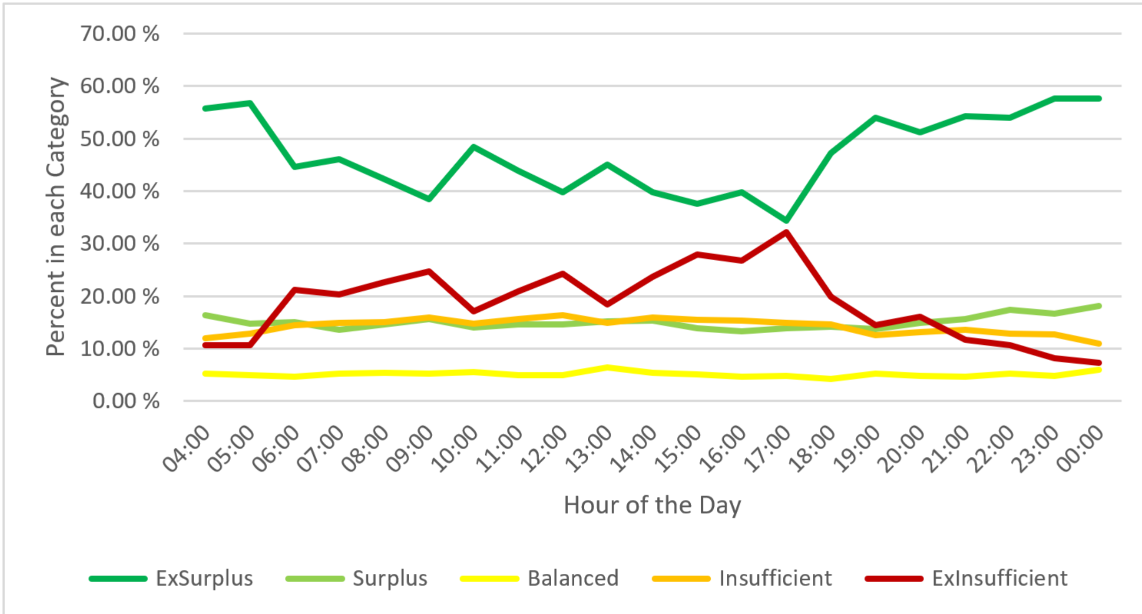

Figure 4.4: Systemwide stop to stop level travel time classification count plot.

increased, while the other cases remained relatively the same. This suggests that the actual travel time decreases faster than the scheduled travel time during those hours.

The system wide visualizations, Figure 4.5 and Figure 4.6 also reinforce a result from trip level analysis, which is that vehicles tend to run slower than scheduled travel time near major intersections, especially where two bus lines intersect. This visualization also shows that at 5pm, most of the southern part of downtown is classified as extreme insufficient. One possibility is that commuting traffic trying to get access to the highway on ramps located in the southern edge of downtown.

By contrasting Figures 4.5 and 4.6, which show systemwide classification at noon and 5 pm respectively, we can visualise the facts observed from the classification count shown in Figure 4.4. The number of red extreme insufficient cases increase, resulting in a map that is more red. Also, it also point out a fact that if a segment is classified

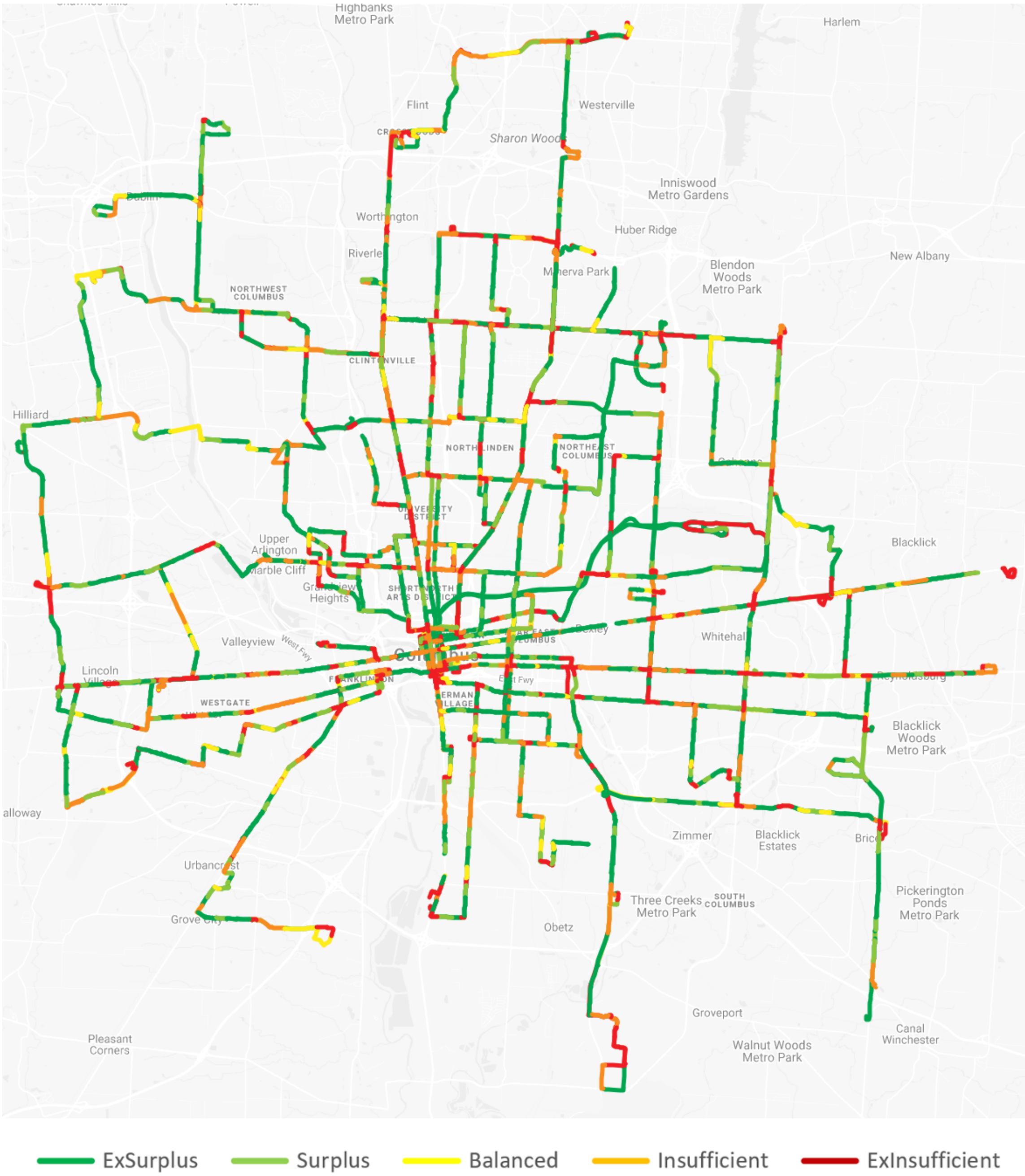

Figure 4.5: Systemwide Stop to Stop Level Travel Time Visualization for Weekdays at Noon, where each segment is shown on the map.

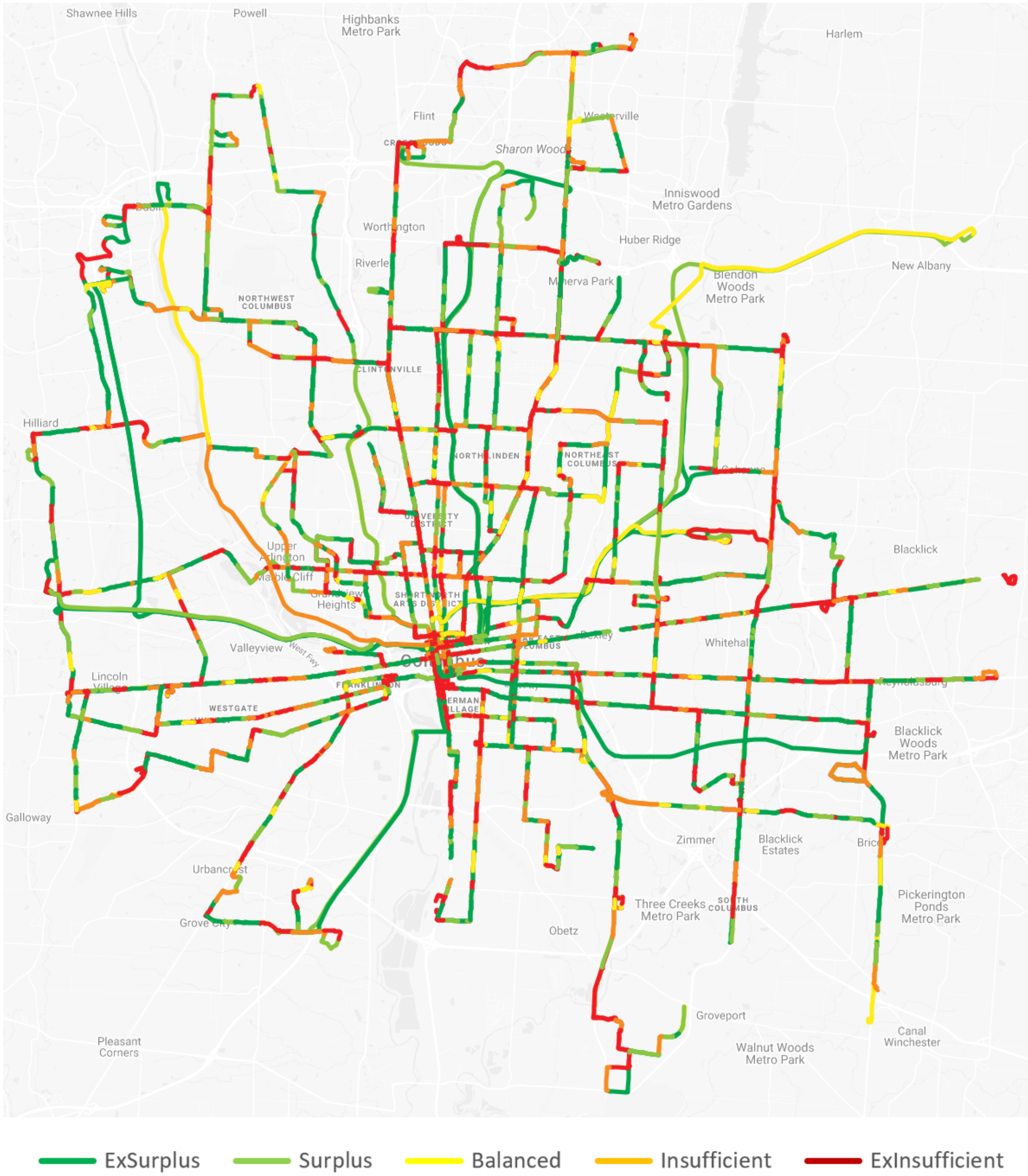

Figure 4.6: Systemwide Stop to Stop Level Travel Time Visualization for Weekdays at 5 PM, where each segment is shown on the map.

as extreme insufficient, it is very likely to stay that way even with the additional travel times allocated to the schedule. Green sufficient sections during off peak hours are also likely to shift to insufficient classifications. This is consistent with the spike of the number of extreme insufficient segments during PM peak, due to the overall shift towards insufficiency.

In conclusion, transit agencies and their software vendors might want to adopt more detailed analysis to provide a more robust transit schedule. Current practices have led to mostly adequate terminal to terminal travel times in our examples. However, the stop to stop level travel times still have larger deviations from the observed data. With more trip planning apps showing passengers arrival times at stop level, it becomes more important for agencies to adhere to their stop level arrival and travel times for passengers as well. In addition, since summing up stop to stop travel times and stop dwelling times should yield terminal to terminal travel time, and percentiles don't add up to percentiles, agencies need a more statistically sound way of providing stop arrival times. As a result, analysts might potentially want to look at means and their confidence interval or credible intervals when assessing travel times, and look at variances when determining layover time. This implication is also consistent with previous literature by Levinson (1991).

4.2 Stop to Stop Travel Time Distribution

Typically, agencies determines its layover times by using a high terminal to terminal travel time percentile to ensure most vehicles can start their next trip on time (Kimpel et al., 2008). Previous research focused on fitting terminal to terminal travel times using either symmetrical distributions or skewed distributions (Mazloumi et al., 2010).

However, as seen from previous sections, more aggregated view, terminal to terminal, masks large variations at smaller scale, the stop to stop level. Since mean, variance, and empirical percentiles do not describe skewness or potentially multimodality, fitting stop to stop level travel time distributions would support agencies if they decide to use percentiles similar to Muller and Furth (2001) and allow researchers to understand travel time variations more clearly. In addition, transit agencies also need to account for travel time distributions to determine their schedule, especially when planning for slack times to account for larger variations at certain segments. Finally, researchers or agencies might be able to use these theoretical distributions to develop simulations that help further analyse different operation strategies (Kieu, 2015).

As the central limit theorem points out that the sum of independent random samples tends to converge to normal distribution even when original variables are not normally distributed, making normal distribution a good approximation. However, as the previous literature pointed out in the literature review section, the travel times between terminals and timepoints are highly skewed (Mazloumi et al., 2010). Here, this study examines whether travel times for stop to stop level travel times recorded from COTA are also skewed. Refer to the literature reviews, previous studies have mainly examined uni-modal skewed distributions without providing comparisons among the available skewed distributions, and the multimodal tests shows the majority of the corridor travel times do not show significant multimodality (Kieu, 2015). In addition, previous literature have not used mixture distributions to model stop to stop travel times. Here, additional skewed distributions and mixture distributions were fitted and compared to obtain better fittings to stop to stop travel times.

Also, due to the higher data resolution, we can observe some more details, such as buses waiting for traffic lights. A bus get stopped by traffic lights is a random event, which depends on how long a bus wait at the traffic light, and resulting travel times between stops would be different due to different events related to the traffic light or other common random events along the line.

Mixture distributions are commonly used to model a random variable selected by chance from a collection of distributions according to given probabilities of selection, then the value of the selected random variable is realized. In the travel time case, travel times might be modeled as a random variable selected by the probability of whether the traffic is red or green, which would result in different travel time distributions. This lead to another question, whether mixture distributions provides relatively better fit for these stop to stop travel times.

Data after the outlier removal process, mentioned in Chapter 3, were selected from the database, and were fitted with various distributions. Common distributions that were used in this analysis are normal, logistic, generalised extreme value, generalised Pareto, log-logistic, log-normal, epsilon skewed normal, uniform, and mixture distributions with 2 and 3 normal distributions. Obviously, there are some common distributions that are not defined nor suitable to model the travel time distribution, and these distributions were not fitted in this study. One example of such distribution is exponential distribution, which measures the time between events that occur independently and continuously at a certain rate (i.e. a Poisson process).

To compare each distribution fit, I calculated the log likelihood. However, due to differences in parameter numbers among these models, log likelihood function will tend to show better results for larger models. Therefore, I also need to use Akaike

Information Criterion (AIC) to penalise the additional parameters in larger models when comparing them. In other words, AIC will deal with both overfitting and underfitting of data, and provides a relative measure for the quality of models being considered. Thus, AIC values were used when ranking these fitted distributions.

The smaller the AIC, the better a model is compared to other models. Here I focus on common distributions that provides best AIC for different stops. Most common best fit distributions are Generalised Extreme Value (GEV), Epsilon Skewed Normal (ESN), and Mixture distributions with 2 and 3 normal distributions (refer to as BiNormal and TriNormal distributions for simplicity).

Using Trip 697167 as example again, the results, listed in Table 4.5, show consistent results to previous literature that the travel time distribution is highly skewed. 58% of the stop pairs falls into the single-mode skewed distribution categories, represented mostly by the Generalised Extreme Value and Epsilon Skewed Normal distributions, whereas 39% of the stop segments fall into the mixture distribution categories. For more detailed AIC and Negative Log-Likelihood results for this trip, please refer to the appendix Table A.1.

Table 4.5: Travel Time Distribution Counts for Trip 697167.

TripId	Name	NumberOfDist
697167	GEV	17
697167	ESN	22
697167	BiNormal	10
697167	TriNormal	16
697167	LogLogistics	1
697167	Uniform	1

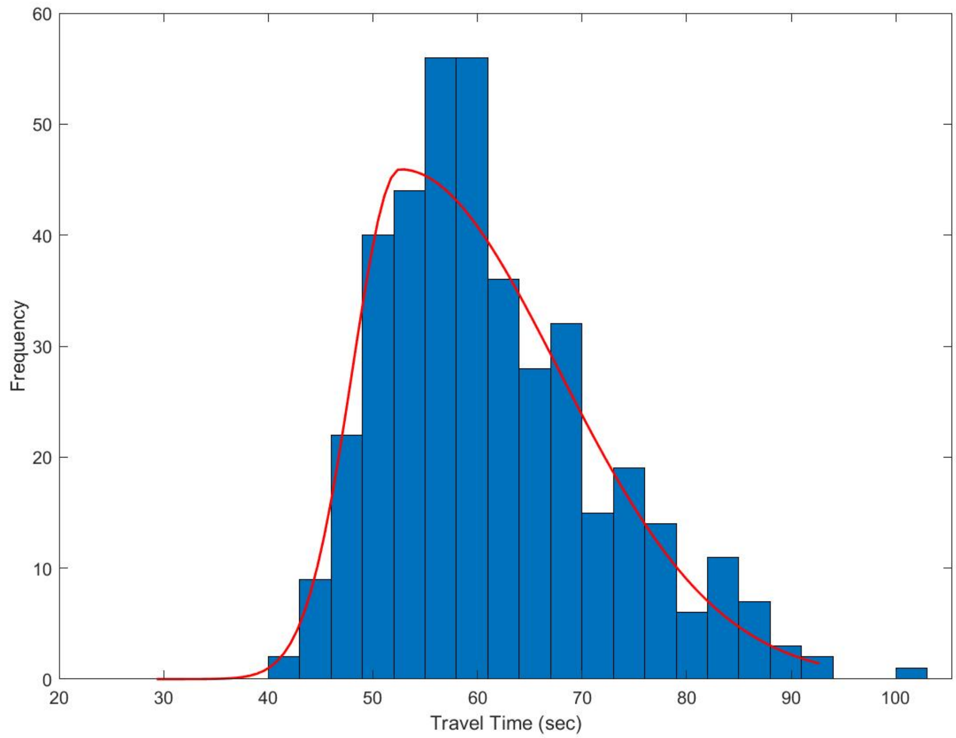

Figure 4.7: Travel Time Histogram and Fitted PDF for Trip 697167 from North Star & Presidential to 1197 Kinnear.

Similar to previous research, the majority of data do show significant skewness. To demonstrate the skewness of the distribution, one example of single-mode skewed data is shown in Figure 4.7, where the x-axis is the travel time, y-axis is the counts in one bin, the blue bars are the histogram, and the red line represents the probability density function fitted using epsilon skewed normal distribution. This data is observed from stop NOSTARN (North Star Rd & Presidential Dr) to stop KIN1197E (1197 Kinnear Rd). The sample mean of the observed data is 66 seconds. However, from the histogram, we can see that the mode is around 56 to 61 seconds, which is less than the mean. This shows a positive skewed distribution (i.e. a distribution with fatter right tail).

In addition to previous research findings, there is also cases where mixture distributions provide better fits. To demonstrate the multimodality, one example is shown in Figure 4.8, where again the x-axis is the travel time, y-axis is the counts in one bin, the blue bars are the histogram, and the red line represents the probability density function fitted using TriNormal distribution. In this case, we can observe two larger split point, one around 50 seconds, and another around 90 seconds.

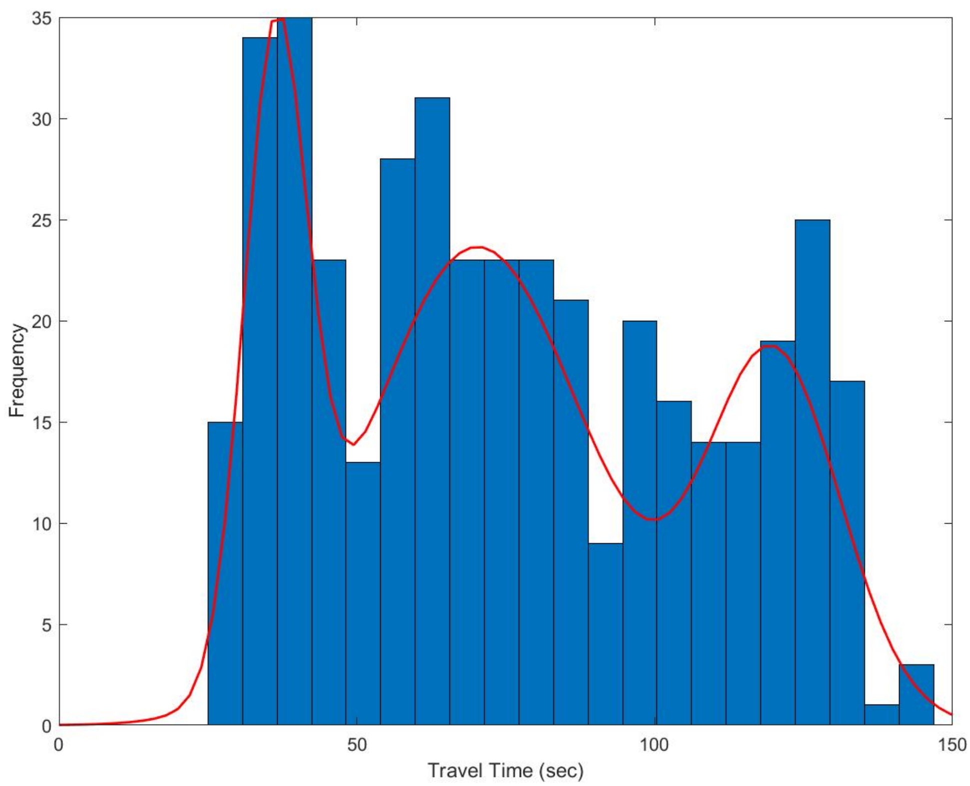

Figure 4.8: Travel Time Histogram and Fitted PDF for Trip 697167 from King & Olentangy River Road to King & Olentangy Trail.

One major difference between these two segments is that the single-mode case doesn't have traffic signals in between, since the intersection is handled by a roundabout. On the other hand, for the tri-normal case, it has one traffic signal in between. The traffic signal in the tri-normal case is a rather complicated one, which contains dedicated left turn signals on all four directions if there are enough vehicle to trigger the traffic sensor. It also has pedestrian light buttons, which will extend the green lights on one street to allow pedestrian crossing. The length of red lights are also dependent on the traffic flow by directions. It would make sense to use mixture distribution in this case.

The above observation leads to a further question, where are these distributions located at, and does traffic signals create these mixture distributions? To answer this, I put these distributions on the map, shown in Figure 4.9. In this map, the different distributions are coded in the following way: GEV (Dark Green), Epsilon Skewed Normal (Light Green), Other (Yellow), BiNormal (Orange), TriNormal (Red).

From the map, the answer is not so obvious. Major intersections, such as the tri-normal example above, Hudson and Summit, and Easton area does have mixture distributions. However, some other major intersections, such as High and Lane, has a single mode, but skewed distribution. For segments between intersections, they are not necessarily all belong to the two single-mode skewed distribution, such as the segments on Hudson.

As for the system level stop to stop travel time distributions, the best fit distributions were aggregated by hour, and the percentages for each distribution are shown in Table 4.6 and are plotted in Figure 4.10. Again, percentages are used here to account for difference in service levels across hours. For example, there are 31.76% segments

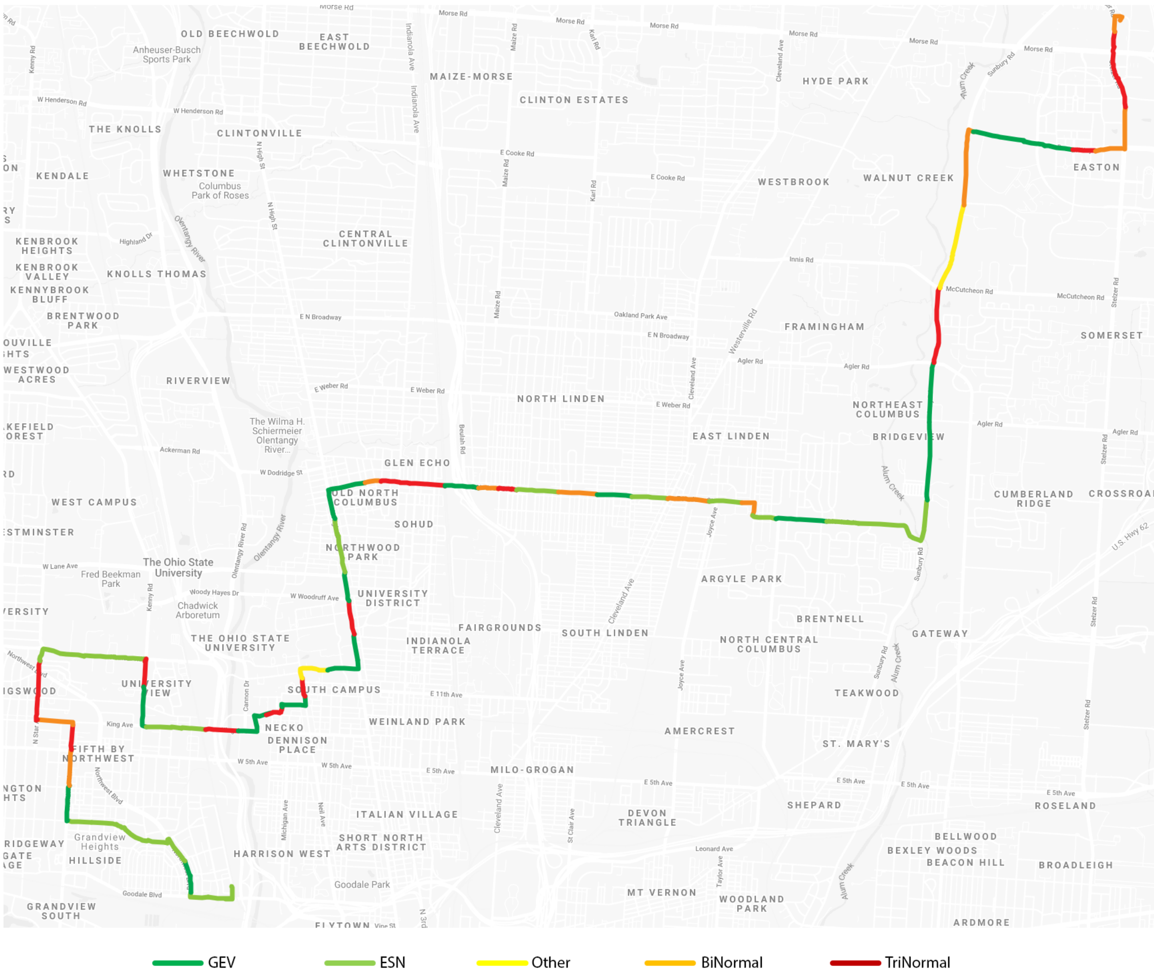

Figure 4.9: Stop to Stop Travel Time Distribution for Trip 697167.

with ESN as its best fit distribution at 4 AM, as shown in the table, which corresponds to the left end of the light green line that represents ESN.

Once again, the result show the stop to stop travel times are not normally distributed, since none of the best fit distribution is normal distribution. Single mode distributions are predominant, while tri-modal mixture distributions also accounts for a significant share. Overall, there are five distributions that accounts for more than five percent of the distributions, namely ESN, GEV, TriNormal, BiNormal, and Uniform distributions. All other percentages, except the top three, TriNormal, ESN, and GEV, remain roughly the same throughout the day.

ESN distribution is shown to be the best fit regardless of hour. ESN and GEV alone accounts for more than 60 percent of travel times. Interestingly, from the plot, we can observe that the percentage of ESN and GEV distributions are going the opposite way during morning peak. More precisely, the percentage of ESN distribution increases, while the percentage of GEV decreases from system opening to the end of morning peak. However, after 9 AM, the percentage of two distributions, ESN and GEV, have the same trend. The reason for why this is the case is left for future research.

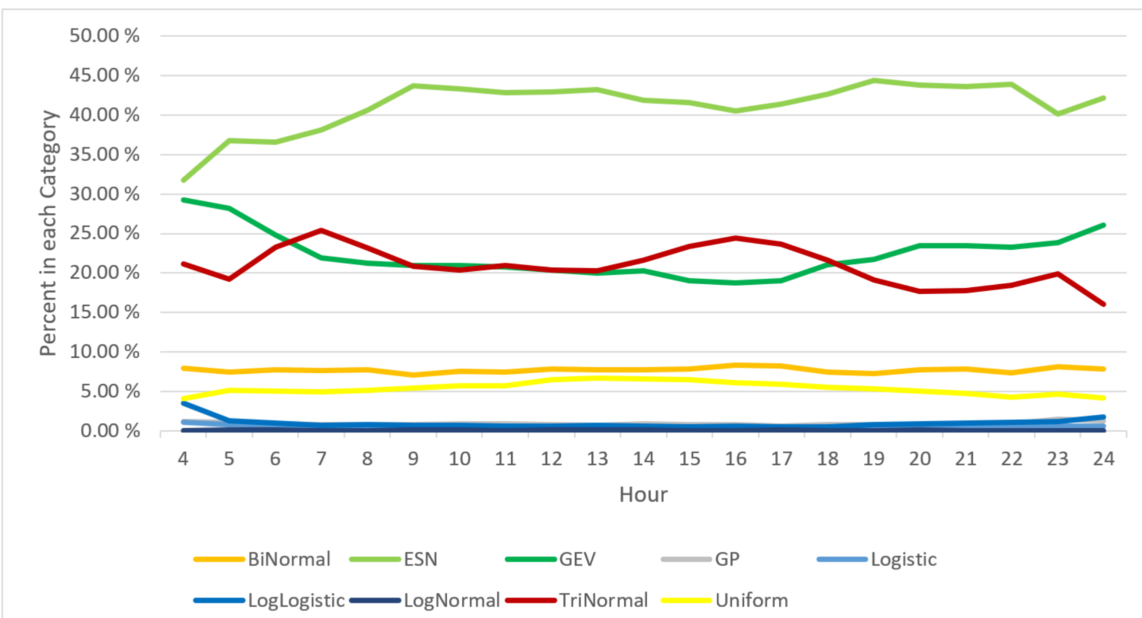

Figure 4.10: Travel Time Distribution Plot By Hour.

Table 4.6: Travel Time Distribution Counts by Hour.

Hour	BiNormal	ESN	GEV	GP	Logistic	LogLogistic	LogNormal	TriNormal	Uniform
4	7.94 %	31.76 %	29.25 %	1.21 %	1.11 %	3.52 %	0.00 %	21.11 %	4.12 %
5	7.44 %	36.79 %	28.18 %	1.10 %	0.79 %	1.29 %	0.09 %	19.20 %	5.11 %
6	7.73 %	36.61 %	24.77 %	0.90 %	0.60 %	1.00 %	0.09 %	23.26 %	5.05 %
7	7.65 %	38.08 %	21.90 %	0.78 %	0.54 %	0.70 %	0.07 %	25.37 %	4.91 %
8	7.71 %	40.64 %	21.23 %	0.80 %	0.48 %	0.83 %	0.03 %	23.15 %	5.13 %
9	7.07 %	43.67 %	20.99 %	0.84 %	0.49 %	0.67 %	0.08 %	20.81 %	5.38 %
10	7.54 %	43.35 %	20.99 %	0.84 %	0.48 %	0.69 %	0.07 %	20.38 %	5.67 %
11	7.47 %	42.86 %	20.80 %	0.90 %	0.51 %	0.64 %	0.07 %	20.99 %	5.76 %
12	7.85 %	42.94 %	20.39 %	0.81 %	0.51 %	0.58 %	0.12 %	20.34 %	6.46 %
13	7.70 %	43.27 %	20.00 %	0.73 %	0.52 %	0.70 %	0.09 %	20.31 %	6.69 %
14	7.69 %	41.86 %	20.24 %	0.85 %	0.42 %	0.64 %	0.11 %	21.59 %	6.60 %
15	7.82 %	41.63 %	18.98 %	0.75 %	0.42 %	0.47 %	0.07 %	23.36 %	6.50 %
16	8.28 %	40.54 %	18.78 %	0.82 %	0.44 %	0.58 %	0.07 %	24.39 %	6.11 %
17	8.27 %	41.36 %	19.03 %	0.66 %	0.47 %	0.53 %	0.08 %	23.69 %	5.92 %
18	7.44 %	42.61 %	21.09 %	0.78 %	0.35 %	0.49 %	0.06 %	21.67 %	5.52 %
19	7.28 %	44.38 %	21.76 %	0.81 %	0.48 %	0.82 %	0.06 %	19.08 %	5.33 %
20	7.71 %	43.77 %	23.45 %	0.81 %	0.57 %	0.89 %	0.09 %	17.68 %	5.04 %
21	7.84 %	43.57 %	23.47 %	0.83 %	0.64 %	1.03 %	0.05 %	17.81 %	4.76 %
22	7.34 %	43.93 %	23.27 %	0.90 %	0.62 %	1.11 %	0.06 %	18.49 %	4.26 %
23	8.14 %	40.13 %	23.81 %	1.48 %	0.60 %	1.18 %	0.05 %	19.91 %	4.70 %
24	7.83 %	42.14 %	26.06 %	1.29 %	0.65 %	1.79 %	0.04 %	16.07 %	4.14 %

There are significant amount of tri-modal cases, which account for 15 to 25 percent of the best fit distributions. There are two noticeable increase in tri-modal percentage during morning and evening peaks. These increases coincide with the two decreases for the single modal distributions. To reiterate, mixture distributions are commonly used to model a random variable, selected by chance from a collection of distributions. The result show that during peak hours, travel time becomes more complex with different sub-category distributions that a travel time might fall into. This suggests that these travel times are affected by different factors on the road, such as traffic light patterns or different passenger flow. However, this does pose a theoretical challenge. What is a representative distribution that can be used in a public schedule, since transit schedule can only list one travel time. This question is left for future research.

The above results highlights the need to evaluate stop to stop travel times using skewed distributions and mixture distributions. The AICs also show that current distribution fittings used by current literature, such as Log-Normal distribution, do not provide better fitting for the data, and additional distributions such as Epsilon Skewed Normal distribution should be considered when modeling stop to stop travel times. The results also calls for further research on the potential causes for these multi-modality distributions, as these research might requires additional datasets from transit agency and other government agencies. More importantly, since the mixture distribution is a collection of distributions weighed by different probabilities, more research is needed to assist agencies in determining one single number that is representative for passenger reference.

4.3 Stop to Stop Travel Time Correlations

If agencies wish to schedule for stop level arrival times instead of timepoint level, they might want to know whether analysts can make simple assumptions regarding travel times during the scheduling process to start with. To reiterate, it is reasonable to speculate the stop level schedules are generated by software assuming a constant travel speed between two timepoints. However, based on previous results regarding inconsistencies between mean travel times and scheduled travel times, a relatively constant travel time assumption may not hold. Regardless, analysts might still want to examine whether there is any correlation between consecutive stop to stop travel times, so that these travel speed assumptions can be applied to some segments to reduce their work load. In addition, correlations may also be helpful for dispatchers to make decisions if increased travel time on one segment would lead to increased travel time on other segments. Dispatchers could then potentially instruct buses to run limited stop service to catch up the schedule.

Theoretically, the sum of stop to stop travel times and stop dwelling times should add up to the terminal to terminal travel time. Again, the mean travel times used in this thesis can be added directly regardless of distribution. Agencies also use travel time variances during their scheduling process to determine layover or recovery times. They may also use the travel time variance to determine whether to include slack times in schedules to help accommodate unexpected delays. Therefore, it is important to also know whether the travel time variances can also be added directly to be timepoint to timepoint level or terminal to terminal level variance, so that they can potentially put slack times on certain consecutive segments with larger travel time variation. If these segments are independent, variances can be summed up. Agencies could prioritize reducing variation on specific segments with larger variances to reduce their overall terminal to terminal variance, and thus improving service and operation reliability. Examining the correlation of between travel time segments can help address the above questions.

Correlations can also show whether vehicles travel times are affected consistently by traffic. For example, if a bus gets slowed down by traffic on one segment due to higher than normal congestion, it could potentially get slowed down in the upcoming segments. Knowing where these correlated segments are can potentially help the agency to gather several stops together under a certain traffic assumption when designing the schedule, which would help them alleviate some workload given their limited resources.

The data processing and analyses is a fairly simple one. I can simply query for stop to stop travel times on two consecutive segments recorded on the same trip on the same day. However, if there are outliers that were removed during the outlier removal process mentioned in Chapter 3, then travel times on those two segments for that trip on that date were removed from the correlation analysis. In other words, only two consecutive stop to stop travel time without any outliers were used in the analysis. This demonstrates once again that, if there are outliers in one segment, analysing travel times at stop to stop level can help preserve some stop to stop level data compared to the timepoint to timepoint level analyses. After obtaining the travel time pairs, correlations can be calculated.

These correlations were then put into a map format. Once again, using trip 697167 as an example, the correlation map is shown in Figure 4.11. The visualisation shows the correlation between the coloured segment and the upcoming segment. The map is colour coded in the following way: <-0.5 Dark Green, (-0.5, -0.2) light green, (-0.2, 0.2) yellow, (0.2, 0.5) orange, and >0.5 is red.

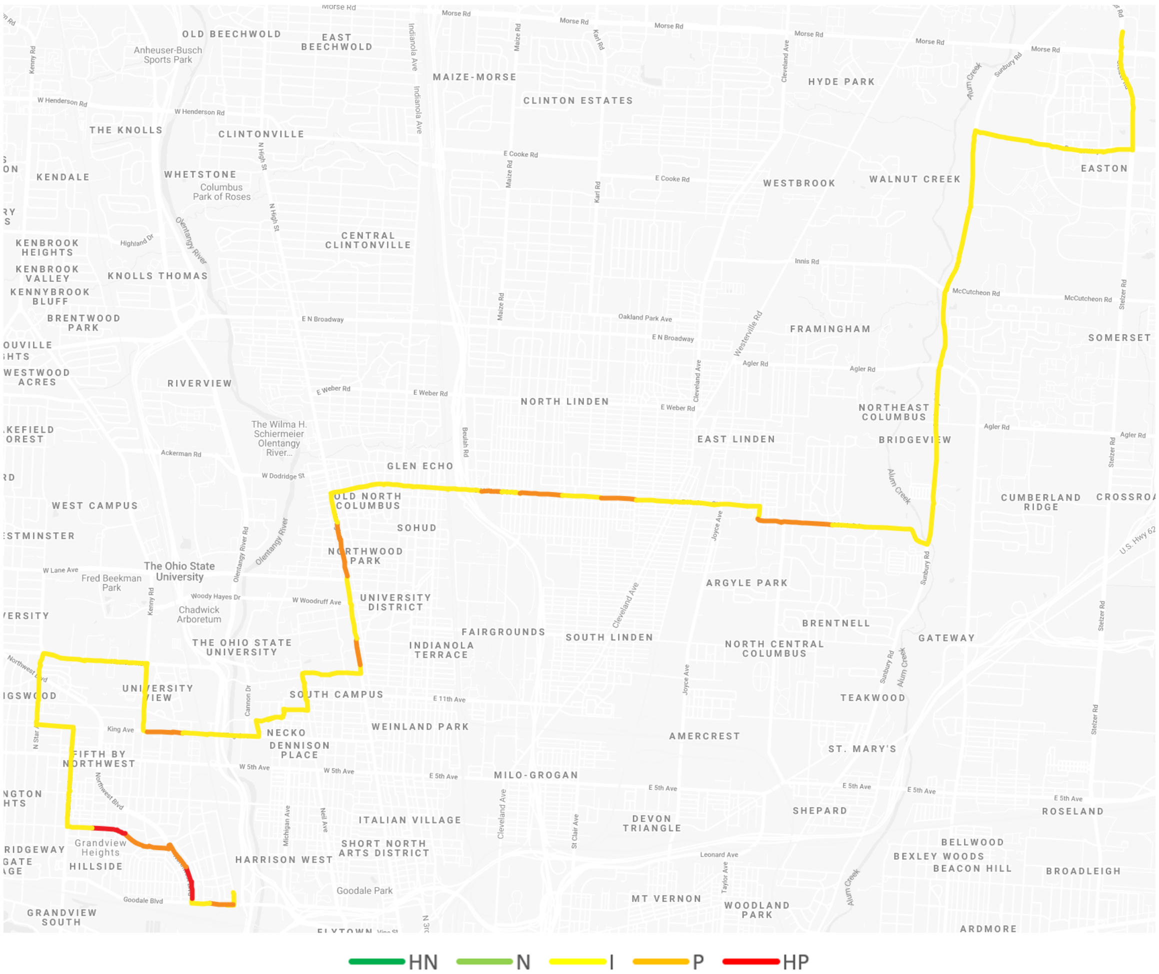

Figure 4.11: Stop to Stop Travel Time Correlation Map for Trip 697167. (HN stands for highly negatively correlated, N for negatively correlated, I for independent, P for positively correlated, and HP for highly positively correlated).

From the map, we can observe that most of the stop pairs have a relatively low correlation, (-0.2, 0.2). However, there are consecutive segments where the travel times are positively correlated, especially in the Grandview area. This makes sense, because those segments don't have many passenger pick ups and drop offs, and there are very limited traffic light interventions (mostly controlled by sensors on quieter streets). This allows buses to travel at a consistent speed in that area, thus resulting in a higher correlation.

Another observation is that, there isn't any highly negatively correlated segments. This also makes sense, since most of the times travel time increase on one segment is unlikely to be followed by travel time decrease on the next segment. However, these cases do exist. One example is the first weekday trip on Line 1 heading to Reynoldsburg Park and Ride. Travel times from High and State and from High and Town are negatively correlated, and the correlation is -0.614. It will be interesting to further examine these negative correlations from far apart segments in future research.

For overall correlation of each stop to stop segments for each trip in the system using the metrics mentioned above, there are 0.1% consecutive segments classified as highly negative, 2.6% as negative, 60.8% as independent, 31.4% as positive, and 5.1% as highly positive. These results suggest that most stop to stop travel times are independent from each other, even with a relatively low threshold for classification. When these percent in each correlation categories were aggregated by hour, shown in Table 4.7 and plotted in Figure 4.12, the independent segments are still the majority. For example, there are 43.77% independent segments at 4 AM shown in the table, which corresponds to the left end of the yellow line in the figure. In other words, knowing one stop to stop travel time, will not likely yield any useful information on the travel time for next segment. Thus, agencies might still need to analyse for each individual stop to stop segments during their travel time analysis.

Table 4.7: System-wide Correlation Classification by Hour. (HN stands for highly negatively correlated, N for negatively correlated, I for independent, P for positively correlated, and HP for highly positively correlated).

Hour	HN	N	I	P	HP
4	0.0029	0.0177	0.4377	0.3690	0.1727
5	0.0020	0.0264	0.4536	0.3770	0.1409
6	0.0012	0.0268	0.5181	0.3532	0.1007
7	0.0015	0.0238	0.5800	0.3275	0.0673
8	0.0016	0.0230	0.5928	0.3293	0.0533
9	0.0015	0.0231	0.6023	0.3219	0.0512
10	0.0014	0.0242	0.6181	0.3129	0.0434
11	0.0017	0.0265	0.6270	0.3055	0.0394
12	0.0019	0.0269	0.6346	0.3001	0.0366
13	0.0014	0.0275	0.6351	0.2995	0.0365
14	0.0013	0.0277	0.6469	0.2912	0.0330
15	0.0014	0.0274	0.6569	0.2838	0.0305
16	0.0012	0.0276	0.6608	0.2797	0.0306
17	0.0019	0.0286	0.6397	0.2964	0.0334
18	0.0016	0.0246	0.6100	0.3214	0.0425
19	0.0011	0.0238	0.5831	0.3419	0.0500
20	0.0010	0.0230	0.5585	0.3583	0.0592
21	0.0030	0.0250	0.5392	0.3663	0.0664
22	0.0034	0.0177	0.5089	0.3950	0.0750
23	0.0035	0.0231	0.4870	0.3956	0.0908
24	0.0023	0.0125	0.3823	0.4654	0.1374

The terminal to terminal travel time equals to the sum of stop to stop travel times and dwell times. Setting the terminal to terminal travel time is important to determine layover times and is important in labour contracts. Since we are discussing stop to stop travel times in this thesis, summing these individual stop to stop segments as random variables should also give us a random variable about terminal to terminal travel time. The expected values, the mean, can be added regardless of distribution, thus giving a clear picture of mean terminal to terminal travel time. The variance or standard

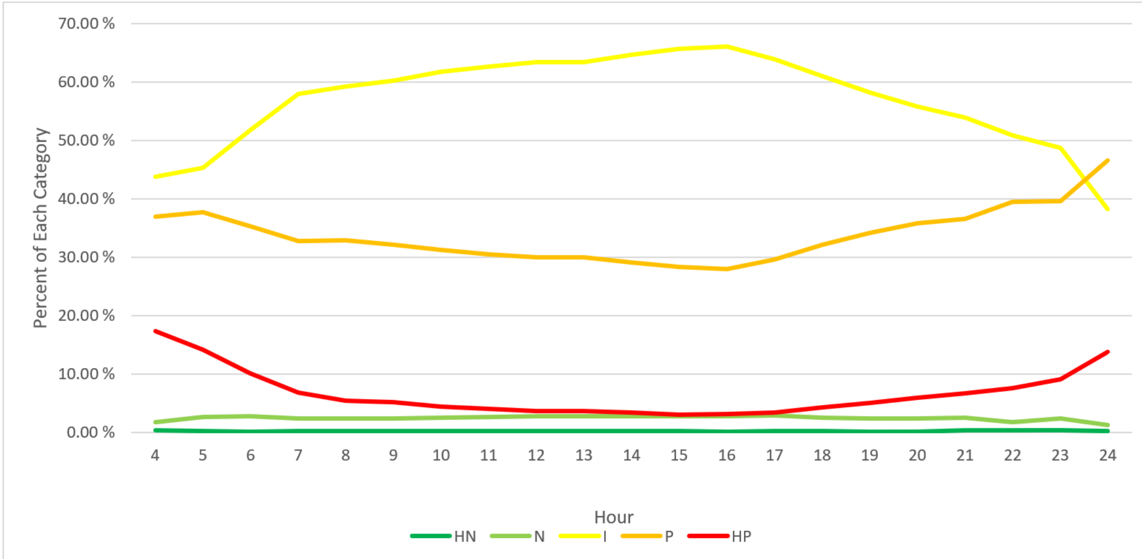

Figure 4.12: System-wide Correlation Classification Plot by Hour. (HN stands for highly negatively correlated, N for negatively correlated, I for independent, P for positively correlated, and HP for highly positively correlated).

deviation of terminal to terminal travel time, however, cannot be summed up directly. The variance equals to the sum of individual variances and their co-variances. The above result shows that most stop to stop travel times can be considered uncorrelated due to their low correlation coefficient, and most of the correlation coefficients are above 0. Thus, most of the variances can be summed up directly. In other words, summing up the variances will slightly underestimate the overall variance. This result suggests that if agencies want to reduce their travel time variability, they might have to look at reducing travel time variability for each stop to stop segment. In addition, since the stop to stop pairs are mostly independent and variances for independent variables can be summed up directly, reducing variance for one segment can reduce overall trip variance. Transit agencies might want to focus their resources on the segments with highest variances.

Overall, the result show most of the stop to stop travel times are not positively nor negatively correlated from the upcoming stop to stop travel time. Travel time in areas with little traffic and passenger activities are often positively correlated. This is also consistent with the systemwide trend shown in the Figure 4.12. When traffic and passenger activities pick up from the beginning of the operation, more segments become independent, and fewer segments are positively correlated. To the contrast, when passenger and traffic activities dies down after PM peak, there are more positively correlated segments and fewer independent segments. This means that analysts can only make assumptions on consecutive stop to stop travel times when there are knowledge regarding how traffic and passengers behave in that area. Otherwise, analysts might have to examine each stop to stop travel time independently.

4.4 Conditional Stop to Stop Travel Time Given Delay

Another concern is regarding the use of all travel time samples on a given trip when analysing its schedule. Travel times are affected by multiple reasons, such as the driver's driving style, and their relationship to the schedule. COTA does not allow drivers to depart a timepoint earlier than the scheduled arrival time, and drivers' bonuses depend on not crossing the timepoint early. If a schedule is sufficient and the bus is running close to being exactly on time, some drivers might opt to slow down to avoid being held at timepoints, while some other drivers would prefer to be held at timepoints so that they can stretch their legs. Since drivers sign up for their runs based on seniority, some senior drivers would consistently sign up for one run, while some other runs will see different drivers from time to time. This would result in inconsistencies when analysing the travel time, and might bias the travel times to suit one specific driver. For example, if a driver that prefers to run fast and stop consistently signed up for one trip is not available on a given day, the schedule might not be consistent for a substitute driver who might be accustomed to run slower. However, due to the lack of operator number in AVL data, I will only look into conditional travel times given a certain delay. This can also be left for future research once the operator data is available from the agency.

Trip 697167 used throughout this thesis is one example of one senior driver operating consistently on this trip. Based on additional interview with the operator, the operator will slow down when running early, otherwise the operator would follow the posted speed limit. However, this might not be the case for all operators. Again, more research on driver behaviours at stop to stop level is needed to make this result more generalisable. The scheduled terminal to terminal travel time is 3900 seconds, and based on the actual AVL data, the average travel time is 3820 seconds without any explicit holding times, while 73 percent of the recorded trips will complete faster than the scheduled 3900 second travel time, and 100 percent of the recorded trips will complete within 4200 seconds (5 minutes longer than the schedule, which corresponds to being on time at COTA). In addition, the trip's on time performance at stop level recorded in January 2020 is presented in Table 4.8 (some days are not recorded due to COTA system upgrading and potential internet connection issues). From the table, it is obvious that the trip consistently operate either on-time or early without being late during January 2020. It is very reasonable to raise the hypothesis from above results that this trip can operate at a higher speed to reduce holding time and the slow operating speed when running early. Travel time is generally considered as a disutility for passengers, decreasing travel time would increase the utility of passengers, thus resulting in higher transit users.

Table 4.8: On Time Performance at Stop Level for Trip 697167.

Date	Early	On Time	Late
20200102	0.0597	0.9403	0
20200103	0.0441	0.9559	0
20200108	0.2059	0.7941	0
20200109	0.3382	0.6618	0
20200113	0.2206	0.7794	0
20200114	0.2687	0.7313	0
20200115	0.1940	0.8060	0
20200116	0.2206	0.7794	0
20200117	0.0735	0.9265	0
20200122	0.2206	0.7794	0
20200123	0.1029	0.8971	0
20200124	0.2206	0.7794	0
20200127	0.1324	0.8676	0
20200129	0.1667	0.8333	0

Using an example on campus, from stop High and 13th to High and 15th, where High and 15th is a timepoint which drivers are not allowed to depart earlier than the scheduled arrival time. A split point of 60 seconds is chosen as an example, which corresponds to the agency's AVL data rounds down delay to the nearest minutes in the published real time feed. Anything delay between 0 and 60 seconds would be considered as 0 in the real time feed. If the bus is running less than 60 seconds behind schedule, the average travel time is 66 seconds. On the other hand, if the bus is running more than 60 seconds behind schedule, the average travel time is 56 seconds.

Two sample T-Test is used to determine whether these two sub-samples have the same mean. The T-Statistics is -4, with p-value less than 0.0001. Therefore, it is reasonable to reject the hypothesis that the two samples have equal mean, and that the operator do slow down significantly when they are closer to running early. This is especially the case when the drivers are biased towards running behind schedule, since they are not allowed to run early. The only option left for the drivers is to slow down and wait for the schedule to catch up while adding excessive travel time to the passengers. This result shows once again that using on-time performance alone does not always show segments where travel time can be improved.

However, if we look at another example on the trip, from Kenny and Steelwood to King and Kenny, where King and Kenny is a timepoint, the conditional travel time shows the opposite story. Once again, if the bus is running less than 60 seconds behind schedule, the average travel time is 94 seconds. On the other hand, if the bus is running more than 60 seconds behind schedule, the average travel time is 101 seconds. Two sample T-Test gives a T-Statistics of 2.85 and a p-value of 0.0046. Based on this result, it is reasonable to conclude that the two sub-samples do not have equal mean, and travel time increases if the bus is running more than 60 seconds behind.

Most of the segments on this trip is consistent with the first example, where if the bus is running more than 60 seconds behind schedule the expected travel time would be lower. Whereas the second case is less common. The explanation for this phenomenon should be explored in future research. However, one correlation for this phenomenon can be tied back to the previous discussion on travel time classifications. Referring back to Figure 4.1, the High St segment is classified as extreme sufficient, while the Kenny Rd segment is classified as extreme insufficient. For now, this analysis confirms the personal observation that buses is slowed down if the trip has a sufficient schedule, and that agency should consider reallocating travel times to allow more consistent travel speeds.

Chapter 5: Transfer Time Analysis

Like I mentioned in the last chapter, there are other aspects in the schedule a passenger expect agencies to meet. In a transit system, one line can only take a passenger to a limited amount of destinations. However, with intersecting lines, a passenger will have access to more destinations served by other lines. Therefore, transfer activity is also an important aspect when passengers plan or experience their trips.

Transfer activities are not coordinated nor guaranteed by COTA most of the times during the day, except there will be line ups at night to provide transfer opportunities when services are running not as frequent as during the day, as mentioned in Chapter 1. Passengers normally infer their transfers, or a trip planning software will, by looking at the published schedule. Either way, the information inferred from the schedule would include the arrival time of the first bus and the arrival time for the transfer bus at the transfer stop.

The importance of transfers and the disutilities caused by the transfers raises questions regarding whether a passenger will make their expected transfer, as well as the observed transfer times between the two trips in this system. To reiterate, the first chapter, COTA does not coordinate nor guarantee passenger transfers except during the late night line ups. However, transfer stops, which are not necessarily timepoints, are clearly marked in their official maps and the trip planning functions from the agency's website. Similar to previous chapter, passengers can easily treat these times as schedules the agency tries to adhere to. Thus, this thesis will refer to them as scheduled transfers. Missing a transfer becomes more important if the scheduled transfer time is short, where passengers risk missing their transfer bus. On the other hand, if the transfer time is too long, passengers might avoid travel on transit for that specific trip, or using other transit lines to get to their destination.

This chapter describes the transfer activities in addition to COTA's existing metric mentioned in Chapter 1, on time performance, that does not describe transfer activities and the relationship between two trips. More specifically, this chapter first explorers the probability of missing a transfer given a scheduled transfer time, which could potentially aid passengers in planning their transfers. Next, the probability of missing a transfer given first bus's arrival at transfer point, which might help passengers identify potential risks associated with their transfers while on board the first bus. Third, identifying transfer time inconsistencies between the observed mean transfer times and the scheduled transfer times, which could help the agency identify and adjust their services to improve transfer reliability. Finally, a potential holding strategy to ensure transfers and its potential consequences relating to holding, which could help agencies in making day to day operational standards.

The last two analyses in this chapter will focus on transfer opportunities provided to passengers by the schedule. If the first bus ran late and missed the planned transfer bus, a later transfer bus could potentially provide shorter transfer time than scheduled to the passenger. However, in this case, despite the potentially shorter transfer time, passengers might still arrive at their destination late due to the delay on the first bus. To better illustrate the effect of missing a transfer in the third analysis, the time a passenger miss their planned transfer is set to negative and used in the analysis. Otherwise, all transfer times would be positive, thus not showing passengers missed their planned transfer clearly. Determining whether a later transfer bus has a shorter transfer time to the passenger can easily be done by comparing the headway between the planned transfer bus and a later transfer bus to the time a passenger missed the transfer. As for holding analysis, holdings are demonstrated up to 15 minutes, which excludes the extreme delays cases on COTA's lower frequency lines. However, one limitation is that determining which transfer bus would provides less delay for both agency and passengers is not explored by this thesis when frequent lines were included in the analysis due to the lack of ridership data. In addition, although some trips require more than one transfer to complete, due to the lack of ridership data, this chapter examines one transfer point at a time and the results are not weighted by ridership. This can be left for future research once the ridership data from APC or AFC become available from the agency.

5.1 Missed Transfer Probabilities Given Scheduled Transfer Time

One simple way to look at these transfer activities is treat the transfers as Bernoulli trials, where we measure whether the transfer is successful or not. When passengers make their travel plans and looking at their transfer times, one question that often comes to mind is that, can passengers make to their transfer bus given a scheduled transfer time?

From the recorded arrival time data for the past two years, transfer point arrival times can be extracted and examined. Due to AVL or internet availability issues, actual arrival times might not be available for the both trips. In this case, only transfers with arrival times for both trips are considered. Then, for walking times needed from one stop to another is measured in Google Maps. If the from stop and to stop is the same one, i.e. no walking is needed, walking time is set to 0. Finally, transfer success rates were calculated, where if the first bus arrives before the transfer bus then the transfer is considered as successful.

Table 5.1 presents the detailed results for missed transfers given a scheduled transfer time. It shows the scheduled transfer time in seconds rounded down to the nearest minutes to be consistent with COTA's system, total recorded missed transfers, total recorded successful transfers, total recorded transfers, and the percentage of missed transfers. To illustrate the trend, the percentage of missed transfers are plotted on a chart, shown in Figure 5.1, where x-axis shows the scheduled transfer time in seconds rounded down to the nearest minutes, and the y-axis shows the percentage of missed transfers. For example, given 3 minutes scheduled transfer time, 15.35% transfers will be missed, meaning 84.65% transfers are successful, as show in the figure and the table.

From the table and the figure, we can observe that the changes of missing a transfer decreases quickly, then slows down as the scheduled transfer time increase. This result is expected. If the scheduled transfer time increase, the first bus will have a lesser impact on the transfer times, and thus passengers are more likely to make their transfers.

On the other hand, if the scheduled transfer time is between four and five minutes, there is still a 10% chance a passenger will miss the transfer bus. This again highlights the insufficiency of only using on time performance to monitor their service. The first

Table 5.1: Observed Missed Transfer Probability Given a Scheduled Transfer Time in Minutes.

Scheduled Xfer Min	Missed	Success	Total	% Missed
0	497999	735152	1233151	40.38 %
1	949555	2029368	2978923	31.88 %
2	686885	2425452	3112337	22.07 %
3	470439	2594364	3064803	15.35 %
4	317360	2626204	2943564	10.78 %
5	237379	2863409	3100788	7.66 %
6	157741	2788458	2946199	5.35 %
7	114735	2968499	3083234	3.72 %
8	78855	3011559	3090414	2.55 %
9	55598	2931613	2987211	1.86 %
10	68392	2929819	2998211	2.28 %
11	19636	2278691	2298327	0.85 %
12	12630	2190946	2203576	0.57 %
13	7732	2151234	2158966	0.36 %
14	4786	2006774	2011560	0.24 %
15	1393	953437	954830	0.15 %
16	676	650664	651340	0.10 %
17	156	577250	577406	0.03 %
18	72	582990	583062	0.01 %
19	23	572516	572539	0.00 %
20	20	539414	539434	0.00 %
21	36	491775	491811	0.01 %
22	3	551875	551878	0.00 %
23	8	497318	497326	0.00 %
24	3	486905	486908	0.00 %
25	4	472636	472640	0.00 %
26	2	453305	453307	0.00 %
27	1	449431	449432	0.00 %
28	1	442201	442202	0.00 %
29	1	396951	396952	0.00 %
30	0	103059	103059	0.00 %

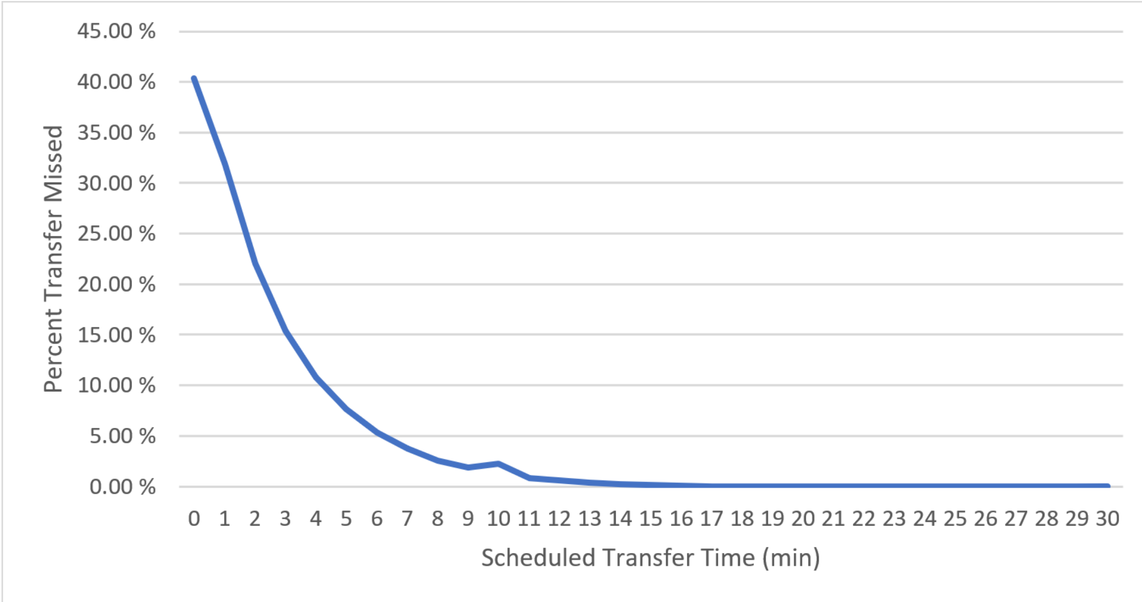

Figure 5.1: Missed Transfer Given Schedule Time.

bus is still considered to be on time by COTA when they run five minutes behind schedule, while the transfer bus is also considered on time when they run exactly on time. This demonstrates once more that on time performance does not describe the transfer aspects passengers expect transit system to behave. This becomes especially important when the transfer bus is not frequent. Missing the bus will delay the passenger significantly longer. This lead to the question of how many transfers would be successful if the transfer bus were held by a certain amount of time, which would be explored in later sections.

From these observations, agencies should come up with a service standard for their passengers on what to expect about their transfers. Using the available AVL data and potentially AFC data or OD matrix, more detailed missed transfer counts and percentages can be calculated, which can be used in conjunction with on time performance to improve the understanding of the transit system's performance. More detailed analyses can be done at trip level to examine whether a specific transfer is considered risky based on the agency's standard, such as how many percent of transfers missed at a given transfer stop. If the transfer is risky, then the agency might want to consider changing the service schedule to allow more time for transfers, or to consider some sort of holding strategy to improve the transfer.

Another interesting pattern is that this graph almost looks like an exponential distribution. More research is needed to closer examine the distribution that generated these data.

5.2 Missed Transfer Probabilities Given Arrival Time of First Bus

While a passenger is on a bus heading to the transfer point, a common question would be, given delay on the first bus, what are the chances to catch the transfer bus? As a passenger, we normally calculate our chances looking at the difference between the expected arrival time on the first bus based on the delay and the scheduled arrival time for the transfer bus since the real time tracker might not work. In other words, passengers want to look at the delay on the first bus and compare it to the scheduled transfer time to determine whether they can make their transfer or not. Again, like the previous section, this section will look at this problem with a simpler view, by treating the transfer points as Bernoulli trials, whether a passenger will successfully make their transfer or not.

From the recorded arrival time data for the past two years, transfer point arrival times and their delays can be extracted and examined. Since there are two trips involved at one transfer point at one time, actual arrival times might not be available for both trips. In this case, only transfers with arrival times for both trips are considered. Then, for simplicity and to be consistent with agency's published AVL data, transfer success rates were aggregated by minutes.

Table 5.2 presents the detailed results for missed transfers given the difference between scheduled transfer time and delay on the first bus. It shows the difference in seconds rounded down to the nearest minutes, to be consistent with COTA's system, total recorded missed transfers, total recorded successful transfers, total recorded transfers, and the percentage of missed transfers. To illustrate the trend, the percentage of missed transfers are plotted on a chart, shown in Figure 5.2, where x-axis shows the difference in seconds rounded down to the nearest minutes, and the y-axis shows the percentage of missed transfers. Here negative numbers show the minutes the first bus arriving before the scheduled transfer bus arrival, whereas positive numbers show the minutes the first bus arriving after the scheduled transfer bus arrival. For example, if the first bus arrives 1 minute before the transfer bus is scheduled to arrive, 4.35% transfers were missed as shown in the table and the figure.

Given more information, we can see the percentage of missed transfers decrease. For example, the percent transfer missed for when the first bus arrives less than one minute before the transfer bus is scheduled to arrive is 14%, less than 40% given only less than one minute scheduled transfer time. This makes sense, since getting to the transfer point one minute before the transfer bus is scheduled to arrive is different from only having one minute of scheduled transfer time. A trip might have a scheduled transfer time of five minutes, but the bus might be running four minutes after the scheduled arrival time.

Table 5.2: Missed Transfer Probabilities Given Scheduled Transfer Bus Arrival and Actual Arrival of First Bus.

ArrivalScheduleDiff	MissedXfer	SuccessXfer	TotalXfer	% Missed
-15	0	773258	773258	0.00 %
-14	39	1154252	1154291	0.00 %
-13	213	1429766	1429979	0.01 %
-12	570	1692230	1692800	0.03 %
-11	791	1918512	1919303	0.04 %
-10	868	2215095	2215963	0.04 %
-9	722	2475183	2475905	0.03 %
-8	548	2637231	2637779	0.02 %
-7	614	2777628	2778242	0.02 %
-6	1735	2794062	2795797	0.06 %
-5	4520	2907526	2912046	0.16 %
-4	11310	2885428	2896738	0.39 %
-3	22426	2916571	2938997	0.76 %
-2	47016	2890755	2937771	1.60 %
-1	123400	2715163	2838563	4.35 %
0	332000	1960298	2292298	14.48 %
1	528938	1140784	1669722	31.68 %
2	571124	613435	1184559	48.21 %
3	508757	314788	823545	61.78 %
4	402490	159051	561541	71.68 %
5	306735	83892	390627	78.52 %
6	227971	45565	273536	83.34 %
7	171880	25553	197433	87.06 %
8	126745	14866	141611	89.50 %
9	93777	8609	102386	91.59 %
10	68642	4836	73478	93.42 %
11	49399	2867	52266	94.51 %
12	34264	1412	35676	96.04 %
13	22951	685	23636	97.10 %
14	13793	301	14094	97.86 %
15	6427	78	6505	98.80 %

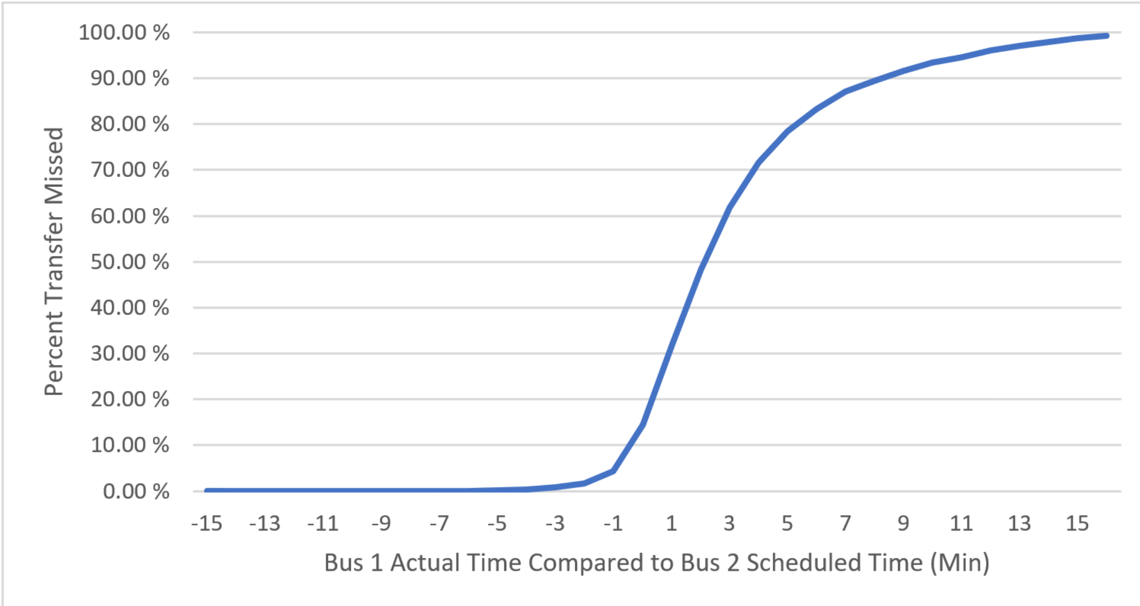

Figure 5.2: Missed Transfer Given Difference Between Schedule Time and Delay.

Notice there are missed transfers when the first bus is arriving before the schedule arrival time of the transfer bus. This is possibly due to buses running early between time points. After all, on time performance is measured at timepoints only and not all transfer points are timepoints, this creates the possibility for operators to leave transfer points early at the transfer stops.

The chances of missing their transfer bus are small but not zero, when arriving a few minutes before the scheduled transfer bus arrival. This is could either potentially related to AVL system issues. There are times when the on board computer does not switch trip directions or freezes and reports wrong information. It could also be contributed to drivers not knowing their schedule very well when driving on an unfamiliar line. However, the general trend still applies.

From the figure, it is easier to observe the rate of change, or slope. Once again, we observe a sharp increase in percent missed transfers centered around 0. However, the sharp increase start to flatten out when the first bus arrives three minutes or more after the scheduled transfer bus arrival. This suggests that if the first bus is significantly delayed, there is a slight chance that the transfer bus is also significantly delayed. Furthermore, for the first two minutes, passengers can still transfer to the next bus with more than 50 percent probability. This also makes sense, since most transfer points, although not all, are timepoints. To reiterate, COTA do not allow operators to cross timepoints earlier than scheduled, and drivers' bonuses depend on this standard. Thus, most operators will avoid crossing the timepoints early, and arriving later than scheduled time, creating more time for transfer passengers even when the first bus run late and arrives a minute or two after the transfer bus is scheduled to arrive.

Since there is still a five percent chance to miss the transfer bus when the first bus arrives one or two minutes before the scheduled arrival time of the transfer bus, due to the transfer bus running early. Agencies might want to hold early buses at transfer points in addition to the timepoints to improve the probability of passengers making to their transfers successfully. Agencies should also keep the transfer points as timepoints and could potentially add more transfer points to timepoints where they evaluate their operators and performances. This would help passengers in two ways. One, timepoints are printed on the schedules, therefore allowing passengers to determine their transfer times easily. Two, timepoints would prevent operators to run early, and could potentially increase passengers' chance to catch their transfer buses.

Again, this analyses can be done in greater detail. However, due to the lack of data, I will leave that to future consideration.

5.3 Identifying Transfer Time Inconsistencies Across Time

After examining the transfer probabilities at a system level, let's turn our attentions to the individual transfer times. Similar to the stop to stop level travel times, this section will use some transfers as examples to highlight the issues.

The following subsections examines the inconsistencies between scheduled transfer times and observed mean transfer times by using t-tests. The null hypothesis used in the test is that observed mean transfer time is equal to scheduled transfer time, whereas the alternative hypothesis is that observed mean transfer time is not equal to the scheduled transfer time.

Since in reality, if a passenger misses the transfer bus, they had to wait for the next one to show up. However, by simply looking at the actual passenger transfer times, it is hard to tell transfers that are more likely to be missed. For the analyses below, missed transfers are represented as a negative time to better illustrate them.

5.3.1 Transfer Level

For each individual transfer, the scheduled arrival time for the first trip and the transfer trip are extracted. Then, the actual arrival times for both trips are matched to these transfer points. From these times, I can calculate the scheduled transfer time and actual transfer time for each day for each scheduled transfer. Finally, using one sample t-test, I can test the hypothesis of whether the scheduled transfer time is the expected value observed from these samples. I will show two examples to illustrate transfer time inconsistencies, one with longer and one with shorter than scheduled transfer time. More specifically, transfer time will be represented in positive number showing the time for passengers to walk and to wait for their transfer bus. On the other hand, negative times show passengers miss their transfer bus by how many seconds.

For longer than scheduled transfer times, using our trip 697167 on Line 31 which tends to run early, the example is the transfer point at High and 13th. This trip is scheduled to arrive at High and 13th at 7:13:41 am, and the southbound Line 2 is scheduled to arrive at 7:22:05 am. By calculation, the scheduled transfer time is 504 seconds. After matching the actual arrival times, the average transfer time is 814 seconds with a 122 second sample standard deviation, resulting in a 12.68 t-statistics. Even without the t-statistics, we can see that the average transfer time significantly more than scheduled transfer time, more than two sample standard deviations away to be more precisely. In fact, if we calculate all other transfers from trip 697167, it has no significant shorter than scheduled transfer time.

On the other hand, transfer times to trip 697167 will be tighter, since it tends to run early. For example, trip 691475 on southbound Line 2 is scheduled to arrive at High and Hudson at 7:16:00 am, and trip 697167 on northeast-bound Line 31 is scheduled to arrive at High and Hudson at 7:20:00 am. The scheduled transfer time is 240 seconds. By calculating the actual transfer times on these two trips, we get an average transfer time of -75 seconds and a standard deviation of 93 seconds, resulting a -16 t-statistics. Again, negative transfer times are used here to better illustrate missed transfers. Where in this case, the passenger miss the Line 31 trip on average by 75 seconds. It is also easy to observe this significantly shorter than scheduled transfer time without the t-test. The difference between scheduled transfer time and average transfer time is more than 5 minutes, and the average transfer time is more than 3 standard deviations away from the scheduled transfer time.

For the agency, it is important to identify these cases, especially at more important transfer points in the system such as High and Hudson. By examining why these transfer times have significant deviations, whether it is due to travel time insufficiency or the operators leave the terminal behind schedule, the agency can improve the transfer reliability throughout the system, by either holding the bus slightly to allow more reliable transfers or addressing operation issues on specific trips.

5.3.2 System Level

For the system level, all transfer times were compared to their scheduled transfer times, similar to the transfer level analysis. Then, t-statistics were aggregated to represent the system. If the t-statistic for a specific transfer is not able to reject the null hypothesis, it is classified as balanced. If the t-statistic is able to reject the null hypothesis and it is positive, meaning the observed mean transfer time is larger than scheduled, the segment is then classified as longer. Finally, if the t-statistic is able to reject the null hypothesis and it is negative, the segment is classified as shorter.

First of all, the overall classification for the past two years are 20.77% transfers with significantly shorter transfer time than scheduled, 48.9% transfers with balanced transfer time, and 30.33% with significantly longer transfer times than scheduled.

However, if we look at the percentage of transfers in each category by hour. The time pattern is less obvious than the transfer time section. Percentage of transfers in each category by hour is presented in Table 5.3. These results are plotted in Figure 5.3 for better illustration, where x-axis represent the hour, and the y-axis represent the percentage in each class. For example, in the table, there are 44.38% transfers classified as Balanced at 4 AM, which corresponds to the left end on the yellow line in the figure.

Table 5.3: Transfer Point Classifications by Hour.

Hour	Shorter	Balanced	Longer
4	24.62%	44.38%	31.01%
5	22.32%	47.36%	30.33%
6	23.08%	42.45%	34.48%
7	20.90%	42.55%	36.55%
8	19.43%	47.22%	33.34%
9	17.89%	51.20%	30.91%
10	17.46%	52.41%	30.12%
11	18.38%	51.66%	29.95%
12	19.38%	50.72%	29.90%
13	19.10%	51.51%	29.38%
14	19.05%	51.24%	29.71%
15	20.15%	49.45%	30.40%
16	22.65%	47.13%	30.22%
17	21.68%	47.82%	30.50%
18	20.31%	49.48%	30.21%
19	19.05%	52.80%	28.15%
20	19.93%	49.96%	30.10%
21	27.50%	49.04%	23.46%
22	34.76%	43.89%	21.35%
23	25.63%	48.97%	25.40%
24	31.41%	44.28%	24.30%

From the table and the figure, we can observe that the temporal component is not as strong as the travel time section. Here, we see an increase in longer than scheduled transfer time category and a decrease in equal transfer time category during the morning peak. To the contrary, for the PM peak, there is a slight increase in shorter than scheduled transfer time category and decrease in equal transfer time category. However, there is a significant decrease for the longer than scheduled transfer time

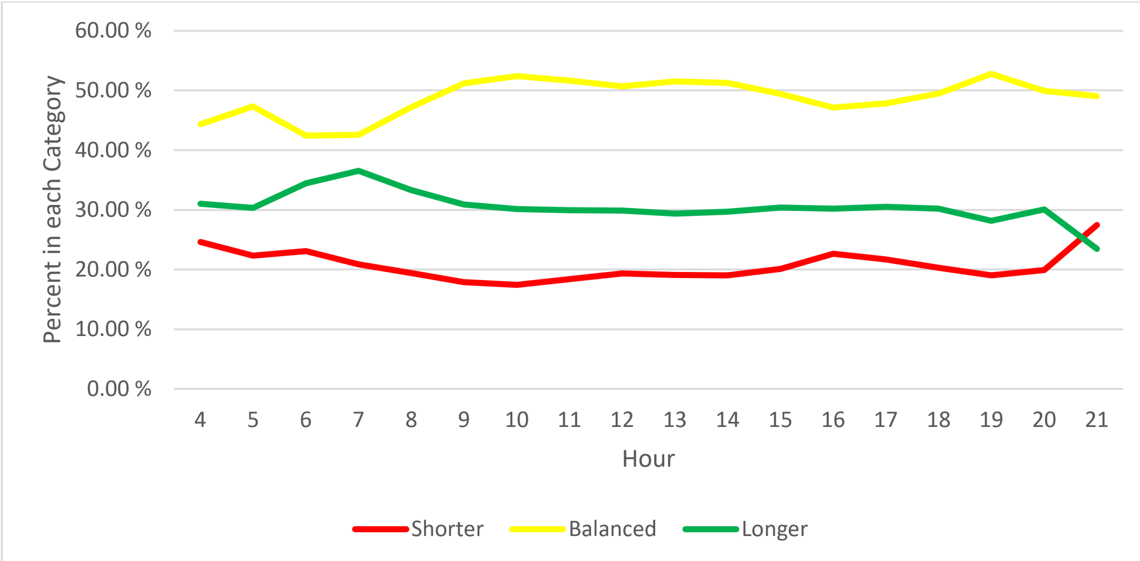

Figure 5.3: Transfer Classification Plot by Hour.

category and a significant increase in shorter than scheduled transfer time category around 9 pm. These temporal results indicate that transfer time classifications are relatively stable, with a longer transfer time during the morning and shorter than schedule transfer times right before the line up starts.

5.4 Holding buses for transfer passengers

For passengers, missing a transfer is not a good service experience, since they had to wait for the next trip. This is especially unpleasant when missing an infrequent service, since the wait time will be very long.

From the results above, it is clear that a passenger is most likely to miss a transfer when the scheduled transfer time is tight, or when the first bus is delayed resulting a lower transfer time. Since the chances of missing a bus decrease very quickly with minute closer to 0, this led to a question, what if the transfer bus is told to wait for first bus? Assuming everything else, such as travel times, stays the same, what will happen to the transfers and subsequent operations?

Since there is no passenger transfer data, I came up with a simplified simulation, where I assume there is transfer activity at every scheduled transfer point. I picked a day as an example in this subsection, more specifically January 23rd 2019, since almost all of the trips were recorded on that day. Then, I came up with some sample holding strategies. The holding categories include holding all buses, holding standard buses, and holding buses with more than 30 minute headway for a given maximum amount of time, ranging from 0 (no holding, day progress as recorded) to 15 minutes. For each holding strategy, number of missed transfers, number of delayed subsequent trips, and number of times buses was held or hold count of the daywere calculated to describe the impacts to passengers and to the agency.

First, all the scheduled transfer points were sorted from earliest to latest as a starting point. This is to simulate the day starting from the beginning of the operation when nothing was held. Next, as the day progress, if there a trip in the corresponding holding catalogue, running with the sum of total holding less than maximum holding time, delays will be added to this trip to simulate buses being held until the second bus arrives. Repeat this step until 10pm when the line up starts. As a reminder for the Introduction chapter, evening line ups mean all local buses meet downtown for 10 minutes to coordinate and guarantee transfers due to buses running at lower frequency. Finally, as the day ends, the previously mentioned descriptive statistics were calculated, and the results are shown in the following paragraphs.

Table 5.4 shows the detailed simulation results for January 23, 2019. It includes the descriptive numbers for each holding category and the maximum holding time in seconds. To better illustrate these results, they are plotted on three charts, Figures 5.4, 5.5, and 5.6. As an example, the table and the green line in the figures show when holding routes with headway more than 30 minutes for a maximum 5 minutes, 7220 transfers would be missed systemwide, 1 additional terminal departure would be missed, and 213 holdings would be performed.

From Table 5.4 and Figure 5.4, we can observe that as the max holding time increase, there is an initial drop in missed transfers number, and, sooner or later, an increase in missed transfer numbers. Based on the data, the increase in missed transfer numbers is most likely a result from upstream holding, which used up the max holding time waiting for buses with larger delay. The holding caused the first vehicle to miss subsequent transfers further down the line, while the transfer buses, also used up its max holding times, do not hold for the first bus. The trend is especially obvious when only holding the standard lines.

Another observation from Table 5.4 and Figure 5.5 is that as the maximum holding time increase, the delayed future trips also increase. This is especially problematic when holding all bus trips, as the number of delayed future trips increase quickly. Although not shown here in the data, the delayed future trips are more likely to happen during the peak hours. This makes sense, since the layover times at the terminals are shortened due to the travel time increase during peak hours.

Table 5.4: Transfer Holding Strategies Compared for 2019.01.23.

	Hold 30+ min Headways			Hold 30, 45, and 60 min Headways			Hold All Routes		
MaxHold	Missed	FutureDelay	Holds	Missed	FutureDelay	Holds	Missed	FutureDelay	Holds
0	7244	0	0	7244	0	0	7244	0	0
1	7208	0	81	6732	0	479	5564	3	1545
2	7198	1	137	6569	5	750	4548	29	2577
3	7197	1	178	6586	8	922	3882	39	3287
4	7195	1	205	6593	10	1058	3401	45	3965
5	7220	1	213	6727	16	1127	2992	62	4634
6	7240	1	221	6946	19	1199	2797	71	5153
7	7286	1	227	7055	24	1263	2605	81	5751
8	7300	1	226	7267	22	1283	2464	89	6271
9	7347	3	229	7456	26	1332	2262	92	6748
10	7351	3	232	7697	26	1334	2250	98	7262
11	7351	3	234	8165	28	1352	2218	117	7761
12	7361	4	238	8354	30	1357	2218	117	8075
13	7361	4	240	8772	34	1382	2065	110	8461
14	7392	4	235	8938	40	1395	2136	98	8801
15	7392	4	235	9143	40	1432	2182	118	8961

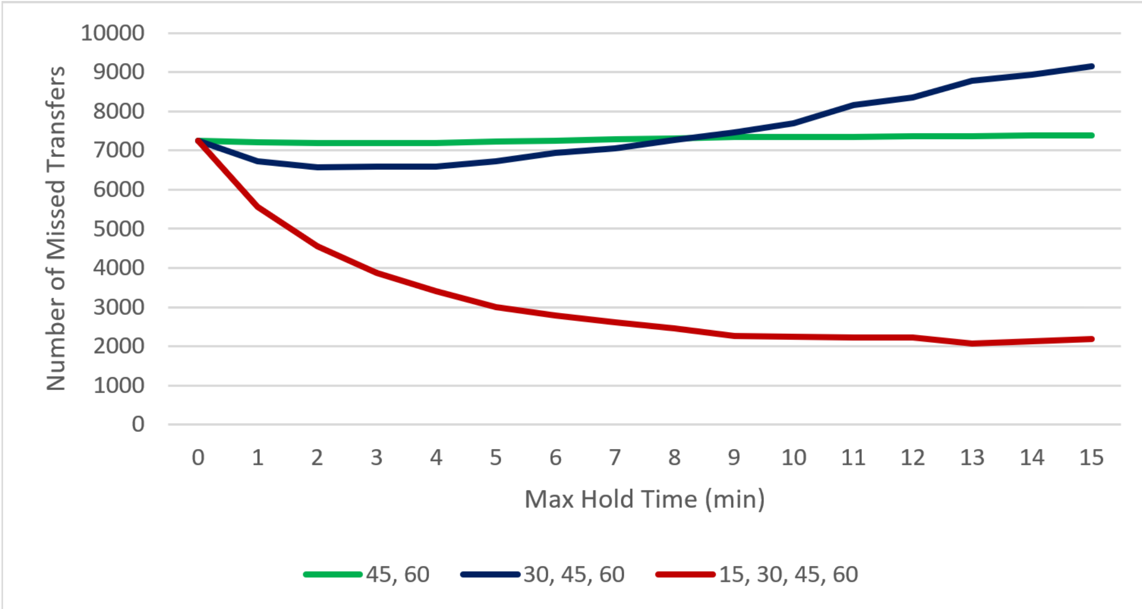

Figure 5.4: Missed Transfer Counts using Different Holding Strategies, where the numbers in the legend corresponds to COTA's headway category. E.g. the blue line represents holding routes that run every 30, 45, and 60 minutes, which corresponds to COTA's standard routes.

From Table 5.4 and Figure 5.6, we can observe that as the max holding time increase, there is an increase in holding counts. This is expected, as the maximum holding time increases, certain trips would be held longer, and might cause extra holdings in their subsequent transfers that would not be held when the trips were not held as long. This again highlights the strong dependence among all bus lines, since holding one bus could potentially cascade delays to its future transfers, as well as cascade through subsequent trips and future transfers. This shows holding buses for transfers can potentially be problematic as the number of holding increases. This is notable when holding all lines. As the maximum holding time increases, the number of holdings grow faster than the other two categories. When maximum holding time

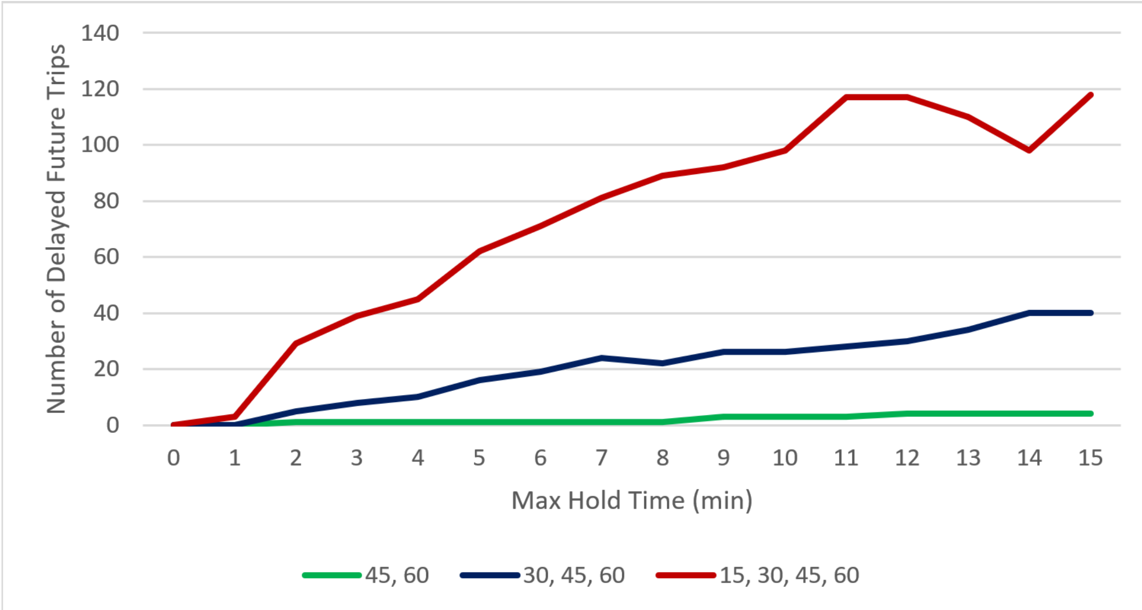

Figure 5.5: Delayed Next Terminal Departure Counts using Different Holding Strategies, where the numbers in the legend corresponds to COTA's headway category. E.g. the blue line represents holding routes that run every 30, 45, and 60 minutes, which corresponds to COTA's standard routes.

is only one minute, the holding counts reach more than 1500, which is more than the maximum for other two categories. Even though the holdings counts are significantly less for the other two holding categories, the counts may still reach several hundreds quickly. Thus, coordinating and tracking these transfers may not be realistic for operators and agencies.

Although not shown in the result section, holding strategies do result in missing other transfers points, but overall transfer risks are lower. The results also highlights the result of TSR. TSR created a network of bus lines, considering only one bus, one line, or one transfer at a time, ignores the strong level of interdependence between

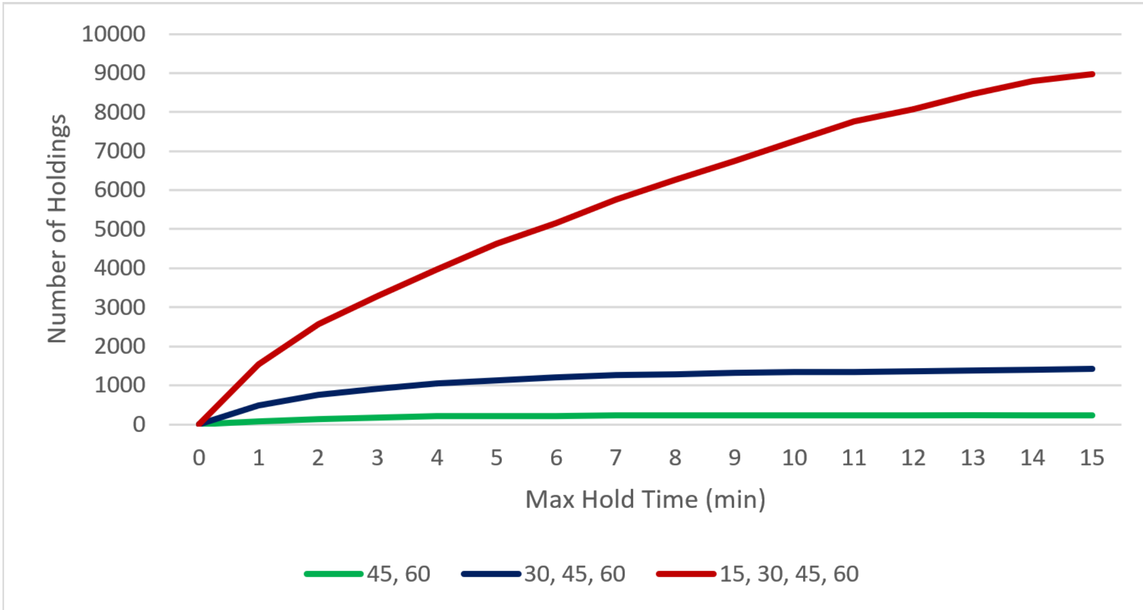

Figure 5.6: Holding Counts using Different Holding Strategies, where the numbers in the legend corresponds to COTA's headway category. E.g. the blue line represents holding routes that run every 30, 45, and 60 minutes, which corresponds to COTA's standard routes.

buses. Holding one bus for a long period of time, will delay and potentially cause passengers to miss future transfers on the bus.

Again, the above simulation is run based on the assumption that there will be some passenger transferring at the given transfer point. The actual numbers will be lower than what is shown in the results. However, based on these observations, it is important to recognize and combine the results from previous sections. Since passengers are most likely to miss tightly scheduled transfers (less than three minutes) and transfers when the first bus is delayed closer to the arrival of transfer bus, along with the initial decrease in missed transfer counts, it is important to recognize the first three or four minutes.

Again based on the observations from Figures 5.4, 5.5, and 5.6, the missed transfer count decrease the fastest when the maximum holding time is less than 2 or 3 minutes while the number of holdings and delayed future trips increase the fastest in the same period. This creates a tension between holding the bus for passengers and ensuring later operation reliability. Since holding more than 2 minutes can cause delays in future trips and 2 minutes is not a long time interval, a holding strategy that provides a good compromise would be that the agency tell the operators to wait for transfer passengers if they see another bus approaching the transfer stop, so that the overall holding time for a specific trip would remain short. The strategy would reduce missed transfer counts, while the holding time would not cause significant delays for later times and holding counts would remain manageable for dispatchers or supervisors to control. Although this process can also be automated by the software vendors, human monitoring is still required to potentially override the system to ensure reliability in later operations. One example of such automated system is from INIT. They developed a transfer support system for planners to pre-plan transfer points and allow passengers to request holding their transfer bus. Most importantly, regardless of whether the process is automated or not, passengers should let their operator know when they will potentially miss their transfer bus, and ask the driver to signal the transfer bus to wait for transfer passengers.

Chapter 6: Discussions and Conclusions

6.1 Research Summary

To summarize, this thesis research was motivated by the importance of transit reliability in general (Carrel et al., 2013), COTA's transit system redesign, and COTA's implementation of automatic vehicle location systems. Once these significant changes were in place, the agency might want to fine tune their system so that it becomes more reliable to passengers. Since previous researchers had issues obtaining and analysing transit data at a more detailed level (Mazloumi et al., 2010), this thesis attempted to set up a data analysis framework for analysing the recorded stop to stop travel time as well as transfer time data.

The main data source used in this thesis is the GTFS and GTFS real time data, published by the agency. The on-going data collection started in January 2018, meaning more than two years of GTFS data were collected, stored, processed, and analysed. The data processing framework was designed to be updated everyday so that the analyses results are up to date and reflects the latest data.

The data processing processes matches the vehicle GPS locations to the route locations and distances. Next, from these route distances stop arrival and departure times can be estimated. These stop time data were used in transfer time section to identify the arrival time of first bus and the departure time of transfer bus. Then, the travel times between stops were calculated. Finally, outlier travel times, such as break downs and reroutes, were removed from the analyses using DBSCAN.

Although this study is based solely on the GTFS data from Columbus, OH, this methodology can easily be applied to other transit systems elsewhere that implemented AVL systems with a few potential modifications based on agency specific requirements. For example, holding strategies might need to be adjusted since other agencies have different holding requirements for drivers. Another potential adjustment would be travel time matching and delay calculation process, since some agency operate based on frequency without a fixed schedule. On the other hand, some agencies use a completely different AVL data format. The basic idea of this study can still be applied for the agencies that uses CEN Network Timetable Exchange (NeTEX) and Service Interface for Real Time Information (SIRI) formats. These modifications will be minimal, since schedule and vehicle position data are provided universally across these data standards.

6.1.1 Travel Time Analysis

For the travel time analyses, they are done at stop to stop level and they have four components, travel time inconsistencies, travel time distributions, travel time correlations, and conditional travel time given delays. First, the travel time inconsistencies were done at stop to stop level, trip level, line level, and system level to reflect temporal and spatial differences.

For stop to stop level, the analysis compares the scheduled travel time to the recorded sample mean, 15th percentile, and 85th percentile, for all weekday trips that serves that segment. The analysis used one segment to illustrate the inconsistencies for this particular stop segment. The result show, for this segment, the schedule does not reflect actual travel time increases around lunch hours, and the schedule does not add enough travel time to accommodate the increased travel time for AM and PM peaks.

Then, for trip level, the analysis aggregates all stop to stop segments for one particular trip. Data visualisation were used to help illustrate the inconsistencies across space for one trip. One particular trip was used as example to demonstrate the analysis. The result for this trip shows that travel times are likely to be insufficient around major intersections, where timepoints are usually located at. The surplus travel times, on the other hand, tend to lie between major intersections.

The line level analysis is the aggregation of trip level analysis, where the amount of sufficiencies and insufficiencies segments are counted and presented for all trips on one line. This step aggregates temporal and spatial variations for one line, which allows agencies to narrow down potential causes for insufficiencies on one line. The line 31 was used to illustrate the process. The result show rapid variations for the line throughout the day.

The system level aggregates all segments in a time period. This allows agencies to control for a certain time and examine the spatial variation at a higher level. The thesis presented two one hour periods, 12 PM to 1 PM and 5 PM to 6 PM. The result is consistent with previous finding, that the peak hours have more insufficient segments despite the increase in scheduled travel times. The result also generalises the finding at trip level, that the insufficient segments are concentrated around major intersections.

Next component is the travel time distribution. This component is mainly to check whether stop to stop travel time distributions are consistent with higher level distributions like terminal to terminal travel time distributions. If the distributions are not consistent with previous used distributions, what are some potential representative distributions that could be used to model stop to stop travel times? The result show that previously used distributions at terminal to terminal level do not fit the stop to stop travel times better. Due to the higher data resolution, more impacts, such as from traffic lights, can be observed from the distributions, making mixture distributions one of the dominant distributions. On single mode distributions, Epsilon Skewed Normal and Generalised Extreme Value distributions tend to perform better. For spatial information, mixture distributions tend to appear near major intersections, which have more complicated traffic light settings. Temporally, mixture distributions tend to appear more often around peak hours, which have more traffic impacts.

For the correlation analysis, sample from one stop to stop segment is correlated with the upcoming segment. This help determine whether stop segments can be grouped under one traffic condition assumption. This also help bridge the gap between the observed data and the theoretical derivation, where variances cannot be added directly without the independent condition. However, the result show most stop to stop travel times are independent, there are more positively correlated segments than negatively correlated, and that more stops segments become independent during peak hours. The spatial finding show that only stop segments that are in the same geographical area without large traffic and passenger activities can be grouped under the same assumption.

6.1.2 Transfer Time Analysis

The transfer analysis tried to describe the overall recorded transfer time data from the passengers' perspective, since there are little literature that systematically analyse the actual transfer times in their analysis. The components for the transfer analysis are the probability of missing a transfer at a given scheduled transfer time, probability of missing a transfer given arrival time of the first bus, deviations between scheduled transfer times and observed transfer times experienced by passengers, and finally a very simple holding strategy that attempts to reduce the probability of missing a transfer.

For the probability of missing a transfer given a scheduled time, the result show an almost exponential decrease as the scheduled transfer time increases, as expected. Even if the scheduled transfer time is around five minutes, the agency's definition of being on time, there is still 8 percent chance that a passenger will miss the transfer. This information could be helpful for passengers to make their transfer decisions and could help agencies in designing a transfer time standard for their systems.

As for the probability of missing a transfer given the arrival time of the first bus, the result show an almost exponential increase starting at the first bus arrives 1 minute before scheduled transfer bus departure. This is could potentially be explained by transfer bus running early.

For the deviations between scheduled and actual transfer times, the analysis is done at two levels, transfer level and system level. The transfer level used 2 transfer points to demonstrate shorter and longer than scheduled transfer times. The system level aggregated the individual transfer within the hour and found most transfer times are as expected during midday. However, AM peak tends to have longer than scheduled transfers times, whereas the PM peak tends to have shorter than scheduled transfer times.

Based on these probabilities, a simple holding strategy was demonstrated by applying it to different service headway categories. The result show the number of transfers decrease the fastest when buses are held by maximum 3 minutes. If buses were to be held longer, the holding could cause missed transfers at future transfer stops further down the line. The result also shows expected increase in missed next terminal departures, meaning the layover times at the terminal are not sufficient for drivers to recover from the holding times and delays. Finally, the result indicates the number of holdings increases as the maximum holding time increases.

6.2 Preliminary Conclusions

This section summarises the preliminary findings from this thesis. Pending further research, these findings may only be applied specifically to COTA. However, similar analyses can be done to other systems without significant changes to the framework.

The results suggest that the difference between the scheduled and actual travel time data varies both temporally and spatially. This study reveals that transit schedules do not accurately reflect the actual variation in travel time at different times of the day for certain sections. There are also a few sections where the actual travel time is either insufficient or sufficient of schedule for all trips during the week. Despite the increase of scheduled travel times, there tend to be more insufficient scheduled stop to stop travel times during AM peak, lunch hours, and PM peak. In addition, regardless of time of the day, travel times near major intersections tend to be insufficient.

Transit agencies will be able to focus their resources on adjusting scheduled travel times based on route segments where actual travel time deviates significantly from the schedule. Analysts at agencies might be able to focus their resources on addressing the increase in insufficiencies during rush hours and lunch hours by reallocating the excessive travel times on certain segments of the trip to the insufficient segments of the trip. On the other hand, agencies could address the insufficiencies around major intersections by advocating for Transit Signal Priorities.

The analysis also provides new information for researchers and analysts on travel time distributions. The result show stop to stop travel times are likely to be skewed, and sometimes tend to be multi-modal. Although more research is needed, there are other distributions, not yet mentioned in the literature, that provides better fit for stop to stop travel times, namely GEV, ESN, and mixed normal distributions. The result show mixture distributions share a similar trend with stop to stop travel time insufficiencies. There tend to be more mixture distributions around traffic lights and during peak hours.

Travel time correlations at stop to stop level also provides some new insights. When summing random variables, stop to stop travel times in our case, variances cannot be added without considering their co-variances unless they are independent. The result show most of the stop to stop travel times are independent with its previous or subsequent segments. This indicate variances can also be added most of the time, which means larger variance on one stop to stop segment would be likely to contribute to the overall terminal to terminal travel time variance. Agencies can focus their resources on addressing these larger variances to improve their overall service consistency. In addition, since most of the travel times are independent, meaning knowing travel time for one segment will not affect the travel time outcomes on other segments, designing stop arrival and departure times might require agencies to look at individual segments separately. However, the result show higher positive correlations in areas or times with little passenger activities or traffic interruptions. This makes assumptions regarding consistent travel speeds more justifiable in these areas.

Although future research is needed, conditional travel time analysis given delays show that operators are more likely to deliberately slow down to avoid running ahead of schedule, since COTA does not allow drivers to cross timepoints earlier than scheduled arrival time. These results can help agencies to design schedules that are more representative for driver behaviours. Agencies might want to use data where vehicles are more than a minute behind schedule to avoid drivers consistently slowing down to avoid running early. Since in vehicle delays are viewed more negatively than waiting time at their origin stop (Carrel et al., 2013), designing schedules that are more representative could help improve passengers' happiness.

It is also important for agencies to design a transfer time standard, so that transfer time performances can be analysed and improved for passengers. This could potentially be done by synchronizing their services schedule better at transfer points, so that transfer times experienced by passengers are not "tight", i.e. with more chances to miss their transfers. Since a common on time performance goal is 85%, agencies might also want to set a success transfer goal of 85%, which corresponds to 3 minutes of scheduled transfer time plus walking time between stops. Agencies might want to use 3 minute transfer time plus walking time when scheduling a transfer point to improve their transfer reliability. Given the arrival time of the first bus, the 85% successful transfers can be achieved when the first bus arrives before the scheduled transfer bus arrival. Dispatchers might want to avoid the first bus arriving later than the scheduled transfer bus arrival. Some form of holding strategies should also be considered to improve the connections between services. The result show greater improve when buses are held less than 3 minutes, suggesting an easier improvement by instructing drivers to wait slightly or signal the other vehicle if they see another bus approaching the transfer point. Agencies might also want to prioritise transfer points that has significantly shorter than scheduled transfer times to improve their transfer reliability, which corresponds to 20% of the total transfer opportunities in COTA system.

6.3 Potential Use of the System

The research findings could also provide new insights and implications for transit planners to improve the reliability and consistency of scheduled travel times.

Stop level information could also help focus transit planners to the sections where the schedule is significantly insufficient, instead of trying to analyse multiple stops between time points. They can also infer the likely causes for why vehicles are running significantly slower than scheduled travel time, based on time of the day, day of the week, and the location of the stops and road layouts.

In addition, the methodology developed in this study may also help traffic engineers improve the effectiveness of traffic light settings in order to facilitate more consistent transit operations. Especially, this allows transit planners to advocate for the effective means of improving transit reliability, such as transit signal priorities. This would also allow the agencies to explore different strategies when installing additional transit priority signals based on their resources and their most problematic segments. Agencies could consider the historical travel time data and travel time variations to determine which intersections are more suitable or should be prioritised for transit priority signals. If historical data of similar implementations or testing projects are available, agencies could also evaluate the effects of different strategies by comparing detailed travel times and travel time variations before implementation to travel times and travel time variations after implementation.

This study can also be applied to improve the transparency of transit agencies. This visualization helps transit riders understand where buses run tend to catch up or get delayed, and this information would help riders plan their trip accordingly. It would also help agencies provide more justification when advocating for transit investments, such as additional transit priority signals or bus lanes.

These travel time analysis could also help agencies determine a better operational schedule for the drivers. There are only three service changes relating to published schedule per year at COTA. The operational schedule change more frequently, sometimes daily operation determined in the morning will be changed in the afternoon due to interruptions such as vehicle outages and driver assignment changes. Agencies might be able to determine a better daily operational schedule based on driver and vehicle availability.

6.4 Limitations and Future Research

There are still several limitations and assumptions that needs to be studied, and several future improvements are planned for the research.

The most obvious limitation for this study is that this does not apply to agencies without AVL information. If agencies do not have AVL implemented, then there is no vehicle position and timestamp available for any computerized or automated travel time analysis.

One thing to note is that although the mathematics can be very precise and the patterns found can help planners focus their resources, they cannot be used to replace specific planning process. This study does not imply that the numbers generated by this algorithm are superior to existing methods when analysing system performance, since there are other operational constraints that this paper did not consider. This study does not suggest that mean travel times should be used in real world schedules. It is up to transit planners to recognize the patterns described in this paper and determine their analysis methodology in order to achieve their specific operational goals.

One limitation of this study is the impact of schedule revisions, which will be further studied. This study merged all past travel time data into one analysis. However, actual travel times might change due to schedule revisions. If the schedule is tightened, drivers might have difficulties keeping up with tight schedules, therefore they might drive faster to recover. On the other hand, if the schedule added too much time, the vehicles might arrive early, and the dispatchers might instruct the drivers to slow down or to stop.

Another limitation is the effect of other scheduled vehicles on the same corridor, since their interactions will affect the travel time as well. If the two vehicles on two different routes are scheduled closely on the same corridor, the first vehicle might take more the passengers, thus increasing dwelling time for the first vehicle while reducing the dwelling time for the second vehicle. This might cause differences when estimating the dwelling times and actual travel times between the two trips. Similar interactions between local and limited service should also be taken into consideration. Local services might delay express services if there is no space for the limited service to overtake, therefore increasing the travel time on the limited service. An area of future research is to study the impact of these interactions on the travel times. This could potentially help agencies determine a better spacing between vehicles based on travel times as well as a better schedule that reflects travel times more precisely.

Various regression analyses can be applied systematically on travel times at stop level to determine the factors that are more likely to cause deviation from both scheduled and actual travel and transfer times. As described in the literature review, current studies do not consider the entire system and they analysed at time point level. Data mining algorithms can be used to group trips with similar travel times together to aide transit planners when developing new schedules.

From above, future research will design an algorithm that uses real time data to produce suitable travel time values for transit schedules. This could potentially be done by minimizing deviations between actual and scheduled travel times while considering the schedule recovery process, driver behaviours, and passenger behaviours.

Appendix A: Appendix

Table A.1: Best Fit Distributions for each Segment in Trip 697167 by AIC.

FromStop	Distribution	NLogL	AIC	Mu	Median
12T153E	GEV	1790.77	3575.55	103.92	102
1STAVOW	ESN	1234.26	2462.51	35.12	34
1STGRAW	GEV	1517.36	3028.72	70.17	69
1STNORW	ESN	1255.66	2505.31	33.54	32
1STOXLW	ESN	1368.67	2731.35	42.09	40
1STVIRW	ESN	649.90	1293.79	10.65	10
1STWILW	ESN	1001.14	1996.28	28.44	28
EASCHAE	BiNormal	2030.40	4070.79	116.48	105
EASEASPE	TriNormal	1390.56	2797.12	27.03	26
EASMORE	GEV	1545.15	3084.29	47.95	48
EASSUNE	GEV	1728.03	3450.06	85.60	84
GOO868W	ESN	357.78	709.56	24.81	24
GOOYARW	ESN	471.28	936.57	19.70	19
GRA3RDN	BiNormal	1576.87	3163.75	76.89	79
GRA5THN	TriNormal	1395.26	2806.53	36.47	35
GRAKINGN	BiNormal	1493.56	2997.12	67.95	67
HIG13TN	GEV	1583.36	3160.73	59.35	57
HIG15TN	TriNormal	1449.65	2915.29	59.00	53
HIGBLAN	GEV	1636.75	3267.50	51.16	44
HIGHUDN	GEV	1678.02	3350.04	72.71	71
HIGLANN	ESN	1460.52	2915.03	29.59	27
HIGNORN	ESN	1506.90	3007.80	34.24	32
HIGPATN	ESN	1334.76	2663.51	26.27	25
HIGWOON	GEV	1742.03	3478.06	73.50	71
HUDADAE	BiNormal	1069.45	2148.90	20.55	19
HUDCLEE	BiNormal	1707.79	3425.59	73.19	69
HUDDAYE	TriNormal	1659.12	3334.24	58.96	59
HUDDREE	ESN	1476.36	2946.72	49.84	49

HUDFIOE	BiNormal	1381.89	2773.77	34.34	32
HUDHIAE	ESN	1535.16	3064.32	50.34	49
HUDJOYE	ESN	1663.43	3320.85	53.17	50
HUDMCGE	BiNormal	1625.62	3261.24	57.65	57
HUDONTE	GEV	1410.40	2814.80	39.59	38
HUDOPPE	BiNormal	1401.27	2812.54	25.03	21
HUDPONE	TriNormal	1390.61	2797.21	26.45	22
HUDSILE	GEV	1680.69	3355.39	61.03	60
HUDSUME	TriNormal	1316.66	2649.33	53.32	52
KENKINS1	TriNormal	1498.56	3013.11	50.40	49
KENSTELS	GEV	1807.87	3609.73	102.42	102
KIN1197E	ESN	1539.11	3072.22	71.34	69
KINDELE	ESN	1413.85	2821.70	31.96	31
KINKENE1	ESN	1401.28	2796.55	43.28	42
KINOLEE	TriNormal	1922.89	3861.78	76.02	72
KINOLEE1	GEV	1134.49	2262.99	88.56	89
KINSTAW	TriNormal	1513.42	3042.84	49.59	48
MEDC9THN	GEV	1848.13	3690.26	111.20	110
MEDCCE	TriNormal	1013.36	2042.71	43.27	42
MOC2315E	ESN	1315.23	2624.46	35.86	35
MOC250E	ESN	1696.33	3386.67	133.61	133
MOCWOOE	GEV	1526.26	3046.52	54.14	53
NEI10TN	TriNormal	1390.34	2796.68	38.13	36
NEI11TN	Uniform	1875.53	3747.07	118.00	118
NORBURNW	GEV	866.21	1726.43	15.12	15
NORCHAN	TriNormal	1548.46	3112.93	66.30	66
NORGOON	GEV	825.93	1645.86	17.03	17
NORWILLN	ESN	1115.90	2225.80	31.06	30
NOSTARN	ESN	1494.16	2982.31	61.08	59
PARMOCS	ESN	961.79	1917.59	27.32	27
RAIBALS	ESN	752.36	1498.72	76.23	69
STE418N	BiNormal	996.05	2002.10	13.96	14
STECOLN	TriNormal	1338.65	2693.29	27.35	23
STEMORN1	TriNormal	1405.86	2827.71	24.38	21
STEWORN	TriNormal	1917.86	3851.71	86.34	87
SUNCASN	TriNormal	1828.76	3673.51	89.82	88
SUNMCUN	LogLogistic	1436.15	2868.31	62.02	61
SUNPATN	BiNormal	1358.95	2727.89	60.42	58
SUNSUNN	GEV	1674.91	3343.83	117.05	116

Table A.2: Complete Trip Travel Time Table for Trip 697167.

FromID	Sch	SSize	SMean	SSD	Diff	TStat	Classification
RAIBALS	37	141	76.53	45.20	-39.53	10.38	Insufficient
GOOYARW	23	122	21.23	5.43	1.77	-3.60	Surplus
GOO868W	32	148	25.94	6.39	6.06	-11.54	Surplus
NORGOON	26	298	18.39	4.46	7.61	-29.50	ExSurplus
NORBURNW	25	206	17.54	3.39	7.46	-31.57	ExSurplus
NORWILLN	47	363	31.04	5.39	15.96	-56.37	ExSurplus
1STNORW	29	365	33.85	9.89	-4.85	9.36	Insufficient
1STOXLW	44	364	41.95	11.37	2.05	-3.45	Surplus
1STVIRW	14	20	18.50	4.64	-4.50	4.34	Insufficient
1STWILW	38	374	28.60	5.58	9.40	-32.55	ExSurplus
1STAVOW	38	374	35.36	8.01	2.64	-6.37	Surplus
1STGRAW	59	358	70.28	17.12	-11.28	12.47	Insufficient
GRA3RDN	68	361	76.89	20.29	-8.89	8.33	Insufficient
GRA5THN	35	374	36.96	9.91	-1.96	3.83	Insufficient
GRAKINGN	57	388	67.95	13.21	-10.95	16.33	ExInsufficient
KINSTAW	41	390	49.59	13.44	-8.59	12.62	Insufficient
NORCHAN	54	390	66.30	13.88	-12.30	17.50	ExInsufficient
NOSTARN	84	403	61.02	10.56	22.98	-43.70	ExSurplus
KIN1197E	99	392	71.84	12.57	27.16	-42.78	ExSurplus
KENKINS1	39	393	50.40	12.27	-11.40	18.42	ExInsufficient
KENSTELS	71	398	102.54	23.02	-31.54	27.33	ExInsufficient
KINKENE1	75	391	43.29	9.31	31.71	-67.34	ExSurplus
KINDELE	52	407	31.94	8.04	20.06	-50.31	ExSurplus
KINOLEE	70	407	76.02	32.45	-6.02	3.74	Insufficient
KINOLEE1	112	271	88.60	16.06	23.40	-23.98	ExSurplus
MEDCCE	31	270	43.27	11.13	-12.27	18.10	ExInsufficient
MEDC9THN	80	391	111.33	27.64	-31.33	22.42	ExInsufficient
NEI10TN	54	388	38.55	13.25	15.45	-22.98	ExSurplus
NEI11TN	108	396	111.60	28.57	-3.60	2.50	Insufficient
12T153E	119	388	104.03	24.84	14.97	-11.87	Surplus
HIG13TN	79	388	60.04	20.55	18.96	-18.17	ExSurplus
HIG15TN	70	388	59.40	16.70	10.60	-12.50	Surplus
HIGWOON	59	389	73.51	21.87	-14.51	13.09	Insufficient
HIGLANN	30	377	30.25	11.56	-0.25	0.41	Balanced
HIGNORN	45	406	34.45	11.10	10.55	-19.16	ExSurplus
HIGPATN	37	390	27.06	10.20	9.94	-19.24	ExSurplus
HIGBLAN	59	388	50.94	23.45	8.06	-6.77	Surplus
HIGHUDN	71	389	72.80	18.51	-1.80	1.92	Balanced

HUDADAE	29	381	21.06	5.26	7.94	-29.47	ExSurplus
HUDDAYE	51	388	58.96	18.32	-7.96	8.56	Insufficient
HUDSUME	63	388	54.48	14.34	8.52	-11.70	Surplus
HUDSILE	60	390	61.12	18.23	-1.12	1.21	Balanced
HUDOPPE	35	355	26.35	10.20	8.65	-15.97	ExSurplus
HUDPONE	32	374	27.10	11.00	4.90	-8.62	Surplus
HUDHIAE	79	402	50.37	11.47	28.63	-50.05	ExSurplus
HUDMCGE	64	393	57.65	15.73	6.35	-8.00	Surplus
HUDONTE	56	390	39.74	10.98	16.26	-29.25	ExSurplus
HUDDREE	60	390	49.92	10.99	10.08	-18.11	ExSurplus
HUDCLEE	48	391	73.19	20.57	-25.19	24.21	ExInsufficient
HUDJOYE	43	402	53.58	16.02	-10.58	13.24	Insufficient
HUDFIOE	24	397	34.34	9.16	-10.34	22.50	ExInsufficient
PARMOCS	28	398	28.03	5.43	-0.03	0.10	Balanced
MOCWOOE	59	407	54.18	10.64	4.82	-9.13	Surplus
MOC2315E	38	397	36.09	9.53	1.91	-3.98	Surplus
MOC250E	122	397	133.57	17.46	-11.57	13.20	Insufficient
SUNSUNN	152	393	117.17	17.43	34.83	-39.62	ExSurplus
SUNCASN	86	400	89.82	25.74	-3.82	2.97	Insufficient
SUNMCUN	86	390	62.42	10.15	23.58	-45.89	ExSurplus
SUNPATN	81	393	60.60	10.00	20.40	-40.44	ExSurplus
EASSUNE	73	390	85.73	20.64	-12.73	12.18	Insufficient
EASMORE	50	390	48.01	12.86	1.99	-3.05	Surplus
EASEASPE	40	394	27.39	8.10	12.61	-30.92	ExSurplus
EASCHAE	127	418	116.48	36.95	10.52	-5.82	Surplus
STECOLN	53	394	27.66	10.66	25.34	-47.19	ExSurplus
STEWORN	59	391	86.34	37.88	-27.34	14.27	Insufficient
STEMORN1	27	303	29.64	16.77	-2.64	2.74	Insufficient
STE418N	64	149	20.28	4.16	43.72	-128.39	ExSurplus

Bibliography

Basak, S., Sun, F., Sengupta, S., and Dubey, A. (2019). Data-driven optimization of public transit schedule. In Madria, S., Fournier-Viger, P., Chaudhary, S., and Reddy, P. K., editors, *Big Data Analytics*, pages 265–284, Cham. Springer International Publishing.

Bates, J., Polak, J., Jones, P., and Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, 37(2):191 – 229.

Bookbinder, J. H. and Désilets, A. (1992). Transfer optimization in a transit network. *Transportation science*, 26(2):106–118.

Carrel, A., Halvorsen, A., and Walker, J. L. (2013). Passengers' perception of and behavioral adaptation to unreliability in public transportation. *Transportation research record*, 2351(1):153–162.

Ceder, A. (1987). Methods for creating bus timetables. *Transportation Research Part A: General*, 21(1):59–83.

Coleman, M., Tarte, L., Chau, S., Levine, B., and Reddy, A. (2018). A data-driven approach to prioritizing bus schedule revisions at new york city transit. *Transportation Research Record*, 2672(8):86–95.

El-Geneidy, A. M., Horning, J., and Krizek, K. J. (2011). Analyzing transit service reliability using detailed data from automatic vehicular locator systems. *Journal of Advanced Transportation*, 45(1):66–79.

Furth, P. (2000). Data analysis for bus planning and monitoring (tcrp synthesis no. 34). *Washington DC: Transportation Research Board*.

Graves, E., Zheng, S., Tarte, L., Levine, B., and Reddy, A. (2019). Customer journey time metrics for new york city bus service using big data. *Transportation Research Record*, 2673(9):1–10.

Halvorsen, A., Wood, D., Stasko, T., Jefferson, D., and Reddy, A. (2019). Passenger-centric performance metrics for the new york city subway. *Transportation Research Record*, 2673(1):417–426.

Kieu, L. M. (2015). *Utilising big transit data for transfer coordination*. PhD thesis, Queensland University of Technology.

Kimpel, T. J., Strathman, J. G., and Callas, S. (2008). Improving scheduling through performance monitoring. In *Computer-aided systems in public transport*, pages 253–280. Springer.

Knoppers, P. and Muller, T. (1995). Optimized transfer opportunities in public transport. *Transportation Science*, 29(1):101–105.

Levinson, H. (1991). Synthesis of transit practice 15: Supervision strategies for improved reliability of bus routes. *TRB, National Research Council, Washington, DC*.

Liao, C.-F. and Liu, H. X. (2010). Development of data-processing framework for transit performance analysis. *Transportation research record*, 2143(1):34–43.

Mazloumi, E., Currie, G., and Rose, G. (2010). Using gps data to gain insight into public transport travel time variability. *Journal of Transportation Engineering*, 136(7):623–631.

McCann, M. (2016). 2016-2040 long-range transit plan. Technical report, COTA.

Muller, T. H. and Furth, P. G. (2001). Trip time analyzers: key to transit service quality. *Transportation Research Record*, 1760(1):10–19.

ODOT (2016). Franklin county functional class routes. Technical report, ODOT.

O'Sullivan, A., Pereira, F. C., Zhao, J., and Koutsopoulos, H. N. (2016). Uncertainty in bus arrival time predictions: Treating heteroscedasticity with a metamodel approach. *IEEE Transactions on Intelligent Transportation Systems*, 17(11):3286–3296.

Parbo, J., Nielsen, O. A., and Prato, C. G. (2014). User perspectives in public transport timetable optimisation. *Transportation Research Part C: Emerging Technologies*, 48:269–284.

Schubert, E., Sander, J., Ester, M., Kriegel, H. P., and Xu, X. (2017). Dbscan revisited, revisited: why and how you should (still) use dbscan. *ACM Transactions on Database Systems (TODS)*, 42(3):1–21.

Shafahi, Y. and Khani, A. (2010). A practical model for transfer optimization in a transit network: Model formulations and solutions. *Transportation Research Part A: Policy and Practice*, 44(6):377–389.

Wessel, N., Allen, J., and Farber, S. (2017). Constructing a routable retrospective transit timetable from a real-time vehicle location feed and gtfs. *Journal of Transport Geography*, 62:92–97.