Understanding Bus Passenger Crowding Through Origin Destination Inference

by

Christopher W. Southwick

B.A., Economics
Haverford College (2011)

Submitted to the Department of Civil and Environmental Engineering in partial fulfillment of the requirements for the degree of

Master of Science in Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author .... Signature redacted ....
Department of Civil and Environmental Engineering
May 18, 2016

Certified by .... Signature redacted
John P. Attanucci
Research Associate, Department of Civil and Environmental Engineering
Thesis Supervisor

Certified by ... Signature redacted ....
Frederick P. Salvucci
Senior Lecturer, Department of Civil and Environmental Engineering
Thesis Supervisor

Accepted by .... Signature redacted ....
Heidi Nepf
Donald and Martha Harleman Professor of Civil and Environmental Engineering
Chair, Graduate Program Committee

Understanding Bus Passenger Crowding Through Origin Destination Inference

by

Christopher W. Southwick

Submitted to the Department of Civil and Environmental Engineering
on May 18, 2016, in partial fulfillment of the
requirements for the degree of
Master of Science in Transportation

Abstract

Comfort is an important aspect of the transit passenger experience. Crowding can significantly decrease passenger comfort and disrupt service delivery, causing passenger travel times to increase and even resulting in passengers being unable to board an arriving vehicle. Reducing crowding is especially important to encourage ridership growth in the part of the system most attractive to customers. However, due to high marginal costs of manual data collection, crowding has not been extensively analyzed. With the advent of automatically collected data systems, it is now possible to gain a more nuanced understanding on how passengers experience crowding as well as monitor conditions as ridership increases. This thesis explores the use of passenger origin-destination inference to measure passenger crowding on buses using the Massachusetts Bay Transportation Authority (MBTA) bus network as a case study. There are three primary components of this research: vehicle trip level origin destination interchange (ODX) scaling; development of passenger centric crowding metrics, and crowding source contribution estimation.

The trip level scaling process enables the reliable estimation of passenger loads (accounting for those passengers not using smart fare media) for approximately 90% of MBTA bus trips. Comparisons of ODX and Automatic Passenger Counter (APC) load estimates show that while there is some inherent variability in the ODX derived estimates, many vehicle trips have similar estimates. These ODX derived load and passenger flow estimates were used to create passenger centric crowding metrics that consider many aspects of the passenger experience. Results showed that the majority of crowding occurs on high frequency routes during the peak periods as a result of building schedules around average peak loads and slow travel speeds due to traffic congestion.

Next, using a classification tree methodology, the relative contribution that different potential crowding sources have on creating crowding situations was estimated for each route/direction/30-minute-time-period combination. While there were variations between routes and time periods, most of the crowding observed appears to be derived from fixed schedules not able to account for day-to-day fluctuations in demand or service reliability problems that result in uneven headways causing loads on successive trips to vary widely.

The research concludes with a review of crowding intervention/mitigation strategies including which strategies are more effective for each crowding source.

Thesis Supervisor: John P. Attanucci
Title: Research Associate, Department of Civil and Environmental Engineering

Thesis Supervisor: Frederick P. Salvucci
Title: Senior Lecturer, Department of Civil and Environmental Engineering

Acknowledgments

To my advisors, John Attanucci and Fred Salvucci, thank you for giving me a chance to do transit research and sharing your wealth of knowledge . I've learned so much from our weekly research meetings. This thesis would not have been possible without your guidance and helpful feedback.

To Laurel Paget-Seekins, thank you for suggesting crowding analysis as a research topic and for your support both over the summer at Park Plaza and during the school year. Thank you to Greg Strangeways, Phil Groth and the many other MBTA staff who have answered my questions and gave me insights into real world transit operations and planning.

To Gabriel Sanchez-Martinez, thank you for cleaning up my SQL code and your helpful suggestions whenever I became stuck on a problem. To Jay Gordon, thank you for your numerous answers to my data questions. To Catherine Vanderwaart and Dave Maltzan, thank you for bringing me up to speed on all things MBTA.

To my fellow MST's, thank you for the fun times on campus.

To my roommates, Joseph, Zach, and Aaron (Woody, Nutz, and Bike), thank you for helping me take my mind off of school after long days of work. I'll never forget alpine skiing in the streets of Somerville.

Finally to my family, thank you for always being supportive, coming up from New Jersey to visit many times and not getting upset when I forget to call home.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Crowding is an important component of the public transit passenger experience. It can significantly reduce passenger comfort and disrupt service delivery causing passenger travel times to increase. In systems where vehicle miles of service have not been increased enough to meet growing ridership, crowding results in a disincentive to use transit on the very routes and at the times of day where demand is highest. However, traditionally crowding has been expensive to measure and, as a result, it has been difficult to manage. This thesis sets out to improve the ability to measure passenger crowding on buses and therefore to manage it more effectively.

It is likely that crowded conditions create a larger negative perception of public transportation disproportionate to the concentrated number of vehicle trips on which it occurs. Crowding on a small number of high demand links affects the experiences of a proportionally larger number of passenger journeys, so identifying when and where crowded conditions arise can help to better understand the scale of the issue and potential remedies for any transit agency.

This thesis proposes a methodology to estimate passenger crowding on buses using several automated data sources, including inferred passenger origin destination interchange (ODX) as well as automatic passenger counter (APC) data. The Massachusetts Bay Transportation Authority (MBTA) bus network serves as a case study, and a procedure is developed to estimate passenger crowding throughout the network. ODX data is used to expand the coverage of load information from a sample of APC trips and estimate trip level passenger flows in addition to loads.

Using the newly estimated measures of crowding and a set of reasonable assumptions, the impact that various potential sources of crowding will be inferred for each route, direction and time period in order to guide recommendations for reducing and mitigating crowding. The goal is to guide a more effective use of resources rather than to simply increase scheduled frequency for all crowded routes. Increasing scheduled frequency is the most direct response to crowding and produces ancillary frequency improvement benefits, but may not always be financially feasible so it is important to develop a broad range of options.

1.1 Motivation

There are three primary motivations guiding this research:

1. Crowding has not traditionally been well investigated due to the large expense of traditional data collection methods. Recent advances in automatically collected data systems (ACDS) allow for a more nuanced understanding of the passenger crowding experience and make it more affordable to monitor increased crowding as ridership grows. Since crowding occurs on the routes and times of day when demand is high it can have a disproportionate impact on discouraging ridership growth.

2. Transit agencies are always budget constrained. Often times there are not enough available funds to improve conditions in all crowded situations, while simultaneously improving frequency, on time performance, and other elements of the passenger experience. Therefore, there is a need to compare conditions across routes and time periods in order to prioritize resource use.

3. While often caused by budgetary limitations, some agencies are also vehicle constrained, making the identification of situations where crowded conditions can be improved without increasing peak vehicle requirements important.

1.1.1 Passenger Centric Measurement

Passenger crowding is an issue that almost every transit agency faces. An American Public Transportation Association (APTA) survey in 2008 found that 85% of responding agencies were experiencing crowding in some portion of their system. (APTA 2008) [1] Crowding in vehicles and stations not only decrease passenger comfort, making a given service less appealing, but can also lead to unsafe conditions. Therefore, it is important for agencies to continuously monitor crowding throughout their system.

In the case of monitoring bus passenger crowding, traditional sampling methods can provide an overview of crowding due to inadequate frequency for individual routes, but crowding caused by headway and schedule adherence variability is more difficult to measure.

As automatically collected data systems mature, agencies are able to develop more passenger centric performance metrics. In addition to metrics based on operational conditions, such as peak vehicle load or trip running time, data is now increasingly available to better understand how the operation of the system affects passengers' experiences. Both complete implementation of APC systems as well as utilization of the ODX algorithm allow for nearly complete coverage of load estimates while ODX also provides important additional planning information including passenger flow estimates.

1.1.2 Resource Prioritization

Transit agencies often have to implement improvement strategies under budget constraints. In the APTA survey mentioned above, 91% of agencies faced limitations of some form in their response to crowded conditions. Of these, 65% were limited in their response by a lack of funds. (APTA 2008) [1] In these cases, it is important to identify the highest priority routes and time periods to address the most prevalent crowding. These situations could receive improved service as soon as schedules are revised while lower priority routes could wait until more funding becomes available, and other operating strategies might be employed as well.

1.1.3 Strategy Assignment

Agencies also work within bus fleet size constraints. Of the agencies that faced limitations in their response to crowded conditions, 28% were limited by a lack of additional vehicles that could be introduced into the network. (APTA 2008) [1] In these situations, it may not be possible to improve conditions by increasing scheduled frequency in the short term.

However, an agency may still be able to improve conditions systemwide by more effectively utilizing the current capacity in operation. If routes with a high proportion of crowding attributed to unreliability can be identified, then potentially systemwide crowding can be reduced by improving reliability on the identified routes.

In addition, one can even imagine that if the operation of a current high capacity but very unreliable route were improved substantially so that the resulting headways were much more regular, the resulting improved service quality might attract more ridership to fill the increased effective capacity.

1.2 Objectives

This thesis has two main objectives: 1) develop alternative monitoring measures and an automated crowding diagnostic tool that allows for comparison and ranking of conditions across a diverse set of routes; and 2) develop a methodology to infer the sources of crowding for a route during a given time period which would help inform the development of various crowding mitigation techniques to a diverse set of routes.

1.2.1 Alternative Crowding Measures and Diagnostic Tool

The first goal is to create a passenger centric crowding metric that moves crowding measurement away from the traditional method of using the bus trip as the unit of measurement. Instead, a set of alternative passenger-based metrics are explored. Although it is important to keep peak vehicle loads in mind when evaluating crowding as there is a finite amount of capacity within each vehicle and bus trips with loads nearing crush capacity have a high probability of denying passenger boardings, a transit agency should be primarily concerned with how passengers experience crowding on their journeys.

Once a set of alternative metrics are proposed, an automated process can be developed to compare conditions among different routes. Converting the unit of measurement from vehicles to passengers has the added benefit of enabling easier crowding comparisons between routes of varying characteristics (length, frequency, ridership patterns, etc...). Therefore, routes can be objectively ranked by the degree in which their passengers face crowded conditions, allowing a transit agency to prioritize routes and route segments on which to implement various crowding reduction techniques.

1.2.2 Crowding Source Determination

The second objective is to develop a methodology to determine the sources of crowding for each route throughout the day. This could be used to help agencies more efficiently address crowding conditions systemwide given finite resources. While increasing levels of service and frequencies on a route will always decrease crowding, it may not be the most efficient or effective method if a route has service reliability issues. In that case, it might be more effective, for example, to work on improving on-time terminal departure performance and introducing all-door boarding instead; and only use additional vehicle resources on routes on which more crowding can be attributed to scheduled frequency.

Crowding is a symptom of a problem, whether it be inadequate frequency or schedule adherence. Identification of the primary source of the problem can enable more effective action to mediate the problem.

1.3 Methodology

To achieve the objectives mentioned in Section 1.2 a three phase process is used:

1. Develop a bus trip origin-destination-interchange (ODX) inference scaling procedure to provide estimates of both unlinked passenger origin-destination flows and vehicle load profiles

2. Create crowding metrics that consider both the intensity and duration of passenger crowding to evaluate crowding conditions across the entire bus network

3. Determine how much crowding can be attributed to various factors (i.e. scheduled frequency, headway variability, etc...).

1.3.1 ODX Scaling

Complementing APC information with ODX output is the basis for most of the analysis in this thesis. For transit agencies like the MBTA with high penetration rates of reusable identifiable fare media and granular automatic vehicle location (AVL) data, ODX provides estimates of passenger origin-destinations for most passengers traveling within the system by combining automatic fare collection (AFC) and AVL information. The resulting inferred OD matrix can be used to develop both a vehicle load profile and a collection of individual and aggregated passenger origin destination flows. These two measures are used in the development of alternative crowding metrics.

However, a percentage of passenger flows are not inferred in the raw ODX output. In cases when a passenger's subsequent fare transaction does not meet a set of criteria established in the ODX inference algorithm (temporal and spatial thresholds) or a transaction is made in cash (therefore not accompanied by a subsequent transaction), a destination for his/her trip cannot be inferred. There are also passenger trips taken where there is no record in the ODX output. This occurs when a passenger does not interact with the farebox on the vehicle or validation station at a bus stop and therefore has no fare transaction. These two cases cause raw ODX output to undercount reality. A process described in Chapter 3 is used to scale the raw output and create more accurate estimates of vehicle load profiles and passenger origin-destination flows.

1.3.2 Metric Development

The scaled ODX output is then used to create various crowding metrics that can be evaluated. The intent is to allow for comparison among all bus route/30-minute-time-period combinations in the network to better understand the nature of crowding throughout the system. At first, a systemwide analysis is done to show the distribution of weekday crowding among routes and time periods. Then a ranking is produced to identify the highest priority routes and time periods.

The primary metric used in this analysis is cumulative passenger crowding time (CPCT), which is the cumulative amount of passenger time for which passengers spend above a given crowding threshold. This metric takes into account both the intensity (number of passengers above the crowding threshold) and duration (the length of time in which passengers spend in crowded conditions) of crowding. This can be calculated for each bus trip over a given time period (e.g., weekdays in March 2015) and combined in various ways to make comparisons (i.e., temporal distribution of crowding on a given route on an average weekday or the distribution of crowding among routes during a given time period.)

Additional metrics are used to estimate other dimensions of crowding. These include the number of unique unlinked passenger trips on which a passenger experienced crowded conditions, the number of unique passengers who had to stand for at least a portion of their trip, and their average standing duration. These metrics are important for further investigation into why a given route and time period has a high amount of cumulative passenger standing time.

This will be accomplished while respecting the limitations of the current ODX inference process, such as not being able to infer the specific vehicle a passenger boards when he/she pays their fare at a faregate instead of onboard a vehicle, and equipment (AFC, AVL) outages which cause entire vehicle trips to not be observed.

Routes and periods for which this methodology does not well work will be identified and noted.

1.3.3 Crowding Sources

Once the nature of crowding throughout the system is understood and routes and times of high priority are identified, the effect that different potential sources have in creating crowding conditions during each route/direction/time-period combination will be explored. This will help agency decision makers select the most effective strategy to address each crowding situation.

The first step is to determine how much crowding can be attributed to scheduled frequency as opposed to "variability" factors. "Variability" factors are phenomena that cause differences in loads between trips. This isolation of causes is accomplished by comparing the amount of crowding we would expect on a route during a given time period if every trip carried the same number of passengers (i.e., the total observed demand divided by the number of vehicle trips scheduled) and passengers on each trip had the same alighting distributions. Under some assumptions, this ratio is the percentage of crowding that can be attributed to scheduled frequency while the remainder can be attributed to variability factors.

Variability factors are then partitioned into a factor measuring the impact of "dropped trips" (i.e., any scheduled trip not operated for any reason) and day-to-day fluctuations in demand as well as within-period load variability factors. The dropped trip factor captures the amount of crowding resulting from the difference in actual versus scheduled frequency on a specific day. Day-to-day fluctuation in demand captures the additional amount of crowding caused by surges in demand on particular days assuming that passengers are evenly distributed among all actual vehicle trips run during a particular length of time. Any additional crowding is attributed to within-period load variability factors, which cause loads on vehicle trips within a short time period on a specific day to vary. These are factors such as uneven headways, poor schedule adherence and varying passenger alighting distributions and arrival rates.

1.4 Thesis Organization

This thesis is composed of seven chapters. The next chapter provides background information and a literature review on this topic. It includes discussions on the theory of passenger crowding, traditional crowding measurement techniques, applications of detailed load information, uses of ODX, and the MBTA context in which these methodologies will be applied. Chapter 3 describes the process used to scale ODX and its validation results.

Chapter 4 describes how ODX can be used to create a passenger centric crowding metric and applies these metrics to the MBTA network in order to gain a better understanding of passenger crowding and identify when and where passengers are facing the worst crowding. Chapter 5 discusses a methodology to determine the source of crowding for a route during a given time period. Chapter 6 discusses potential crowding reduction strategies and proposes crowding reduction programs for the MBTA context under different resource availability scenarios. Finally, Chapter 7 provides a summary of research findings, recommendations for the transit industry, and potential areas for future research.

Chapter 2

Background

A measure of passenger comfort is an important component of transit agencies' evaluation of the level of service they are providing to their customers. A survey of rail passengers in Australia found that a passenger's satisfaction with crowding accounted for over 20% of their overall satisfaction with the service. (Thompson et. al. 2012) [2] Often times this is done through analysis of passenger loads on vehicles throughout the network. This allows agencies to gain a better understanding of the conditions passengers face while they ride. For example, someone who boards a vehicle with a current passenger load above seating capacity will very likely not have access to a seat and will have to stand for at least a portion of their trip, thus having a potentially less comfortable experience than if a seat had been available.

2.1 Crowding

Crowded conditions as a psychological phenomenon can be defined as the point for which a given person is experiencing more social interaction than desired. Often times individuals regulate social interaction by controlling the amount of personal space they provide for themselves. When this is compromised he/she feels crowded. (Lepore & Allen 2000) [3] Many situations can arise throughout a transit system when an individual's personal space is constrained to the degree where he/she might feel crowded both in vehicles and stations.

Being exposed to crowded conditions can have many negative health effects. Mahudin et. al. (2011) [4] found that the more crowded passengers felt on their transit journey the more stressed and exhausted they were likely to feel upon arrival at their destination.

Passengers also experience two kinds of crowding: objective and subjective. (Li & Hensher 2013) [5] Objective crowding is the set of physical conditions that a passenger experiences. This is often what transit agencies measure and use to evaluate the service they are providing. Common metrics are passenger density measures such as vehicle load or passengers per square meter.

Subjective crowding is a passenger's perception of the conditions they face. This is based on his/her objective crowding but is influenced by previous experiences, expectations, and personal preferences. Two passengers facing the same objective crowding could be facing very different subjective crowding if their values differ. The latter is what passengers generally use to inform mode choice and trip making decisions. (Thompson et. al. 2012) [2]

2.1.1 Objective Crowding

Objective crowding conditions have traditionally been measured as passenger density of which there are multiple methods of calculation. Some agencies use vehicle loads as a percentage of seated capacity often called a load factor. This allows for comparisons among routes that use similarly configured but different sized vehicles.

Other agencies use standing density, a measure of how close passengers would need to stand at different vehicle loads. This is likely a more nuanced metric compared to load factors especially when comparing conditions across vehicles with very different configurations. However, it assumes that all floor space is equally desirable and that standing passengers will evenly distribute themselves throughout the vehicle. (Tirachini et al 2013) [6]

2.1.2 Subjective Crowding

Subjective crowding is more difficult to measure as it is very individualized. There are many factors besides passenger density that can affect how someone perceives a situation. Hirsch & Thompson (2011) [7] identified eight factors that affect a passenger's perception of crowding: 1) Expectations of Conditions, 2) Environmental Charateristics, 3) Quality of Communication from Transportation Provider, 4) Perception of Personal Control of Situation, 5) Amount of Crowding Caused by Delays, 6) Perception of Risk to Safety and Health, 7) Passenger's Emotions, and 8) Behavior of Other Passengers. This can also vary between cultures. Hall (1966) claims that people from "contact" cultures where personal contact is an everyday occurrence are less prone to feeling crowded than people from "non-contact" cultures.

However, objective crowding does play a significant role in how passenger's perceive crowding. A meta-analysis of crowding studies performed by Wardman & Whelan (2010) [8] found that crowding increases a passenger's "disutility" of travel time (the amount of money an individual would be willing to pay to decrease travel time) for both passengers standing and sitting. They were able to quantify this effect through value of time multipliers. They found a positive correlation between load factor and the value of time multipliers indicating that as vehicles become more crowded travel time becomes more onerous for passengers. Table 2-1 describes their findings. These higher disutility rates on routes and times of day when demand is high act to inhibit ridership growth.

Recognizing different "disutility" rates can allow resource priorities to be set between increasing frequencies on low frequency routes to improve passenger accessibility versus improving frequencies on high frequency, high demand routes to reduce crowding.

The authors also mentioned that, although there was an inadequate amount of research studying the effect that length of time spent in crowded conditions has on a passenger's value of time to draw significant conclusions, they suspected that the two were positively correlated as well.

Table 2.1: Crowding Value of Time Multipliers by Load Factor

Load Factor	Seated	Standing
100%	1.17	2.14
150%	1.40	2.50
200%	1.66	2.92

Source: Wardman & Whelan (2010) [8]

2.2 Operational Impacts

Crowding also impacts a transit agency's ability to deliver a fast reliable service to its customers. Lin & Wilson(1992) [9] find that passenger loads are positively correlated with dwell time and that its effect increases as loads become higher. This increases passengers' in-vehicle travel time.

Crowding can also affect a passenger's total journey travel time. If crowding is prevalent through out a transit system and causes some vehicles to reach crush capacity there is a possibility of passengers not being able to board the first vehicle that arrives at his/her stop. This causes not only his/her expected travel time to increase but also the variability of his/her possible travel times. Senna (1994) [10] shows that passengers value both aspects.

Uniman (2009) [11] created a metric (Reliability Buffer Time) to evaluate how this unreliability affects the amount of time a person allocates to make a trip. It is the difference between the median travel time and a percentile that reflects their risk aversion to being late (the 95^{th} is commonly used) between an origin and destination during a given time period. The more variable the travel times the more time a passenger is required to allocate as a buffer in order to reach their destination at the desired time. Therefore, even if a given passenger's trip is not delayed by denied boardings, if they have been denied in the past they will have likely allocated more time to make the trip than if travel times were more consistent.

Crowding on one vehicle can lead to crowding conditions across an entire route. Adebisi (1986) [12] shows that variations in load from fluctuations in passenger arrival rates can cause buses to bunch, two vehicles running in closer proximity than scheduled, since higher loads cause longer dwell times. A vehicle on which more passengers than average board will likely operate slower than average due to increased dwell times at each stop. If passenger arrival rates return to average levels the following vehicle is likely to catch up to the first since its dwell times will be shorter and, as a result, the gap (i.e., headway) between the following bus and its leader will also decrease as each moves along the route, resulting in the following bus picking up fewer and fewer passengers and traveling faster than average. The third vehicle is then likely to experience a longer than average headway since the second is traveling faster than average. It is likely to pick more passengers and travel slower than average. This process then repeats itself causing high load and headway variability along a route.

2.3 Passenger Recruitment and Retention

It is in the interest of agencies to maintain a certain level of comfort for their passengers in order to retain riders. Customers increasingly have multiple modes available to make a trip and therefore could be persuaded to switch away from transit if conditions deteriorate. If an occasional passenger has a pleasant experience, they are more likely to continue taking transit and could eventually become a regular passenger. (Thompson et. al. 2012) [2]

2.4 Standards

To ensure some level of passenger comfort, agencies often use vehicle "peak" loads to set frequencies on high demand services. Average maximum vehicle load, often referred to as average peak load, of trips departing over a given time period have been traditionally compared to a peak load threshold to determine whether a route meets a crowding standard. The threshold is often based on load factors. An example of this crowding standard would be : "no route/direction combination should have an average peak load over 140% of seating capacity during any 30 minute period" within the designated peak periods. Any route that fails this standard would be prioritized for an increase in frequency the next time schedules are revised.

There are many variations of this performance standard framework. Some agencies have different thresholds for different time periods. For example, Southeastern Pennsylvania Transportation Authority (SEPTA) uses a 100% load factor during the off-peak periods. This increases to values ranging from 131% to 174% during the peak periods depending on the specific bus model. (SEPTA 2014) [13] This recognizes that in order to effectively meet the high travel demand during the peak periods throughout the entire network some standing passengers must be tolerated. It also notes that the configuration of different bus models leads to different ratios of seats to floor space which allows some models to better handle large numbers of standing passengers comfortably.

Some agencies have various thresholds for different levels of route frequency. The Los Angeles County Metropolitan Transportation Authority has higher thresholds for higher frequency routes. During the peak periods, load factor thresholds range from 140% for routes with headways under 10 minutes to 100% for routes with headways over an hour. (LA Metro 2015) [14] This recognizes that the higher the load factor threshold the more likely a passenger is to be denied boarding along a route. Therefore, this standard reduces the likelihood that a passenger is denied boarding on a low frequency route where the travel time penalty imposed by not being able to board the first bus that arrives would be significant. If an individual is unable to board a bus on a route with headways of 10 minutes they will likely only have to wait another 10 minutes for the next bus while a passenger unable to board a bus on a route with 60 minute headways will have to wait another hour.

Agencies also have different thresholds depending on where a route operates. In addition to the standards mentioned above, SEPTA also restricts routes that utilize limited access highways to a 125% load factor. This accounts for the additional safety concerns that arise from high speed travel.

Finally, as load information becomes more widely available, some agencies are taking a more granular and passenger based approach to creating standards. The San Francisco Municipal Transportation Agency (SFMTA) measures the percentage of bus trips that exceed the crowding threshold in the peak direction during the peak periods. (SFMTA 2013) [15] This reveals some crowding that occurs due to variation in loads between adjacent trips that might be masked by average measures.

Regional Transportation District (RTD), which serves the Denver, Colorado metropolitan area, requires that during the peak periods no passenger should have to stand for longer than 15 minutes. (RTD 2002) [16] This ensures that when passengers do have to stand, it is for a limited amount of time.

2.5 Crowding Measurement

Methods of calculating passenger loads and crowding have changed as the technology available to transit agencies improves. Initially, all measurements were performed manually by human "checkers" on a sample of bus trips. Recently, with the development of automatically collected data systems (ACDS), it became possible to automatically estimate ridership and loads. At first, while these ACDS systems were initially being introduced into the transit agencies, they were only installed on a small percentage of the bus fleet. As ACDS technologies matured and became widely adopted, some agencies have fully implemented these technologies throughout their bus fleets in order to obtain a more granular understanding of loads and crowding through their network.

2.5.1 Manual Measurement

Before the advent of automatically collected data systems, load estimation was done manually. For bus operations, there were three primary methods used to collect the required data: counts done by the bus operator, ride checks, and point checks. (Kittelson Associates et al 2003) [17] Operator counts would entail the bus driver counting boardings and alightings at each stop. This allows an agency to manually collect data without having an employee solely dedicated to data collection but is feasible only on low ridership routes. From these on-off totals, loads can be calculated.

Ride checks have an employee other than the operator ride the bus the entire length of the route. He or she then tracks boardings and alightings at each stop. This can be done potentially along with other tasks such as distributing surveys or taking running time splits.

If the location where peak load of a bus route usually occurs during a given time frame is known, often called the peak load point, it is then possible to gain an approximate understanding of peak load variability during a given time period through the use of point checks. This involves having an employee stationed at the peak load point. He or she would then count the number of passengers on each vehicle that passes through that point on the route. There is some margin of error in this calculation since there is no guarantee that the peak load will occur at the location at which the employee is stationed and without disrupting service it is difficult to get an entirely accurate passenger count. However, it allows an agency to understand peak load variation between trips. Since the marginal cost of these manual data collection techniques performed by a separate employee is large, they are only done on a sample basis.

2.5.2 Automatic Measurement

With the adoption of automatic data collection systems, agencies are able to collect more data at significantly lower marginal cost. The systems most directly related to estimating passenger loads and crowding are automatic passenger counters (APC). They consist of either infrared sensors or tread pads located at all doors of a vehicle that track the number of people who board and alight every time the doors are opened. They are also often linked to a GPS unit that tracks each vehicle throughout its run allowing the assignment of door openings to stops and development of a fine grain running time analysis.

Full deployment of APC systems throughout a bus fleet allows agencies to have a comprehensive understanding of passenger loading and crowding. Variation of passenger loads between trips and days which are often masked by average values can be studied. However, even with these benefits of universal implementation, only a small percentage of agencies have implemented APC systems throughout their entire bus fleet. According to a survey by Daniel Boyle in 2008 (Boyle 2008) [18] of the agencies using APC systems in their bus fleet, only 27% had devices installed in all vehicles with 62% of responding agencies having fewer than 50% of vehicles equipped. The author expects the number of agencies with full deployment to increase as the costs of installation and maintenance of the systems decrease and purchases are included as original equipment in intelligent transportation system (ITS) equipped vehicle purchases.

2.6 Applications of Passenger Load Data

If an agency has passenger load information available for every trip, researchers and agency analysts have the ability to empirically estimate how passengers experience crowding. For example, Furth et. al (2006) [19] describe how peak load variation can drastically change the passenger experience in terms of crowding. They use the example of a theoretical bus route with an average peak load of 40.3 passengers operating with vehicles of a capacity of 42. The average peak load would indicate that passengers on the route are not experiencing crowding, however, there is a significant amount of peak load variation, meaning that there are a significant number of very crowded and relatively empty trips. On the trips with high peak loads crowding occurs. Figure 2-1 shows an example of this type of peak load distribution. Almost 50% of trips have a peak load over seated capacity.

They also show that passengers experience crowding differently even within a given crowded vehicle trip. At the peak load point of trips with peak loads over the seating capacity, there are two general groups of passengers: standees and seated passengers. For seated passengers, additional passengers above the seating capacity do not affect their experience much besides making it more difficult to exit the vehicle. For standees, each additional passenger above the seating capacity means they will have less personal space in the vehicle. For the same trips mentioned above, Furth et. al. grouped passengers into different crowding experiences (seated, seated without a neighbor, and standing at various levels of passenger loads) at the peak load point.

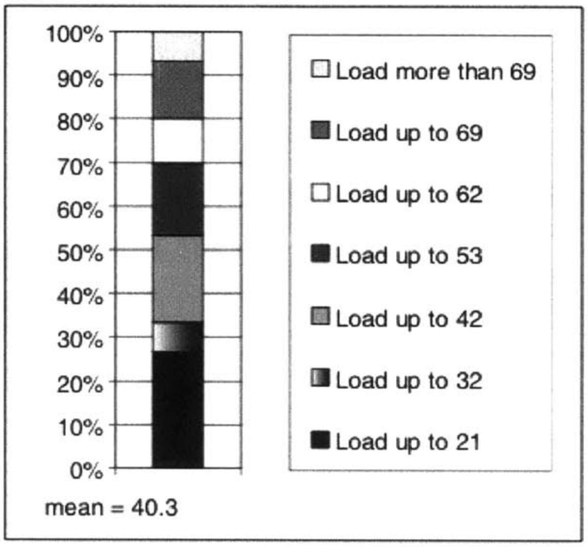


Figure 2-1: Potential Distribution of Trip Peak Loads. Source (Furth et. al. 2006) [19]

The distribution shows that even though almost 50% of trips have peak loads over seating capacity, most passengers are seated on a crowded bus. Therefore only 20% of passengers on these trips are standing at the peak load point. Figure 2-2 shows this distribution. By examining a service from the individual passenger perspective, Furth et. al. present a more nuanced and perhaps more important system estimate of comfort than the traditional methods would generally produce.

2.6.1 Headway Variation Affect On Load Variation

As shown in the previous section, routes with significant load variation generally have more passengers facing crowded conditions than routes with relatively consistent loads. This leads to the question of what is causing load variation. Combining passenger load information with automatic vehicle location (AVL) records that track the location of vehicles can help to address this question. Strathman & Kimpel (2003)[20] used AVL and APC data to look at the relationship between deviation from scheduled headway at the peak load point and peak load. They found that deviation from the scheduled headway was positively correlated with peak load meaning that vehicles that faced a longer than scheduled headway at the peak load point would likely be

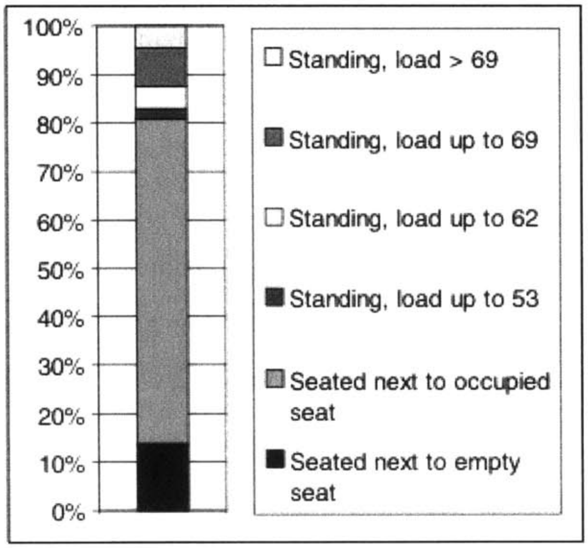


Figure 2-2: Potential Distribution of Passenger Crowding Experiences. Source (Furth et. al. 2006) [19]

more crowded than trips that faced shorter than scheduled headways. Assuming constant passenger arrival rates, this occurs since passengers have more time to arrive at bus stops as headways increase. They also discovered that much of this headway deviation could be attributed to schedule deviation at the origin terminal.

Milkovits (2008) [21] found that changes in the percentage of bus trips operating in bunched situations are positively correlated with the percentage of passenger time spent in crowded conditions. As headways become more regular passengers are more evenly distributed among vehicles, reducing the amount of crowding.

2.7 Uses of Passenger Origin Destination Interchange Inference Data

Another application of ACDS is estimating the origin and destination of passengers' linked and unlinked trips through a process known as origin destination interchange (ODX) inference. ODX combines automated vehicle location (AVL) data with automatic fare collection (AFC) data that records the fare collection device and timestamp for every fare transaction. Transactions on the same fare media can then be grouped together, and are assumed to be a single passenger. Subsequent fare transactions on the same fare media are used to determine origins and destinations of a passenger throughout the day. In an "open" fare payment system, where the passenger only interacts with the farebox upon entry to the system, a destination is inferred for each unlinked trip by determining the closest possible destination that can be reached on the mode of the current transaction to the location of the next transaction. On the last trip of the day, the target destination is the location of the first transaction. The algorithm assumes that passengers end the day at the same location from which they started.

Unlinked trips are combined together to create multistage linked trips if two adjacent unlinked trips meet certain criteria. These criteria include things such as the destination of the first trip being within a distance threshold of the origin of the second, the departure time of the second trip occurring within a time threshold of the arrival of the first, etc... Figure 2-3 shows a graphical representation of this process.

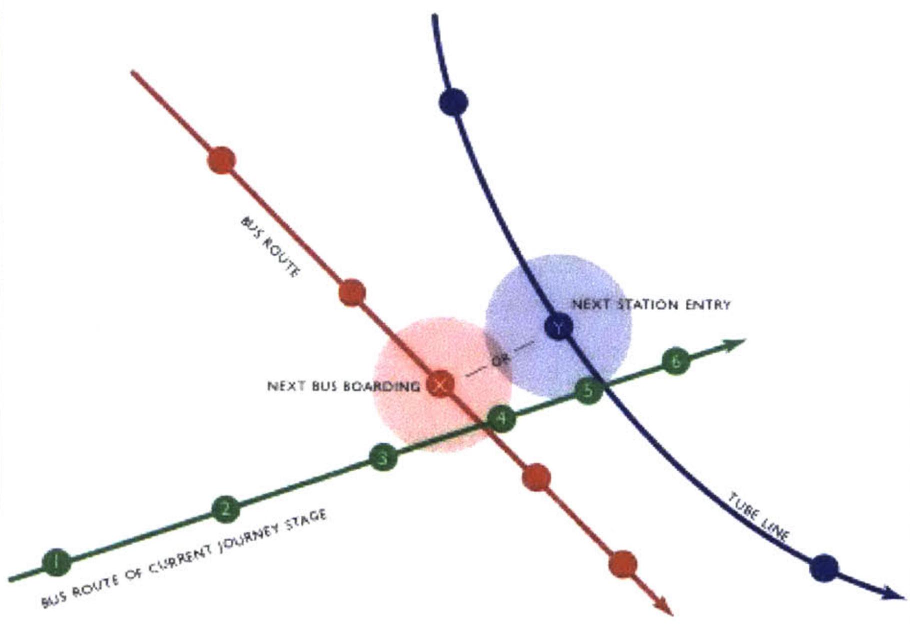


Figure 3-8. Example of destination inference for a passenger trip on a bus route (green) followed by another bus trip (red) or by a rail trip (blue).

Figure 2-3: Example of ODX Destination Inference Process. Source (Gordon 2012)

This wealth of passenger OD information allows agencies to better understand how passengers use their network. Zhao (2004) [22] uses this process to produce an OD matrix for rail passengers in the Chicago Transit Authority's heavy rail network. Cui (2008) [23] modifies this process to produce route level and subnetwork OD matrices for the CTA bus network. Gordon (2012) [24] combines both processes to link bus and rail trips together to infer a passenger's linked trip OD within the entire Transport for London network. This provides a granular estimate of how passengers use a transit network to move throughout a metropolitan area.

Also for TfL in London, Wang (2010) [25] uses ODX to develop bus route aggregate load profiles. She uses this to look at load and ridership variation between days and time periods. However, since there were a significant number of passengers who paid their fare with a media other than an Oyster Card (TfL's smartcard fare media) and some passenger trips did not have inferred destinations Wang needed to develop a scaling process to estimate the total load for each trip. She develops a two-step process that uses electronic ticket machine (ETM) trip ridership counts as control totals. ETM totals are representative of total ridership since bus drivers are instructed to record all boarding passengers regardless of whether a fare is collected. This process enables the calculation of period average flows as well as trip-by-trip distributions.

Two Step ODX Scaling Process

1. Passengers whose destinations are not inferred with ODX are assumed to have the same alighting distribution (probability of alighting at any stop further along the route given that a passenger boarded at a specific stop) as the passengers for whom a destination was inferred and are assigned destinations with this distribution.

$$F'_{ODT} = F_{ODT} + B_{OT} * \frac{\sum_T F_{ODT}}{\sum_T \sum_D F_{ODT}} \tag{2.1}$$

Where

- F_{ODT} is the passenger flow between origin stop O and destination stop D on bus trip T

- B_{OT} is the number of passenger boardings at origin stop O on bus trip T for which a destination is not inferred

2. Passengers for whom there is no inferred origin, and therefore also no destination, are assumed to have the same origin destination distribution as the passengers in the previous step. Therefore, the flows calculated in step 1 are multiplied by the ratio of ETM boarding totals to ODX boarding totals.

$$V_{OD} = F'_{ODT} * \frac{T_T}{\sum_O \sum_D F'_{ODT}} \tag{2.2}$$

Where

- V_{OD} is the scaled passenger flow between origin stop O and destination stop D
- T_T is ETM boarding total for trip T

ODX can also be used to analyze the travel behavior of individual passengers. Dumas (2015) [26] adapted the methodology Gordon used in London for the MBTA network. He then used this to compare passenger travel times of residents of minority neighborhoods to residents of non minorities neighborhoods. Viggiano (2013) [27] used ODX to analyze passenger behavior on corridors with multiple routes.

2.8 MBTA Context

The case study for this thesis will be the bus network of the Massachusetts Bay Transportation Authority (MBTA). The MBTA operates the public transportation system in the Greater Boston, Massachusetts Metropolitan area which includes bus, light, heavy and commuter rail, and ferry services. It's large and diverse service area includes 175 municipalities with a combined population of almost 5 million. It is heavily used with almost 1.3 million unlinked passenger trips systemwide and almost 400,000 bus passenger trips on a typical weekday. (MBTA 2014)[28] Ridership growth has out paced increases in vehicle revenue miles over the past 20 years causing increased crowding. [34] In addition to crowding that occurs on the bus network there is also significant crowding on two of the three rapid transit lines, (Red and Orange) and the Green Line light rail during the peak periods at certain locations.

2.8.1 MBTA Bus Network

Of the modes that the MBTA operates, the bus network is by far the most diverse. There are 170 different routes ranging from high demand urban crosstown and arterial routes to express commuter routes to neighborhood and suburban feeder routes to early morning routes.

This diversity leads to a wide range route characteristics. Typical weekday ridership ranges from over 14,000 passengers for Route 39, an urban arterial route to 45 for Route 431, a suburban feeder route. The length of routes also varies significantly from 2.2 miles one way for Route 26, a neighborhood feeder route, to over 16 miles for Route 34E, a suburban feeder route with many North Shore express routes also exceeding 15 miles one-way. As one may expect, differing levels of demand and route length lead to varying frequencies. During the morning peak period, route headways vary from as low as 4 minutes for Route 7, an urban arterial route, to over an hour for some suburban express routes. (MBTA 2014) [28] Ridership patterns vary as well with some routes having significant amount of rider turnover while on others most passengers board towards the beginning of the route and alight at the end.

2.8.2 Current Crowding Measurement and Standards

The MBTA currently has APC systems installed on approximately 14% of its bus fleet. In order to get load estimates for the entire system, these vehicles are rotated among different routes ensuring that almost every scheduled bus trip (e.g., Inbound Route 1 trip that departs at 8:30 AM on weekdays) has at least one APC record during a given season. This provides a sample of loads that can then be used to evaluate passenger crowding and comfort.

Like most transit agencies, a crowding standard is derived from peak load analysis using the APC data collected through the process described above. The standard is as follows (MBTA 2010) [29]:

Average peak load cannot exceed 140% of seating capacity of a vehicle for any 30 minute duration during the following high demand periods of a weekday:

Early AM (6:00-6:59AM)
AM Peak (7:00-8:59AM)
Midday School (1:30-3:59PM)
PM Peak (4:00-6:29PM)

For all other weekday and weekend periods, average peak load cannot exceed 100% of the seating capacity of a vehicle for any 60 minute duration. The standard also requires that loads at the beginning and end of the service day for each route be evaluated to see if the span of service needs to be adjusted.

While this standard enables analysis of crowding caused by inadequate frequency, it does not take into effect crowding due to variability factors and makes comparison of conditions among routes with different characteristics more difficult.

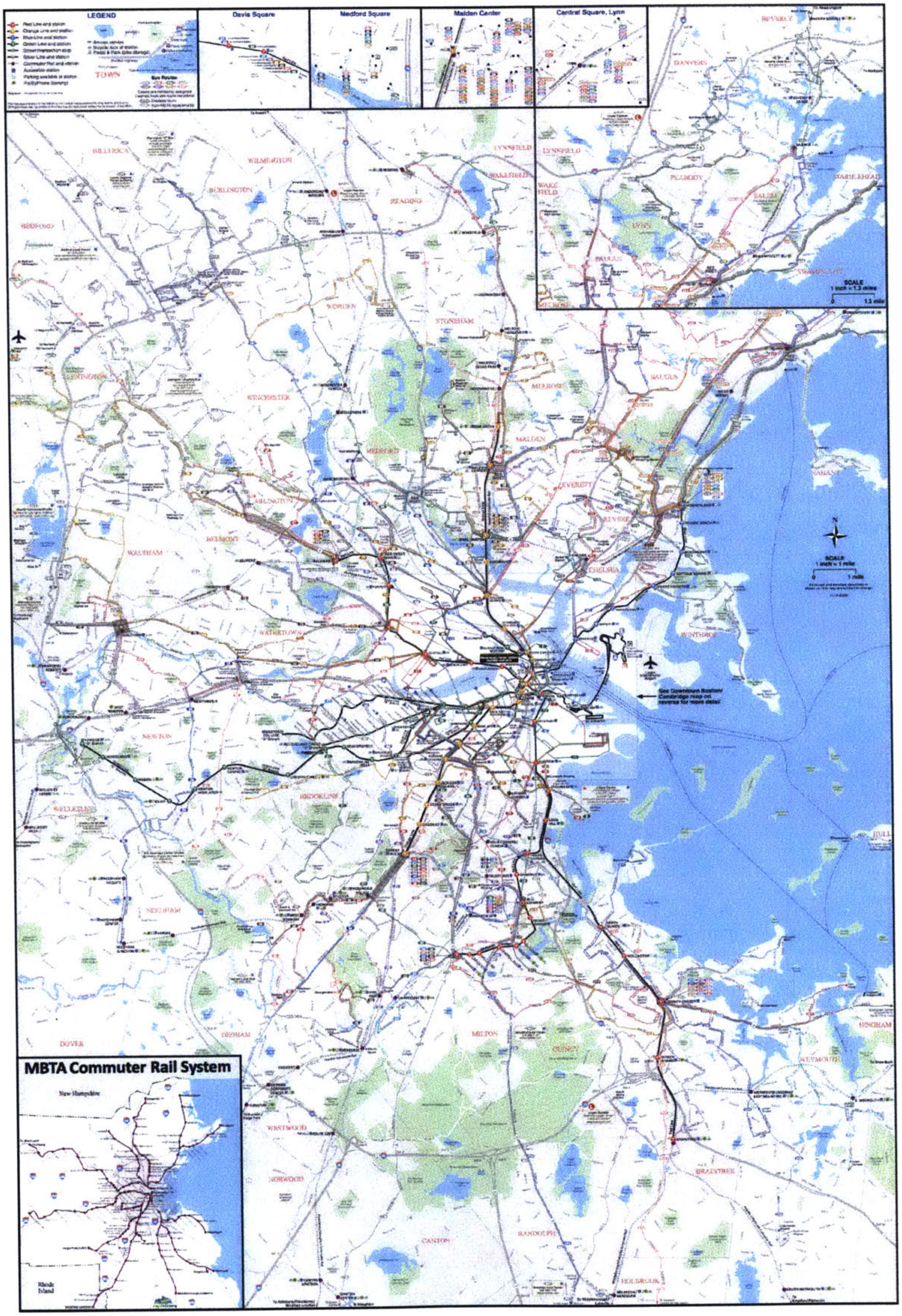


Figure 2-4: MBTA Network. Source (MBTA System Map 2014)

Chapter 3

Procedures for Trip Level ODX Scaling

Although raw ODX output can show relative magnitudes of passenger flows among routes and trips, in order to get accurate estimates of passenger loads and flows, the output from the ODX inference procedure must be scaled. Without this process, all flows are likely underestimated. This is due to two main reasons: some passenger trips only have an origin inferred and some passenger trips are not observed at all.

The first instance occurs when either a fare media is only seen once during a given day, as in the case of a passenger paying their fare in cash or only taking a single MBTA ride; or when their succeeding tap does not meet the criteria needed to infer a destination. In these cases the ODX algorithm does not assign a destination for this passenger trip.

The second instance occurs when a passenger does not interact with the farebox on the vehicle or a fare validator at a bus stop. Unlike on the MBTA's heavy rail system where passengers must pay their fare before passing through faregates and into a station almost all bus passengers pay their fare on the vehicle on which they ride. This leads to a significant number of passengers not interacting with the farebox for a number of reasons: some may be exempt from paying such as children and blind individuals, some may be purposely evading fare payment, and some may be waved on by the operator due to either farebox malfunction or operational efficiencies.

To account for these instances of undercounting passengers, a two step scaling process was developed. It is similar to one Wang implemented on the Transport for London bus network. (Wang 2010) [25] The first step, referred to as destination scaling, allocates destinations to passenger trips for which only origins are inferred based on a seed matrix of passenger flows of final stage trips. The second step, referred to as boarding scaling, accounts for passengers who did not interact with the farebox by multiplying the passenger flows calculated in the first step by a scaling factor derived through comparing trip boarding totals calculated with the ODX inference algorithm to trip boarding totals calculated with APC on bus trips for which both systems are operating. This process results in the expected passenger flow between any two stops on a given bus trip.

3.1 Destination Scaling

Destination scaling entails allocating destinations for every boarding recorded in ODX. This is necessary as approximately 40% of bus passenger trips only have origins inferred with no destination identified. Passenger flows are scaled by first finding the probability of a passenger without an inferred destination alighting at any stop along a bus route given that they boarded a given bus trip at a given stop, referred to as the alighting distribution. This probability is used to allocate passenger flow among all possible origin-destination pairs. Then these newly calculated passenger flows are added to the passenger flows inferred through ODX to produce destination scaled passenger flows as shown in equation 3.1.

$$F_{ODT} = F_{ODTI} + B_{OT} * P_{ODFT} \tag{3.1}$$

Where

- F_{ODT} is the passenger flow between origin stop O and destination stop D for bus trip T
- F_{ODTI} is the passenger flow inferred by ODX between origin stop O and desti-

nation stop D for bus trip T

- B_{OT} is the number of passenger boardings at origin stop O on bus trip T
- P_{ODFT} is the probability that a final stage F passenger boarding bus trip T at origin stop O alights at destination stop D

3.1.1 Seed Matrix Development

The alighting distribution for a given stop on given bus trip is calculated through the development of seed matrices. These matrices are created by aggregating final stage passenger flows along a given route/direction/variation combination for every thirty minute period throughout the day. Final stage trips are defined as any unlinked passenger trip for which there is no successive unlinked trip associated as part of a longer linked trip.

Then, for each boarding location, the percentage of passenger trips with alightings at each of the possible destination stops is calculated. This percentage becomes the alighting probability for each stop as shown in Equation 3.2. The distribution of these alighting probabilities across all possible stops is the alighting distribution for a given boarding stop.

$$P_{ODFM} = \frac{F_{ODFM}}{\sum_D F_{ODFM}} \tag{3.2}$$

Where

- P_{ODFM} is the probability that a passenger on their final stage F during time period M boarding at origin stop O alights at destination stop D
- F_{ODFM} is the final stage F passenger flow between stops O and D during time period M

There are three assumptions needed to create a seed matrix based on inferred ODX passenger flows:

1. **ODX inference is unbiased.** The process assumes that the alighting distribution of passengers without inferred destinations is the same as passengers with final stage inferred destinations. While it is possible that some destinations have lower destination inference rates due to physical location or ridership behavior there is no indication that these differences are large enough to substantially change passenger flows patterns.

2. **Passengers without inferred destinations are on the final stage of their linked trip.** The process also assumes that all transferring passengers will have an inferred destination. This assumption is important as transferring passengers are likely to have different alighting distributions than final stage passengers. Transferring passengers are much more likely to alight at transfer points such as heavy rail or bus stations while final stage passengers are likely to have a more "spread" alighting distribution among all possible stops. Figure 3-1 shows the difference in alighting patterns between transferring and final stage passengers on Weekday Outbound Route 93 trips between 5:00-5:30PM. Red cells represent destinations with high alighting probabilities while yellow and green represent destinations with medium and lower alighting probabilities, respectively. The boarding stops with a zero alighting probability for all other stops are stops for which not a single passenger trip with inferred destination was recorded during the seed period. Therefore, no alighting distribution could be derived. Final stage passengers are more likely to alight in Charlestown, a residential neighborhood of Boston while transferring passengers are more likely to alight either in downtown Boston, where there are many transfer opportunities or at Sullivan Station, a heavy rail station.

   Most transferring passengers will have destinations inferred as they will likely have a successive farebox interaction spatially and temporally close to a possible destination of their current stage. However, there is a possibility that a

Alighting Distribution of Final Stage Passengers on Outbound Route 93 Boarding Between 17:00 and 17:30	Destinations																				
Origins	Pearl St @ Congress St	Congress St @ State St	Congress St @ North St	Congress St @ Haymarket Sta	N Washington St @ Thacher St	N Washington St @ Commercial St	Chelsea St @ Warren St	Chelsea St @ Constitution Rd	Chelsea St @ Fifth St	Vine St @ Moulton St	121 Bunker Hill St opp Lexington St	Bunker Hill St @ Polk St	Bunker Hill St @ Pearl St	Bunker Hill St @ Sackville St	Bunker Hill St @ N Mead St	Bunker Hill St @ St Martin St	Bunker Hill St @ Baldwin St	Bunker Hill St @ Medford St	529 Main St	Cambridge St @ Maffa Way	Sullivan Station - Upper Busway
Devonshire St @ Milk St	0.04	0.01	0.01	0.03	0.01	0.00	0.10	0.03	0.09	0.13	0.11	0.08	0.08	0.05	0.02	0.00	0.01	0.01	0.00	0.00	0.18
Pearl St @ Congress St		0.02	0.01	0.04	0.03	0.00	0.19	0.03	0.04	0.09	0.07	0.10	0.09	0.09	0.04	0.01	0.01	0.03	0.00	0.00	0.10
Congress St @ State St			0.00	0.02	0.03	0.00	0.12	0.03	0.07	0.15	0.10	0.07	0.13	0.05	0.04	0.00	0.00	0.01	0.00	0.00	0.14
Congress St @ North St				0.07	0.01	0.00	0.17	0.03	0.09	0.08	0.12	0.10	0.06	0.07	0.04	0.02	0.01	0.01	0.00	0.00	0.13
Congress St @ Haymarket Sta					0.04	0.00	0.20	0.03	0.12	0.12	0.15	0.12	0.07	0.02	0.03	0.00	0.01	0.00	0.00	0.00	0.09
N Washington St @ Thacher St						0.13	0.25	0.00	0.08	0.17	0.04	0.04	0.00	0.08	0.00	0.04	0.00	0.00	0.00	0.00	0.17
N Washington St @ Commercial St							0.11	0.11	0.05	0.11	0.16	0.00	0.05	0.00	0.11	0.00	0.05	0.05	0.00	0.00	0.21
Chelsea St @ Warren St								0.00	0.10	0.20	0.00	0.00	0.10	0.10	0.10	0.00	0.10	0.00	0.00	0.00	0.30
Chelsea St @ Constitution Rd									0.17	0.17	0.00	0.00	0.00	0.00	0.00	0.17	0.00	0.00	0.00	0.00	0.50
Chelsea St @ Fifth St										0.06	0.00	0.00	0.06	0.06	0.00	0.00	0.00	0.06	0.00	0.00	0.76
Vine St @ Moulton St											0.04	0.04	0.08	0.00	0.00	0.00	0.04	0.13	0.00	0.00	0.67
121 Bunker Hill St opp Lexington St												0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.86
Bunker Hill St @ Polk St													0.13	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.80
Bunker Hill St @ Pearl St														0.00	0.00	0.00	0.00	0.33	0.00	0.00	0.67
Bunker Hill St @ Sackville St															0.00	0.00	0.00	0.00	0.00	0.00	0.00
Bunker Hill St @ N Mead St																0.00	0.00	0.00	0.00	0.00	1.00
Bunker Hill St @ St Martin St																	0.00	0.00	0.00	0.00	0.00
Bunker Hill St @ Baldwin St																		0.00	0.00	0.00	0.00
Bunker Hill St @ Medford St																			0.00	0.00	1.00
529 Main St																				0.00	0.00
Cambridge St @ Maffa Way																					0.00

(a) Final Stage Passenger Alighting Distribution

Alighting Distribution of Transferring Passengers on Outbound Route 93 Boarding Between 17:00 and 17:30	Destinations																				
Origins	Pearl St @ Congress St	Congress St @ State St	Congress St @ North St	Congress St @ Haymarket Sta	N Washington St @ Thacher St	N Washington St @ Commercial St	Chelsea St @ Warren St	Chelsea St @ Constitution Rd	Chelsea St @ Fifth St	Vine St @ Moulton St	121 Bunker Hill St opp Lexington St	Bunker Hill St @ Polk St	Bunker Hill St @ Pearl St	Bunker Hill St @ Sackville St	Bunker Hill St @ N Mead St	Bunker Hill St @ St Martin St	Bunker Hill St @ Baldwin St	Bunker Hill St @ Medford St	529 Main St	Cambridge St @ Maffa Way	Sullivan Station - Upper Busway
Devonshire St @ Milk St	0.01	0.02	0.01	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.00	0.72
Pearl St @ Congress St		0.03	0.01	0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.51
Congress St @ State St			0.01	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.03	0.07	0.01	0.63
Congress St @ North St				0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.81
Congress St @ Haymarket Sta					0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.03	0.03	0.90
N Washington St @ Thacher St						0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.93
N Washington St @ Commercial St							0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Chelsea St @ Warren St								0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.90
Chelsea St @ Constitution Rd									0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.08	0.88
Chelsea St @ Fifth St										0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.98
Vine St @ Moulton St											0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.00	0.96
121 Bunker Hill St opp Lexington St												0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.00	0.94
Bunker Hill St @ Polk St													0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.99
Bunker Hill St @ Pearl St														0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.97
Bunker Hill St @ Sackville St															0.00	0.00	0.00	0.00	0.00	0.00	1.00
Bunker Hill St @ N Mead St																0.00	0.00	0.00	0.00	0.00	1.00
Bunker Hill St @ St Martin St																	0.00	0.00	0.00	0.00	1.00
Bunker Hill St @ Baldwin St																		0.00	0.00	0.00	1.00
Bunker Hill St @ Medford St																			0.00	0.00	1.00
529 Main St																				0.00	1.00
Cambridge St @ Maffa Way																					1.00

(b) Transferring Passenger Alighting Distribution

Figure 3-1: Passenger Alighting Distributions for Weekday Outbound Route 93 Trips, 17:00-17:30

transferring passenger's destination is not inferred or they are categorized as a final stage passenger instead.

The first instance could occur if a passenger does not interact with the farebox on their subsequent unlinked trip. This occurs to some degree on all modes but is most likely to occur when transferring onto either the bus or surface light rail networks due to their onboard fare payment system.

The second instance could occur if the alighting of the first unlinked trip is temporally too far from the boarding of the second trip. This could be someone waiting a significant amount of time for their connecting service or running an errand in between unlinked trips. The ODX algorithm is conservative on its linking of trips so it is possible that some linked trips are unintentionally split though this becomes a philosophical question of what entails a transfer. This assumption potentially skews estimated passenger flows towards final stage distributions.

3. **Expected passenger flows between stops are fractional.** Since the destination scaling process distributes passenger flows of passengers without inferred destinations by the distributions derived from the seed matrices as shown above, it allocates some amount of passenger flow to almost every downstream stop as over the duration for which the seed is developed passengers make trips between most stops. This means that a single passenger likely has his or her flow divided among multiple OD pairs.

   For the creation of average load profiles this does not have a major impact as over the course of many trips there will likely be some portion of passengers traveling between any two possible stops. However, in the creation of trip level load profiles, depending on the ridership patterns of the route, this could lead to more variation in loads than otherwise occurs in reality. Actual passenger flows are composed of whole passengers therefore where a given passenger alights in reality compared to the alighting distribution will affect load calculations.

   This is mitigated to a degree with the development of trip specific seed matrices as described in Section 3.1.2 below. This process limits potential destinations to stops that were made on a given trip thus consolidating passenger flow to some degree.

3.1.2 Trip Specific Seed Matrix

Seed matrices can be disaggregated further from period averages down to the trip level. This is done by updating the alighting distributions based on the set of stops that a vehicle made on a given trip. A stop will only be considered as a potential destination if the vehicle opened its doors at that location. Therefore, the updated alighting probability can be determined by dividing the alighting probability for destination stop D given a passenger boarded at origin stop O by the sum of alighting probabilities at all stops in the set of stops that the vehicle made S_T as is shown in Equation 3.3.

$$\forall D \in S_T : \; P(ODFT \mid S_T) = \frac{P_{ODFM}}{\sum_D P_{ODFM}} \tag{3.3}$$

Where

- P_{ODFT} is the probability that a passenger on their final stage F on bus trip T boarding at origin stop O alights at destination stop D
- P_{ODFM} is the probability that a passenger on their final stage F during time period M boardings at origin stop O alights at destination stop D
- S_T is the set of stops made on bus trip T

The external announcement system onboard each vehicle enables the determination of this set of stops. The system announces the route and destination of the vehicle every time the driver opens a door. It is intended to aide visually impaired passengers in identifying their desired vehicle. A geotagged timestamp is recorded whenever the external announcement system is initiated allowing for both the understanding of the set of stops made during each bus trip and the time at which these stops occurred.

Alighting Distribution of Final Stage Passengers on Inbound Route 39 Boarding Between 8:00 and 8:30

Destinations

Origins	South St @ Spalding St	South St @ Child St	South St @ Carolina Ave	South St @ Sedgwick St	Centre St @ Seaverns Ave	Centre St @ Saint John St	Centre St @ Robinwood Ave	Centre St @ Roseway	S Huntington Ave @ Perkins St	S Huntington Ave @ Bynner St	S Huntington Ave @ Heath St	S Huntington Ave @ VA Hospital	100 S Huntington Ave	S Huntington Ave @ Huntington Ave	Huntington Ave @ Parker Hill Ave	Huntington Ave opp Fenwood Rd	Huntington Ave @ Wigglesworth St	Huntington Ave @ Longwood Ave	Huntington Ave @ Ruggles St	360 Huntington Ave	Huntington Ave @ Gainsborough St	Belvidere St @ Huntington Ave	Dalton St @ Boylston St	Boylston St @ Dartmouth St	Boylston St @ Clarendon St	Back Bay Station
Forest Hills Station	0.32	0.05	0.00	0.06	0.06	0.02	0.04	0.01	0.04	0.00	0.00	0.02	0.01	0.02	0.04	0.08	0.03	0.09	0.04	0.02	0.00	0.02	0.00	0.00	0.00	0.03
South St @ Spalding St		0.08	0.00	0.02	0.02	0.02	0.03	0.01	0.02	0.00	0.00	0.01	0.01	0.03	0.06	0.12	0.06	0.23	0.09	0.06	0.00	0.06	0.00	0.01	0.00	0.04
South St @ Child St			0.00	0.04	0.01	0.01	0.03	0.00	0.01	0.00	0.00	0.01	0.01	0.02	0.07	0.16	0.04	0.19	0.11	0.06	0.00	0.08	0.00	0.03	0.00	0.11
South St @ Carolina Ave				0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.20	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.20
South St @ Sedgwick St					0.03	0.01	0.04	0.02	0.02	0.00	0.02	0.01	0.01	0.04	0.05	0.13	0.07	0.19	0.12	0.05	0.00	0.07	0.00	0.01	0.00	0.10
Centre St @ Seaverns Ave						0.01	0.01	0.00	0.02	0.00	0.00	0.01	0.00	0.04	0.04	0.17	0.09	0.19	0.12	0.05	0.01	0.07	0.00	0.04	0.00	0.12
Centre St @ Saint John St							0.01	0.00	0.01	0.00	0.01	0.02	0.00	0.02	0.06	0.20	0.05	0.19	0.09	0.04	0.01	0.11	0.01	0.03	0.00	0.12
Centre St @ Robinwood Ave								0.01	0.01	0.00	0.00	0.01	0.00	0.02	0.07	0.20	0.07	0.19	0.08	0.04	0.03	0.10	0.01	0.02	0.00	0.14
Centre St @ Roseway									0.00	0.00	0.00	0.01	0.01	0.03	0.05	0.17	0.10	0.08	0.10	0.03	0.05	0.14	0.02	0.01	0.00	0.21
S Huntington Ave @ Perkins St										0.00	0.00	0.01	0.00	0.02	0.05	0.15	0.04	0.17	0.09	0.04	0.01	0.22	0.00	0.03	0.00	0.16
S Huntington Ave @ Bynner St											0.01	0.00	0.00	0.05	0.04	0.12	0.06	0.13	0.07	0.17	0.05	0.06	0.00	0.04	0.05	0.16
S Huntington Ave @ VA Hospital												0.00	0.01	0.05	0.06	0.11	0.02	0.11	0.17	0.06	0.02	0.14	0.01	0.10	0.01	0.13
S Huntington Ave @ Heath St													0.00	0.01	0.04	0.10	0.01	0.07	0.21	0.02	0.14	0.10	0.06	0.02	0.02	0.22
100 S Huntington Ave														0.00	0.04	0.05	0.01	0.05	0.16	0.14	0.04	0.24	0.02	0.05	0.00	0.16
S Huntington Ave @ Huntington Ave															0.03	0.05	0.02	0.10	0.13	0.09	0.03	0.26	0.01	0.10	0.01	0.16
Huntington Ave @ Parker Hill Ave																0.07	0.01	0.04	0.16	0.13	0.08	0.23	0.02	0.06	0.01	0.19
Huntington Ave opp Fenwood Rd																	0.03	0.04	0.08	0.09	0.06	0.31	0.01	0.10	0.03	0.25
Huntington Ave @ Wigglesworth St																		0.08	0.04	0.22	0.00	0.14	0.04	0.08	0.02	0.38
Huntington Ave @ Longwood Ave																			0.07	0.07	0.05	0.36	0.06	0.16	0.01	0.22
Huntington Ave @ Ruggles St																				0.06	0.06	0.33	0.01	0.13	0.07	0.35
360 Huntington Ave																					0.02	0.15	0.05	0.22	0.10	0.46
Huntington Ave @ Gainsborough St																						0.33	0.11	0.33	0.00	0.22
Belvidere St @ Huntington Ave																							0.15	0.19	0.04	0.63
Dalton St @ Boylston St																								0.00	0.00	1.00
Boylston St @ Dartmouth St																									0.00	1.00
Boylston St @ Clarendon St																										1.00

(a) General Final Stage Passenger Alighting Distribution

Alighting Distribution of May 17, 2015 8:00AM Inbound Route 39 Trip

Destinations

Origins	South St @ Spalding St	South St @ Child St	South St @ Carolina Ave	South St @ Sedgwick St	Centre St @ Seaverns Ave	Centre St @ Saint John St	Centre St @ Robinwood Ave	Centre St @ Roseway	S Huntington Ave @ Perkins St	S Huntington Ave @ Bynner St	S Huntington Ave @ VA Hospital	S Huntington Ave @ Heath St	100 S Huntington Ave	S Huntington Ave @ Huntington Ave	Huntington Ave @ Parker Hill Ave	Huntington Ave opp Fenwood Rd	Huntington Ave @ Wigglesworth St	Huntington Ave @ Longwood Ave	Huntington Ave @ Ruggles St	360 Huntington Ave	Huntington Ave @ Gainsborough St	Belvidere St @ Huntington Ave	Dalton St @ Boylston St	Boylston St @ Dartmouth St	Boylston St @ Clarendon St	Back Bay Station
Forest Hills Station	0.00	0.10	0.00	0.00	0.11	0.00	0.08	0.00	0.08	0.00	0.00	0.00	0.01	0.04	0.08	0.15	0.00	0.17	0.08	0.00	0.00	0.03	0.00	0.01	0.00	0.06
South St @ Spalding St		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
South St @ Child St			0.00	0.00	0.01	0.00	0.04	0.00	0.01	0.00	0.00	0.00	0.01	0.03	0.09	0.19	0.00	0.23	0.13	0.00	0.00	0.10	0.00	0.03	0.00	0.13
South St @ Carolina Ave				0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
South St @ Sedgwick St					0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Centre St @ Seaverns Ave						0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.04	0.05	0.20	0.00	0.00	0.15	0.00	0.01	0.09	0.00	0.04	0.00	0.14
Centre St @ Saint John St							0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Centre St @ Robinwood Ave								0.00	0.01	0.00	0.00	0.00	0.00	0.02	0.08	0.24	0.00	0.22	0.09	0.00	0.03	0.11	0.01	0.02	0.00	0.16
Centre St @ Roseway									0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Huntington Ave @ Perkins St										0.00	0.00	0.00	0.00	0.03	0.05	0.16	0.00	0.18	0.10	0.00	0.01	0.24	0.01	0.04	0.00	0.18
S Huntington Ave @ Bynner St											0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Huntington Ave @ VA Hospital												0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
S Huntington Ave @ Heath St													0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100 S Huntington Ave														0.00	0.05	0.06	0.00	0.06	0.19	0.00	0.05	0.29	0.03	0.06	0.00	0.19
S Huntington Ave @ Huntington Ave															0.03	0.06	0.00	0.12	0.15	0.00	0.04	0.29	0.01	0.11	0.01	0.18
Huntington Ave @ Parker Hill Ave																0.08	0.00	0.05	0.19	0.00	0.09	0.27	0.02	0.07	0.01	0.22
Huntington Ave opp Fenwood Rd																	0.00	0.04	0.09	0.00	0.07	0.35	0.01	0.12	0.04	0.29
Huntington Ave @ Wigglesworth St																		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Huntington Ave @ Longwood Ave																			0.07	0.00	0.05	0.39	0.06	0.17	0.01	0.24
Huntington Ave @ Ruggles St																				0.00	0.06	0.35	0.01	0.14	0.08	0.37
360 Huntington Ave																					0.00	0.00	0.00	0.00	0.00	0.00
Huntington Ave @ Gainsborough St																						0.33	0.11	0.33	0.00	0.22
Belvidere St @ Huntington Ave																							0.15	0.19	0.04	0.63
Dalton St @ Boylston St																								0.00	0.00	1.00
Boylston St @ Dartmouth St																									0.00	1.00
Boylston St @ Clarendon St																										1.00

(b) Final Stage Passenger Alighting Distribution for the 8:00 AM Inbound Route 39 on May 17, 2015

Figure 3-2: General and Trip Specific Passenger Alighting Distribution for Inbound 39, 8:00-8:30.

This updating can change the alighting distribution of passengers on board a vehicle. Figure 3-2 is an example of the general seed matrix of Inbound Route 39 during the 8:00-8:30 AM time period compared to the seed matrix specific to the Route 39 trip that departed at 8:00 AM on May 17, 2015. This trip did not stop at ten of the bus stops along the route. The alighting probability for the remaining stops is then increased to account for the reduced number of alighting options for passengers.

3.2 Boarding Scaling

The destination scaling process calculates passenger flows for all passengers with an inferred origin. However, this is still undercounting reality. There are some bus passenger trips that are not recorded at all in ODX. For a passenger to be observed, they need to interact with either the farebox on the vehicle on which they board or a fare validator located at their origin stop. This does not always occur.

There are many reasons why a passenger might not be observed. Some passengers qualify for free rides. For the MBTA, this includes children under 11 and visually impaired individuals. These individuals are likely significantly undercounted. There are also occasions when the farebox on a vehicle might not be working properly, either not collecting any fares or only collecting certain fare media. This will leave some if not all passengers unobserved in the AFC system, and subsequently the inferred ODX database. There are instances when passengers might be waved onboard a vehicle by the operator without paying to increase operational efficiency. This occurs sometimes on very crowded vehicles when an operator might allow passengers to board at rear doors to take advantage of all available space or if a passenger has a pass that can be visually validated. Finally, some passengers may deliberately evade fare payment.

To correct for this undercounting of passengers, scaling factors were developed to bring ODX trip boarding totals closer to APC trip boarding totals. APC totals in this case are used as control totals. Passenger flow totals calculated during the destination scaling process are then divided by the predicted ODX/APC trip boarding total ratio, called farebox interaction rate, as shown in equation 3.4.

$$F'_{ODT} = \frac{F_{ODT}}{I_{RDM}} \tag{3.4}$$

Where

- F'_{ODT} is the boarding scaled passenger flow between origin stop O and destination stop D for trip T.
- F_{ODT} is the destination scaled passenger flow between origin stop O and destination stop D for trip T.
- I_{RDM} is the farebox interaction rate for route R in direction D during time period M.

3.2.1 Farebox Interaction Rate

The farebox interaction rate used in Equation 3.4 is the predicted percentage of passengers that are likely to interact with the farebox for any route/direction/time-period combination. To calculate this rate, trip boarding totals calculated with APC are compared to boarding totals calculated with ODX on bus trips for which both systems were operating.

Figure 3-3 shows comparisons of Inbound Route 23 trips departing between 6:00 and 9:00 AM from February through April 2015. While there are some outlier trips most have slightly fewer boardings recorded in ODX than with APC. Trips with few APC boardings and many ODX boardings are due to the APC system only operating correctly for a portion of a trip causing many boardings to be missed. Trips with few ODX boardings and many APC boardings likely have a low farebox interaction rate due to any of the reasons mentioned above.

The farebox interaction rate is calculated by running an ordinary least squares (OLS) regression on this set of trips. The y axis intercept is constrained to zero as we would expect a bus trip with zero boardings recorded in APC to also have zero boardings recorded in ODX. Outlying trips for which one system has double the number of boardings as the other were excluded from the regression as they were likely a result of equipment malfunction (most likely APC) and not indicative of the true farebox interaction rate. In Figure 3-3, a farebox interaction ratio of 0.914 was calculated with an R^2 predictive value of 0.7487. This farebox interaction ratio means that on average we would expect approximately 91% of passengers on any inbound Route 23 bus trip departing between 6-9 AM to interact with the farebox.

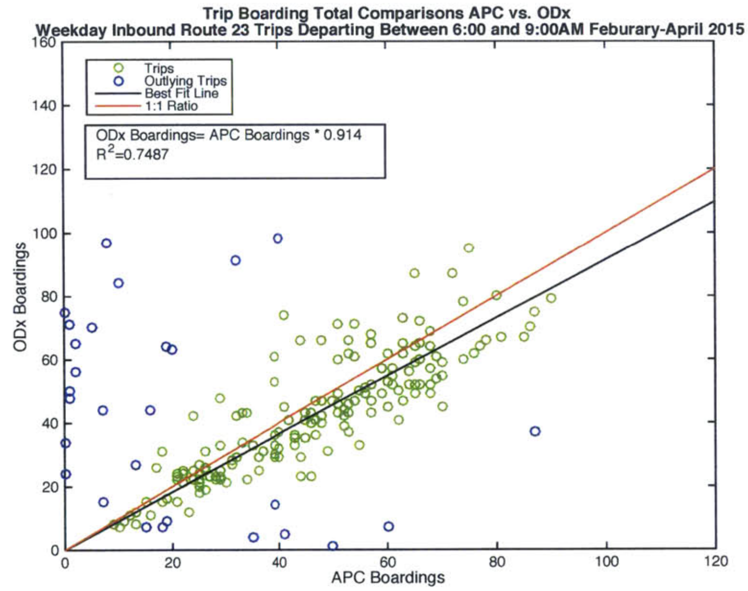


Figure 3-3: APC and ODX Trip Boarding Total Comparisons for Weekday Inbound Route 23 Trips Departing Between 6:00 and 9:00AM Februrary- April 2015

There is some variation between time periods. Figure 3-4 shows Inbound Route 23 between 9:00 AM and 12:00 PM. The farebox interaction ratio is over six percentage points lower than between 6:00-9:00 AM. This could be caused by differing rider demographics between periods. For example, it is possible that young children make up a higher percentage of riders in the 9:00 AM-12:00 PM period than during the AM peak. Usage rates of fare media might change as well. Passengers who travel during the AM peak might be more likely to be regular riders and have a CharlieCard, the smartcard fare media that includes a fare discount, while passengers traveling later in the morning might be less regular riders and pay with a CharlieTicket, a paper ticket fare media, or cash, for which an operator is more likely to wave on a passenger since the transaction time is significantly longer than with a CharlieCard. To account for these variations a farebox interaction rate would ideally be calculated for every three hour time period through out the day for each route/direction combination.

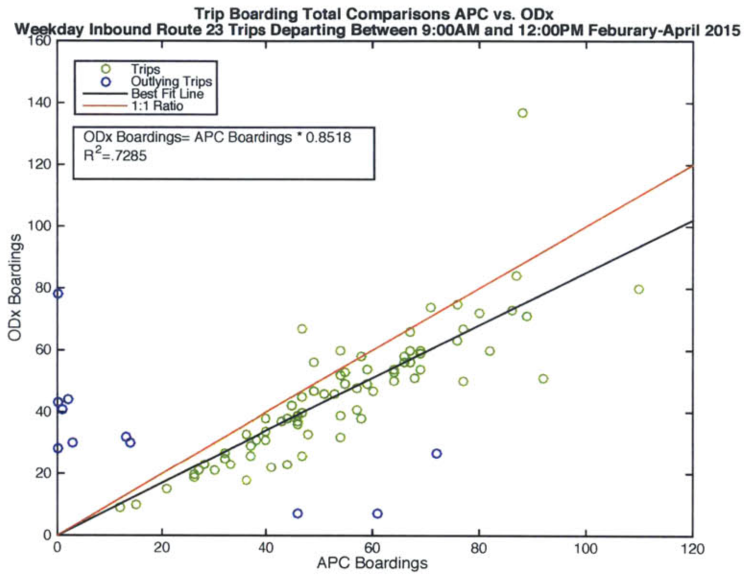


Figure 3-4: APC and ODX Trip Boarding Total Comparisons for Weekday Inbound Route 23 Trips Departing between 9:00AM and 12:00PM February-April 2015.

Systemwide, this period to period variation appears to have a moderate effect on farebox interaction rates. For most route direction combinations with at least two periods having over 50 sample trips, a threshold set to ensure accurate estimation, farebox interaction rates range 5-10 percentage points. Figure 3-5 shows the distribution of these ranges.

In the case of the MBTA bus network, there was an insufficient APC trip sample size to calculate factors for each three hour period on many low frequency routes. Only 38% of route, direction, and 3 hour time period combinations had over 50 sample trips. Thus, only one route direction factor, independent of time period, was calculated in this analysis.

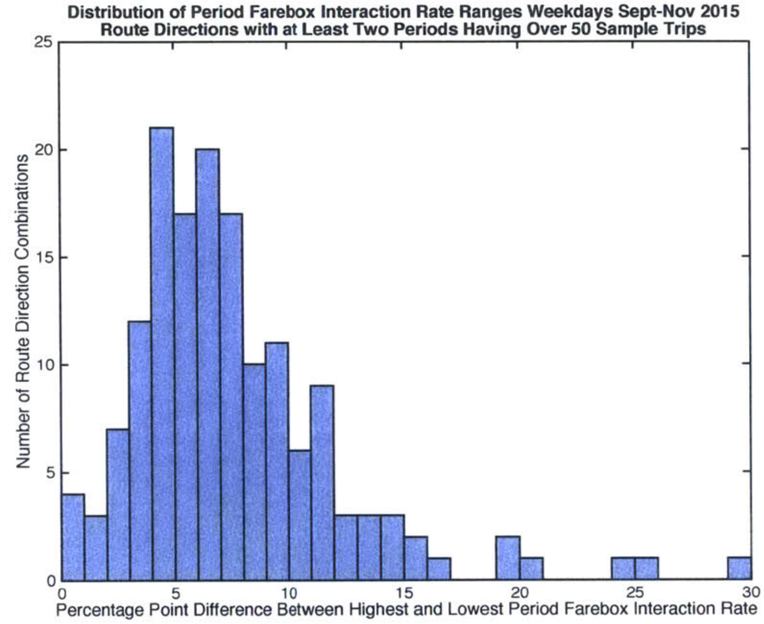


Figure 3-5: Range of Period Farebox Interaction Rates for Route Direction Combinations with at Least Two Periods Having Over 50 Sample Trips- Weekdays September-November 2015

3.2.2 Variation Among Routes

In addition to variation between different time periods there is also a large variation between routes. Farebox interaction rates for route-direction combinations range from as low as 0.5000 for Outbound SLW trips to 1.0224 for Inbound Route 431 trips with most combinations between 0.85 and 0.90. This variation can be attributed to the diversity of the routes in the MBTA network. Routes have different rider demographics, specific fare media usage rates, and fare evasion rates.

Many of the combinations with low interaction rates are routes on which fare payment is made as passengers enter a station instead of onboard the vehicle for a portion of stops on the route; therefore, these station-boarding passengers are not assigned to a specific vehicle trip. At the moment, it is not possible to accurately estimate loads on these routes. Others have a relatively small sample of trips on which to derive a factor, in which case relative outliers may have a larger effect.

Some routes have interaction routes over 1. Similar to many of the combinations with low interactions, these combinations also have small sample sizes on which the interaction rate is based.

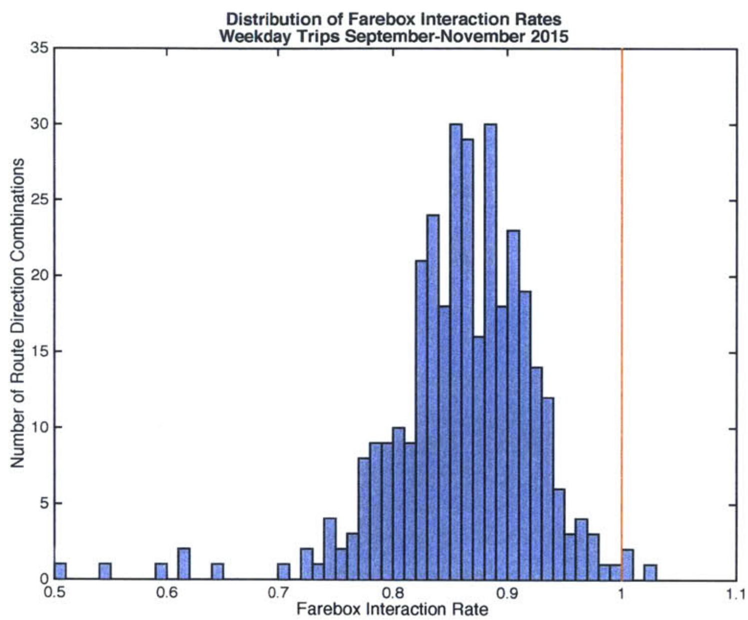


Figure 3-6: Farebox Interaction Rate Distribution for all Route Direction Combinations September-November 2015

3.2.3 Benefits of Calculating Disaggregate Farebox Interaction Rates

Farebox interaction rates alone can provide important information to transit agencies. In cases where AFC is used to calculate ridership totals, these rates can be used as a ridership adjustment factor to account for passengers not included in the initial ridership count or at least identify routes on which AFC totals might be significantly undercounting reality. This could help agencies provide more accurate ridership counts.

It could also help identify routes where fare collection rates are low. While a significant number of passengers who board without interacting with the farebox are likely doing so legally, routes with lower farebox interaction rates are likely collecting less fares than routes with higher farebox interaction rates. This means an agency is potentially missing out on a significant amount of fare revenue. Identifying routes where farebox interaction rates are low allows an agency to prioritize areas of opportunity to increase revenue.

Potential strategies could include promoting off-board fare payment through fare validators on routes where operational issues are causing low farebox interaction rates or increasing enforcement of fare payment on routes with high fare evasion through the use of transit police ride checks. An added advantage to promotion of offboard fare payment with validations is that all door boarding can be used to improve speed of boarding and support on-time performance.

3.2.4 APC Boarding Totals Validation

Validation tests were performed in order to ensure that APC boarding totals did in fact represent reality and were not systematically biased. This was accomplished by comparing boarding totals calculated manually to boarding totals calculated through APC. While there is some chance of measurement error even with manual counts, comparing counts between the two methods should help identify any systematic biases if they exist.

The MBTA had commissioned a set of ride checks to do just this. A total of 64 trips were checked from September 2010 through June 2013. Each ride check entailed having an employee manually count boardings and alightings at each stop on a bus trip for which an APC system was also in operation. Then loads and boarding totals were compared to validate the APC data. For the purpose of validating the boarding scaling process trip boarding totals were compared. Figure 3-6 shows this comparison.

An ordinary least squares regression, similar to the process used in the fare box interaction rate calculation, was run on the boarding totals for these trips. It showed that manual count totals were predicted to be roughly 95% of APC totals. This regression also had a high predictive value with an R^2 value of 0.904.

However, since the estimated model coefficient is close to one and there appears to be trips on either side of the 1:1 ratio line further investigation was required. A one sample t-test was performed to check whether the difference between the boarding totals of the two methods was statistically significant. For each trip, the APC

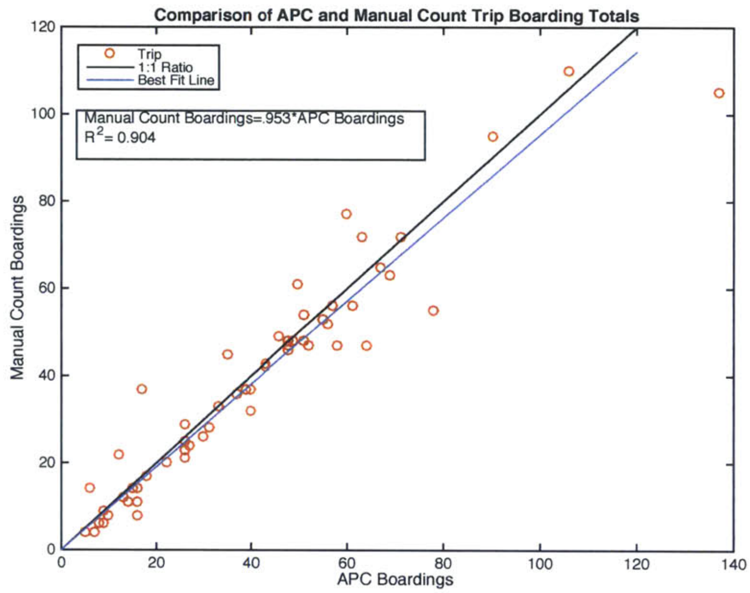


Figure 3-7: Manual Count-APC Trip Boarding Total Comparisons

boarding total was subtracted from the manual count boarding total. Then the t-test was performed on this set of differences to see if the mean of this population was statistically different from zero. The resulting p value from this test was 0.1887 meaning that we cannot reject the null hypothesis that the mean of the differences between APC and manual counts is zero at either the 5 or 10 % confidence level.

Therefore, while the OLS model predicted that manual count totals would be lower than APC totals, the predicted coefficient is not statistically significantly different from one. This suggests that there is no systematic bias in the APC calculation and that no adjustment factor is needed.

3.3 Validation

Validation was completed on the scaled load calculations to measure the accuracy of the estimates. APC load estimates were used as the ground truth during this process. Although these are estimates as well, comparisons to manual counts appear to show relatively small amounts of difference indicating these estimates are highly accurate. It also supplies a large enough sample to produce rigorous validation results.

Loads were estimated for every possible vehicle trip within the MBTA bus network for September-November 2015. Then, systemwide load comparisons were done to identify any systematic errors that might be occurring. Then average and individual trip load profiles were constructed to show any sources of error that may appear at a more disaggregate level. Finally, vehicle trip totals were calculated to determine the completeness of load calculations (i.e., for what percentage of trips operated are ODX based loads able to be calculated) and if biases occur.

3.3.1 Systemwide Analysis

First, load comparisons were made at every stop for trips on which loads were calculated through both ODX and APC processes. This includes over 2.6 million stops distributed among 81,897 individual weekday vehicle trips from September to November 2015. This was done on the original "raw" inferred ODX loads and after each successive scaling process to show its impact on load estimation. The first comparison includes loads calculated by aggregating passenger flows directly inferred through the ODX process. The second comparison adds destination scaling for the passengers without inferred destinations to obtain new aggregated flow totals. The final comparison includes an additional boarding scaling process in addition to the previous processes to obtain the "most comparable" loads and flows. The distribution of the differences between ODX inferred loads and APC calculated loads for the same trips are shown in Figure 3-8.

As mentioned during the scaling process methodology in the introduction to Chapter 3, raw ODX output undercounts passenger flows. This in turn causes load calculations to be underestimates as well. Load profiles constructed solely with ODX inferred passenger flows undercount loads at the vast majority of stops. After destination scaling, loads are still generally undercounted though the accuracy improves significantly. Finally, boarding scaling increases the calculated ODX load at every stop, reducing the amount of undercounting. This is shown in Figure 3-8 by a shift in the distribution of differences to the right, with results more centered around zero difference.

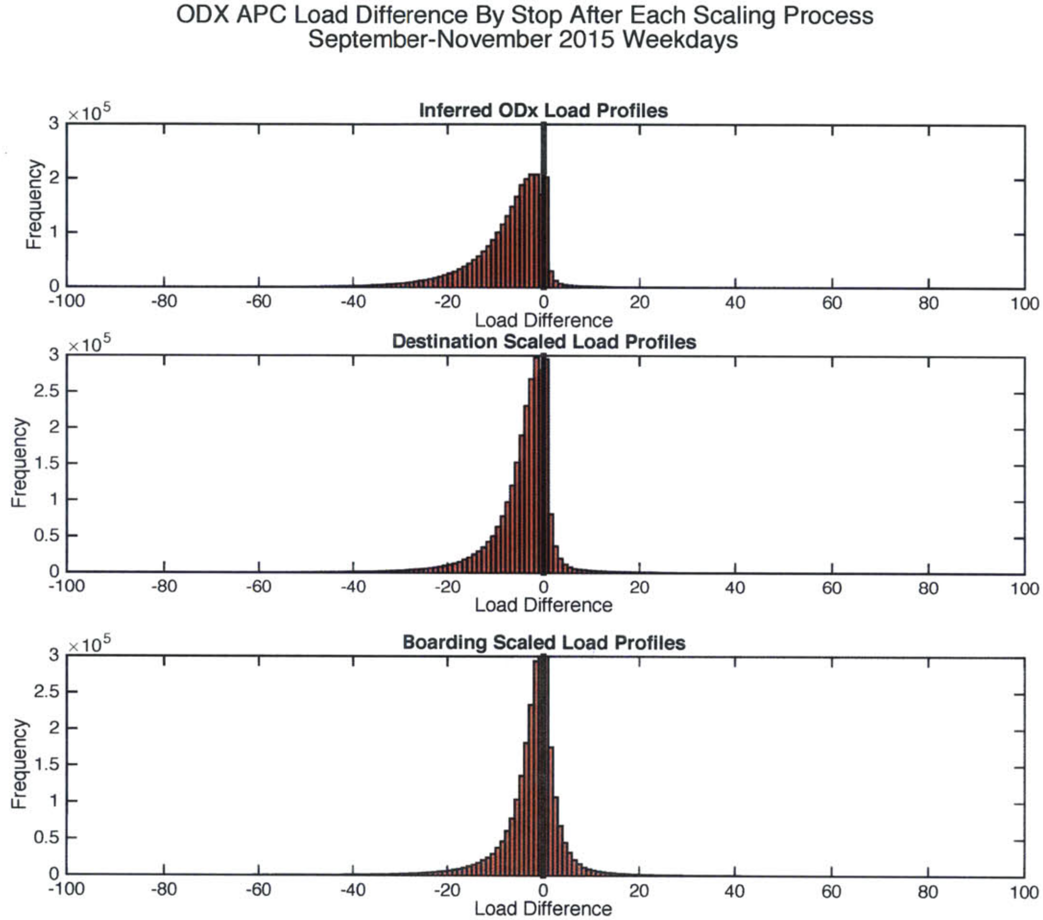


Figure 3-8: Distribution of Load Differences Following Each Stop (ODX-APC) for Trips with both APC and ODX Derived Load Calculations Weekdays September-November 2015

The distributions imply that there is some additional load variability inherently built into the ODX derived load profiles. This is likely due to the use of average seed and boarding scale factors. Farebox interaction rates and alighting distributions have a high probability of varying between trips in reality. The incomplete information available on trips without APC data require the use of average scaling factors independent of time period during the day.

Next, peak load estimates were compared. While load differences are important to note at all stops, it is especially important to note differences at the peak load point since this defines the highest crowding intensity experienced along an entire vehicle trip. Similar to Figure 3-8, peak load estimates were used from trips with both ODX and APC load estimates. A plot of these peak loads is shown in Figure 3-9. The higher on the color scale (yellow being the highest and dark blue the lowest) the more trips that have a given combination of peak load estimates (e.g., trips with an ODX estimated peak of 45 and an APC estimated peak load of 48). The red line represents equal peak loads.

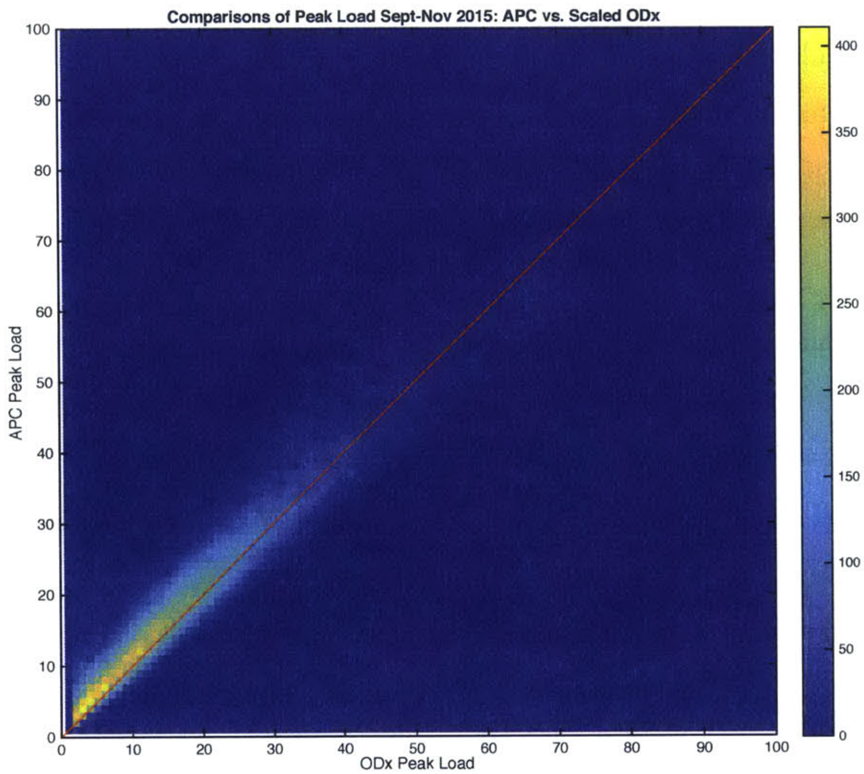


Figure 3-9: Comparisons of APC and ODX Estimated Peak Loads September-November 2015

A similar pattern to the load comparisons at every stop is revealed. The distribution appears to be relatively centered around the equal peak load line with a slight ODX undercount bias. There also appears to be some additional load variability introduced through the scaling process though most ODX peak loads appear to be within ±5 of the APC peak load.

3.3.2 Load Profile Analysis

Period average and individual trip load profiles of trips with both APC and ODX load estimates were also analyzed to determine the accuracy of the scaling process on a more disaggregate level. Average profiles help identify route or period level systematic errors while trip level could identify more sporadic and random errors that affect individual trip estimates. They also show the impact of each successive scaling process in load profile development.

Figure 3-10 shows the average load profile for weekday inbound Route 28 trips that are scheduled to depart during the AM Peak period (7:00-9:00AM) for September through November 2015. The scaling process works well in this instance. APC and boarding scaled ODX estimates never differ by more than three passengers at a given stop. There are some minor differences in the shape of the load profile but generally the two follow the same pattern.

It is also evident that the scaling process improves the accuracy of the load estimate. While both the inferred ODX and destination scaled load profiles provide an accurate representation of the general load profile shape, they underestimate loads. Each additional scaling process increases the load estimate bringing it closer to the APC load totals.

Results vary for across routes. The average load profile for weekday AM Peak inbound Route 1 trips is shown as an example in which the scaled ODX estimates are less accurate in Figure 3-11. Boarding scaled and APC loads are similar for the first half of the route but then differ significantly in the second segment of the route. There are two potential explanations:

1. It is possible that the seed matrix used during the destination scaling process is different than than the alighting distribution in reality. It appears that in the case of Route 1, ODX might be inferring a higher probability of passengers having relatively short trips while in reality they are much longer. This could potentially explain why ODX derived profiles have larger relative drops in load between 84 Massachusetts Avenue (MIT) and Massachusetts Avenue @

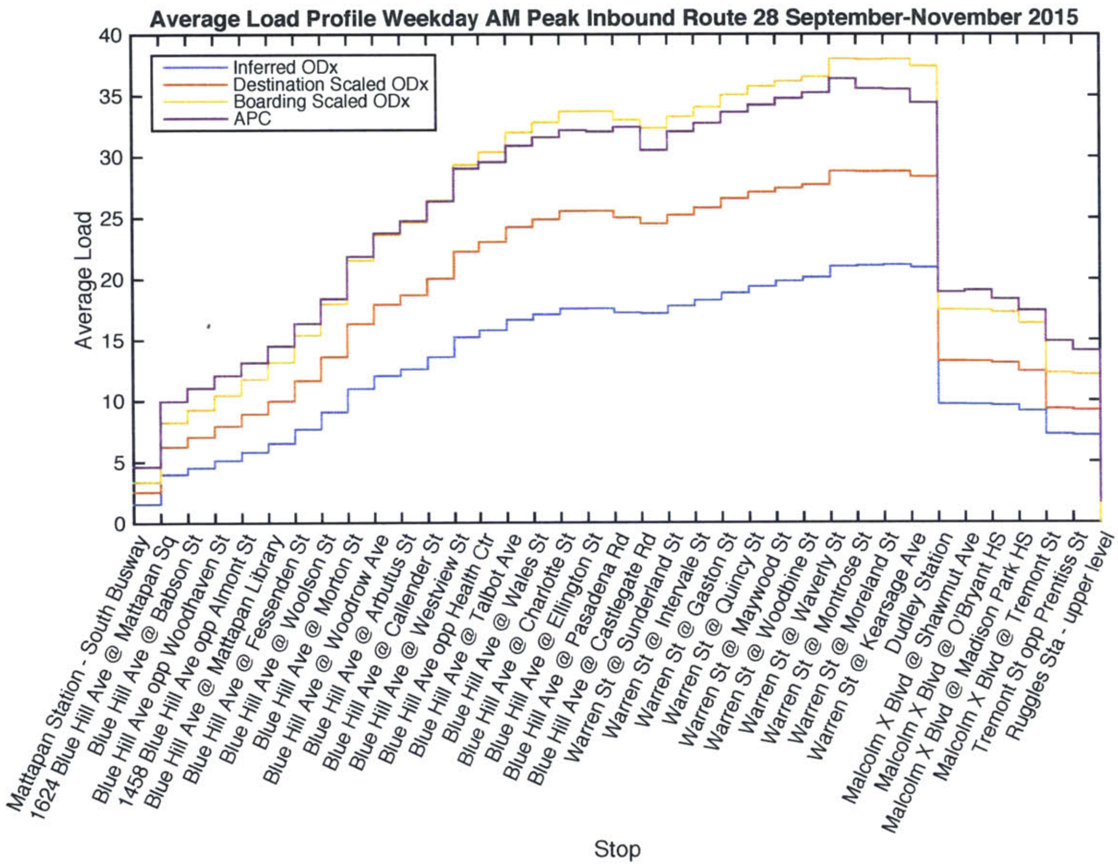


Figure 3-10: Average Load Profile for Weekday Inbound Route 28 During the AM Peak September-November 2015 Estimated with Each ODX Scaling Process and APC.

Massachusetts Avenue Station (a heavy rail stop). This is likely to be a bigger issue for routes with complex distributions of passenger OD's where there are multiple high volume stops compared to feeder or arterial routes where the vast majority of passengers alight at a final destination as shown in Figure 3-10 with Route 28.

2. It is also possible that the boarding scale factor used is incorrect. As mentioned in Section 3.2.1, routes that had large enough sample sizes to derive accurate factors for multiple periods saw moderate variation between periods. If a given period factor differs significantly from the overall factor boarding and load estimates could be inaccurate.

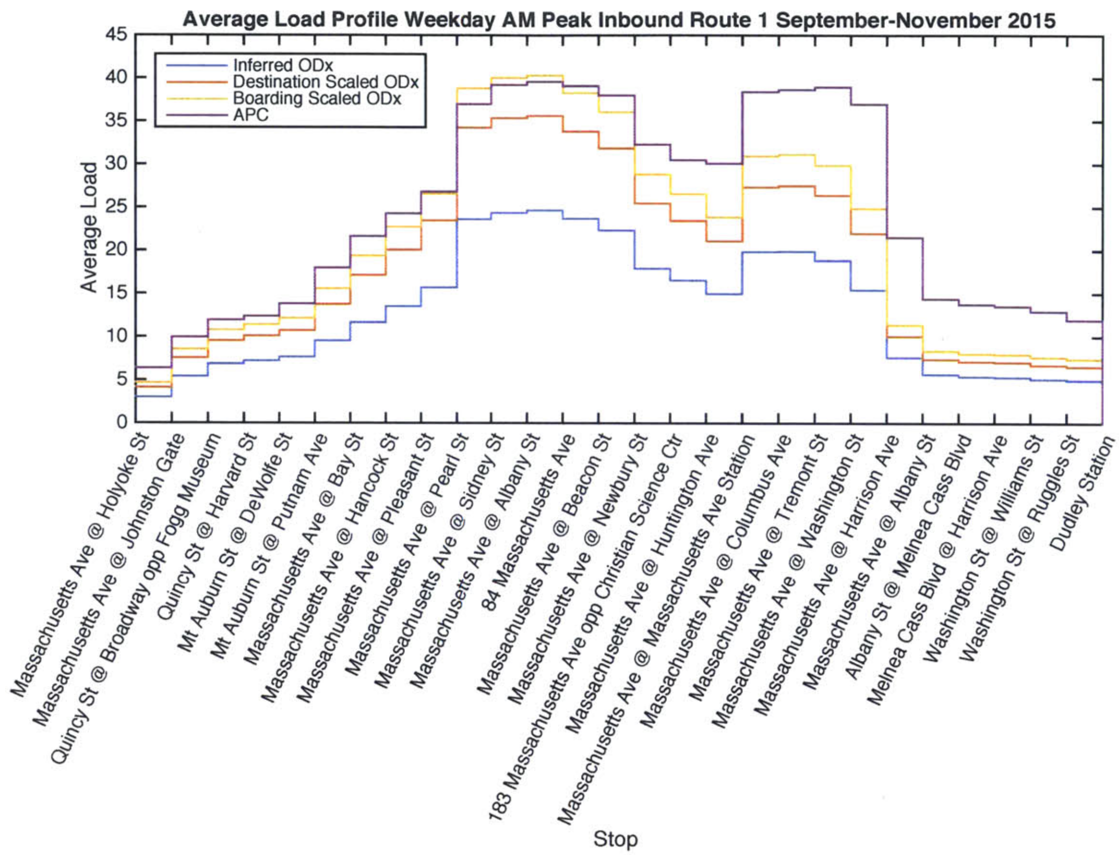


Figure 3-11: Average Load Profile for Weekday Inbound Route 1 During the AM Peak September-November 2015 Estimated with Each ODX Scaling Process and APC.

This difference appears to vary among routes. Some routes may have low interaction rates during the peak periods due to very high loads and relatively higher interaction rates during the offpeak when loads are lower. Other routes might have lower interaction rate during the offpeaks due to a high percentage of children passengers or passengers paying with a method other than smartcard media and higher during the peaks when more passengers are likely to be adults paying with smartcards.

Individual load profiles also help showcase the challenges faced when data collection collection systems are disabled for one reason or another. The 8:30AM scheduled Inbound Route 47 trip on September 21, 2015 provides an excellent example. It can be seen in Figure 3-12. On this trip, the external announcement system was not working properly, causing only the first and last stops to be recognized as potential boarding and alighting locations using the AVL data set. In reality, as can be in the APC profile multiple stops were made on this trip.

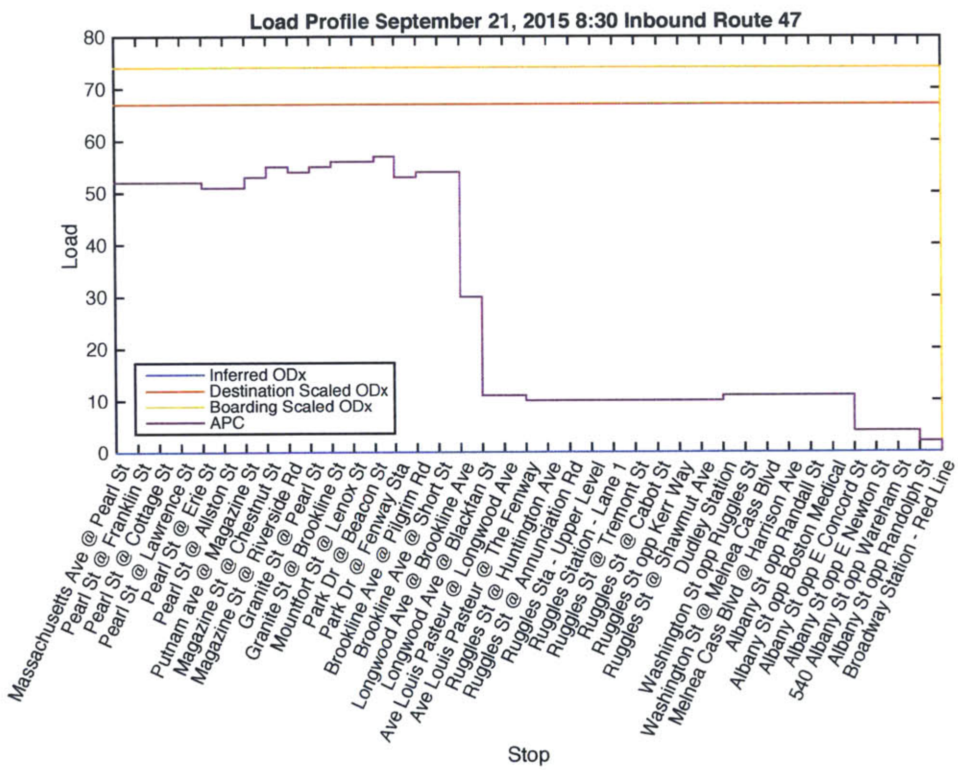


Figure 3-12: Load Profile of the September 21, 2015 8:30AM Inbound Route 47 Trip

Passengers were assigned to this trip through ODX though none had an inferred destination causing the Inferred OD Profile to estimate loads of zero for the entire vehicle trip . The scaling process is able to mitigate this equipment failure to some extent, although it doesn't provide a final resulting load profile that can realistically be used. Since passengers only have one possible boarding location and there is a time buffer around a stop's designated stop time for which passengers can be inferred to the given stop, passengers boarding both at the first stop and stops further downstream are assigned to the first stop. This is why loads are higher with destination scaled and boarding scaled ODX than with APC. These passengers are then estimated to stay on the bus for the entire duration of the trip as the only possible destination is the final stop causing the profile to remain flat. While the boarding scaled load profile shape is not accurate, the scaling process estimates a more accurate peak load than would be estimated otherwise. This equipment failure is relatively rare. Of vehicle trips with ODX derived loads, approximately 2% had external announcement failures.

3.3.3 Completeness

Theoretically this process should be able to estimate loads for every operated trip. However, due to data collection equipment failures and limitations in the ODX and scaling process methodologies, loads are unable to be calculated for all trips. Fare payment characteristics, and specific vehicle types of certain routes cause load calculations to be either inaccurate or unable to be reasonably computed.

On routes where passengers pay at a faregate instead onboard the vehicle, the ODX methodology is unable to assign passengers to specific vehicle trips. Therefore it is impossible to calculate loads for specific trips. In the MBTA case, this includes the Silver Line Waterfront Routes (SL1, SL2, SLW).

Certain routes do not require fare payment. These are often ad hoc shuttles or circulator routes. All passenger trips on these routes are unobserved since the ODX algorithm requires a fare transaction to identify a passenger trip. For the MBTA, this includes the Government Center Shuttle (Route 608), the SL1 Inbound from Logan Airport, and trips "run as directed" which are rail replacement shuttles for instances when portions of the rail network are unable to provide service.

Finally, some routes operate with specialty vehicles on which APC systems are not installed. This inhibits the development of a boarding scale factor. Therefore, loads can be estimated up through the destination scaling process but are likely undercounting by some unknown factor. The MBTA's Routes 71 and 73 falls into this category. They are operated during the weekday with electric catenary trolleybuses which are not equipped with APC systems and because it operates with specialty equipment, other APC fitted vehicles cannot be inserted into its schedule as they are for the vast majority of other routes.

Excluding the identified routes on which this methodology does not work well, 89% of September-November 2015 operated weekday vehicle trips had an ODX derived load estimate. To ensure that there is no bias in the distribution of trips without load estimates, three components were analyzed: day-to-day, route, and scheduled departure time.

Day-to-Day analysis, as seen in Figure 3-13, shows a relatively consistent percentage of operated trips systemwide for which an ODX derived load estimate is calculated. The calculation rate is lower for the first week of September. There was an issue with data collection this week which caused ODX destination inference rates to drop drastically from 60% to 40%. This likely also caused a significantly higher percentage of trips to not have loads calculated. For the rest of the Fall, calculation rates are consistently around 90% and do not show much bias between days of week or specific periods of the Fall.

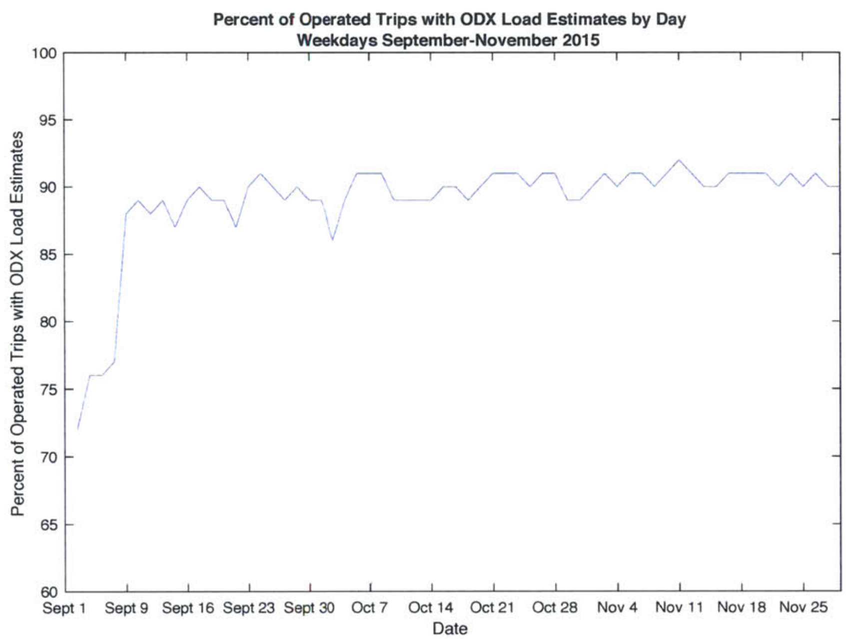


Figure 3-13: Percent of Operated Trips with ODX Load Estimates by Day Weekdays September-November 2015

Trips were then aggregated together by 30 minute scheduled departure periods to determine if there were any biases in calculation rates by time of day. Figure 3-14, shows that for much of the day calculation rates remain around 90%, However, during the early morning and overnight periods rates drop considerably. Approximately 50% of trips that were scheduled to depart during the 1 AM hour had ODX derived load calculations. There are two primary reasons for these lower rates:

1. There is a higher probability that a trip is operated without a single passenger boarding during the entire vehicle trip during the late night hours as there is less passenger travel demand. This would result in a trip not having a load calculation since the ODX algorithm requires fare transactions to infer passenger trips.

2. Since there are also fewer passenger trips overall during these periods, gaps in seed matrices become more likely. (i.e., having no alighting distribution for a given stop) This combined with a trip without any passenger destinations inferred through ODX leads to the trip not being observed.

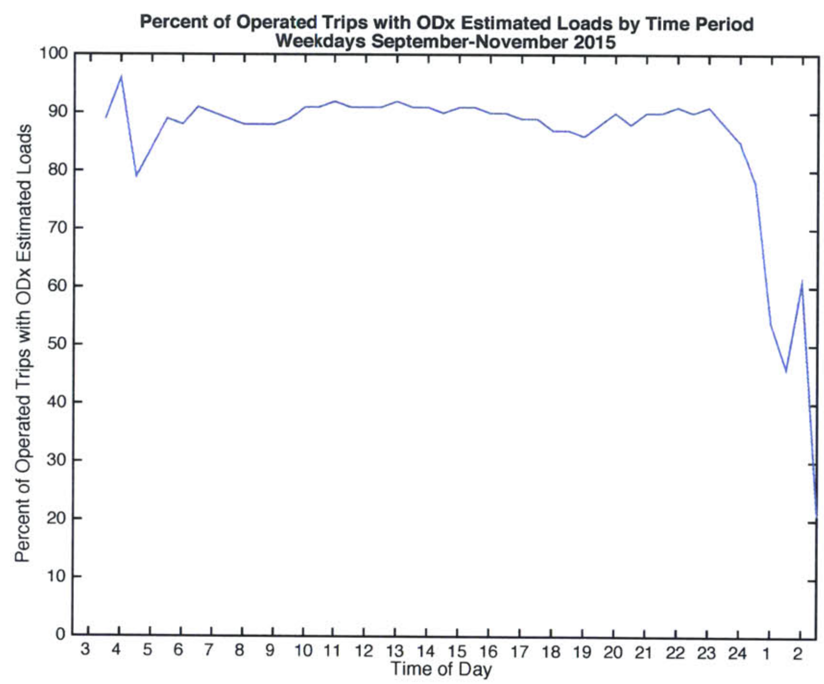


Figure 3-14: Percent of Operated Trips with ODX Load Estimate by Scheduled Departure Time Weekdays September-November 2015

Finally, the distribution of calculation rates among routes was explored to ensure that there was no bias among routes. As shown in Figure 3-15, most routes have similar rates around 95% of trips with the vast majority of routes having rates above

85%. There are a handful of routes with significantly lower rates. The vast majority of these routes are low frequency routes on which reasonable seed matrices cannot be computed . This is similar to the issue which is causing loads not to be able estimated on many late night trips.

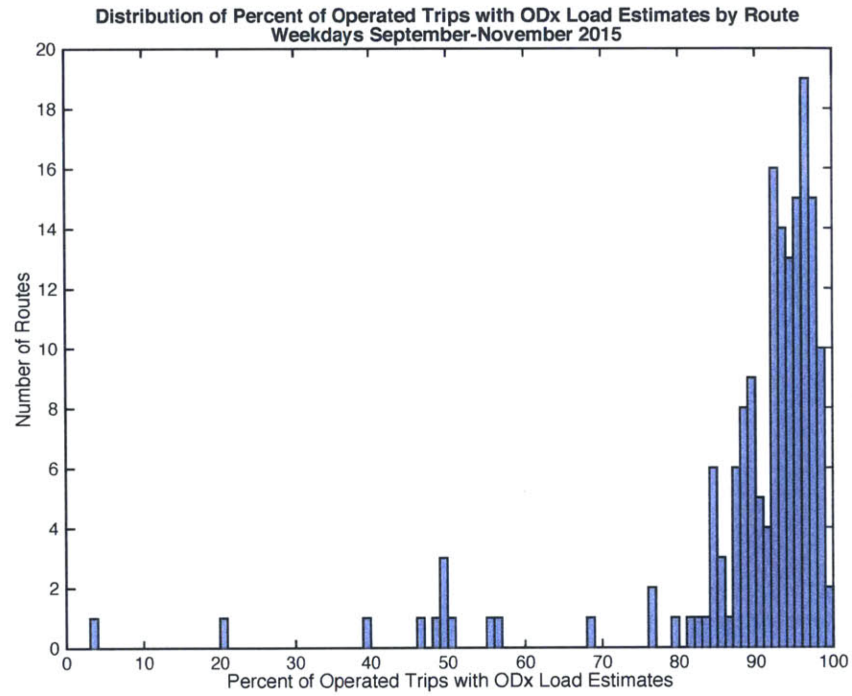


Figure 3-15: Distribution of Percent of Operated Trips with ODX Load Estimates by Route Weekdays September-November 2015

While there are certain routes and time periods for which calculation rates are low, on most routes and time periods, unobserved trips appear to be randomly distributed at constant rates. There is some bias but this occurs mostly on very low frequency and low demand routes.

3.4 Conclusions

In this chapter a process is discussed to scale inferred passenger origin destination information calculated through the ODX method. These scaled trip level passenger flows can be used to construct trip load profiles. It uses a seed matrix created through ODX inferred passenger flows and boarding scale factors derived from APC-ODX trip boarding total comparisons.

While gaps in data collection and methodology cause slightly less than complete coverage of all trips, this methodology allows for load estimates on a vast majority of vehicle trips with a relatively small sample of APC load estimates. Trips for which loads are unable to be calculated appear to generally be equally distributed among routes and time periods though there are some routes with low calculation rates and trips that operate during the overnight period also have lower estimation rates.

Chapter 4

Passenger Centric Bus Crowding Metrics

The process described in Chapter 3 allows for the development of a rich database of both trip level passenger flows and vehicle loads throughout a bus network. This enables crowding and comfort analysis that otherwise would not be possible without full implementation of APC systems in the bus fleet. Passenger centric crowding metrics can be developed to gain a better understanding of how passengers experience crowding throughout the system. These include passenger crowding duration, average standing time, number of unique passenger trips on which a passenger experienced crowded conditions, among others.

Some of these metrics can be selected to enable transit agencies to do comparative analysis to identify the highest priority routes and time periods for crowding mitigation. This helps to focus attention and resources on situations in which passengers experience the worst conditions which is especially important when there are not sufficient resources to meet the needs of the entire system.

Together with the MBTA staff and the Advisory Committee on the MBTA Service Delivery Policy, a passenger centric crowding standard was also developed using the detailed load information afforded by ODX. It identifies route/time period combinations in which a large percentage of passenger time is spent in uncomfortable conditions. This allows for conditions to be evaluated for each combination in isolation of the rest of the network.

4.1 Benefits of ODX Crowding Measurement

There are two primary ways in which ODX derived load and passenger flow estimates can improve the understanding of the passenger crowding experience throughout the bus network: nearly complete estimates without full implementation of APC and the ability to develop passenger centric metrics.

4.1.1 Complete Measurement

With a nearly complete coverage rate on calculable routes, the ODX scaling process allows for a much more granular load estimation compared to traditional sampling methods. In the MBTA case, an APC sample of 13% of operated trips on routes for which ODX loads can be estimated is expanded to almost 90% with the ODX scaling process. For resource constrained agencies, this could be powerful as nearly complete load and stop level running time coverage could be achieved without the installation of APC systems, a system solely designed for data collection, throughout the entire fleet. The inputs for the ODX inference algorithm come from more common ACDS that are required in the day-to-day operation of the network.

This nearly complete coverage allows for more nuanced measurement that accounts for variations in loads between adjacent trips and variations in running time among other considerations. Many of these aspects can be masked through the use of averages which become necessary when samples sizes are small. In addition, with the higher coverage rate cumulative measures can be used that highlight the effect that these variations have in the passenger experience. For example, the number of unique passenger trips on which crowded conditions are experienced, distribution of crowding duration, among others.

4.1.2 Passenger Centric Measurement

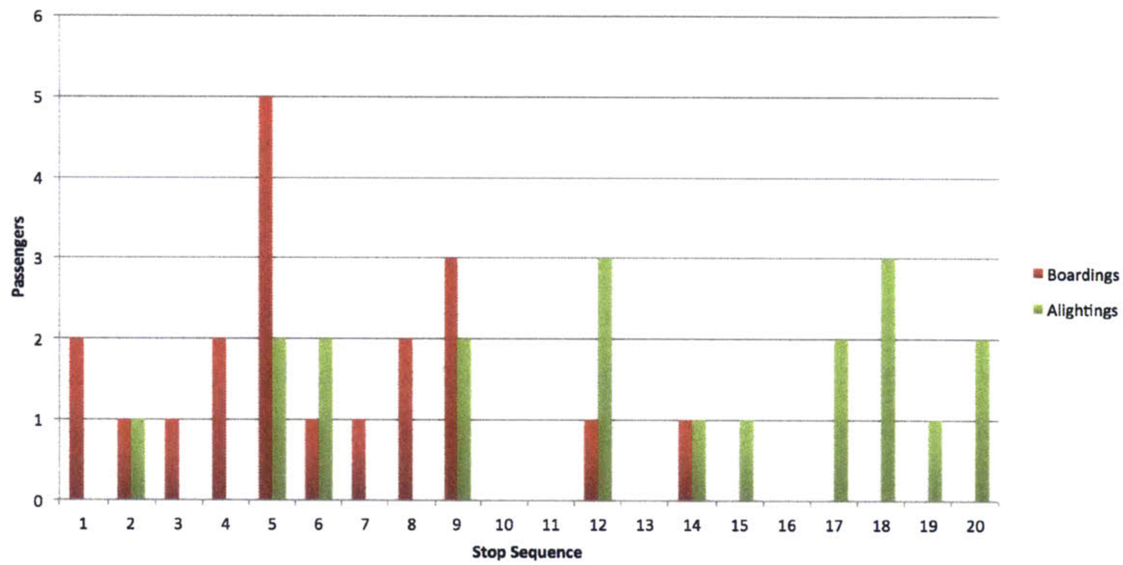
While many of the complete measurement aspects could be achieved with full implementation of APC systems in a bus fleet, ODX derived load profiles have the additional benefit of estimating passenger flows as well. Combining the two outputs enables estimation of the crowding experience for each passenger. As shown in Figure 4-1, traditional load profile development entails aggregating boarding and alighting totals at each stop and using these to construct loads. In this process, passenger OD's are not estimated. Passenger origins and destinations cannot be linked without further estimation through techniques such as iterative proportional fitting.

ODX derived load profiles, on the contrary, are developed with estimated passenger flows. Therefore, a trip on which every passenger had a destination inferred through ODX could be considered as a collection of passenger OD's, as in the bottom image in Figure 4-1. The scaling process makes this slightly more complex as passenger flows that are estimated during the destination scaling process are often fractionalized among possible alighting stops, meaning that a given passenger may not be assigned a unique origin-destination pair when estimating total flows on a route.

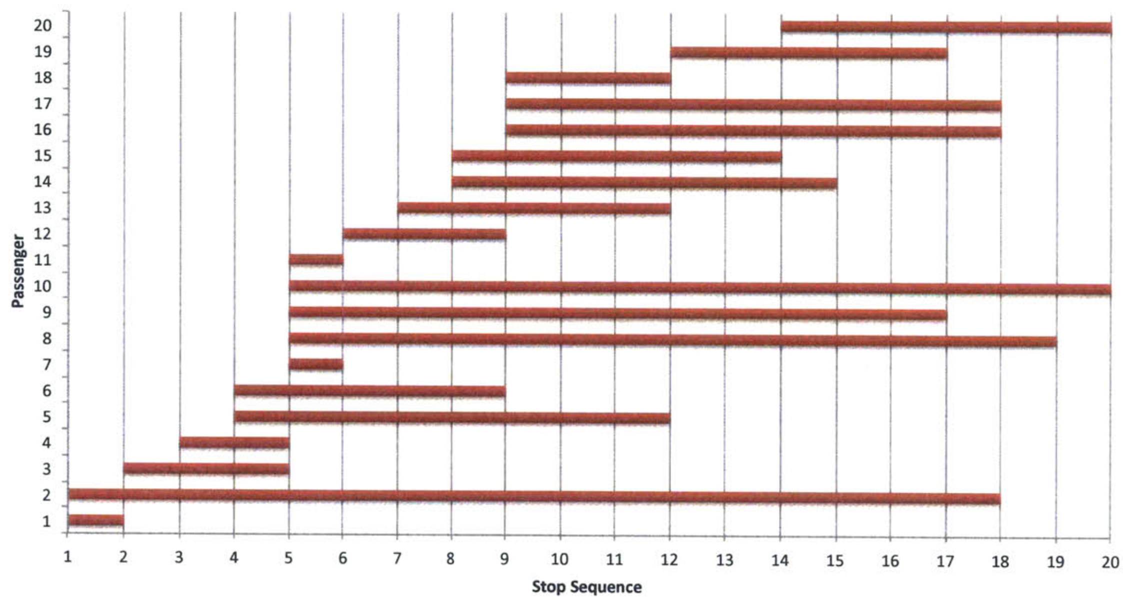
Combining passenger flow information with vehicle loads enables the estimation of each passenger's crowding experience. One can develop metrics that emphasize the passenger experience instead of operational measures such as vehicle loads. Examples of passenger centric metrics include: the unique number of passengers who face crowded conditions or the average length of time for which a passenger has to stand when loads exceed seated capacity on a given route.

4.2 Comparative Crowding Measurement

Many bus networks contain a diverse array of routes. Routes can vary in length, frequency of service, ridership patterns, travel speeds, degree of schedule adherence, etc. This can make comparisons of crowding conditions difficult. As discussed in Section 2.4, crowding conditions have traditionally been estimated through average



(a) Vehicle Trip as Boarding and Alighting Totals



(b) Vehicle Trip as a Collection of Passenger OD's

Figure 4-1: Example Vehicle Trip Showing APC (top) and ODX (bottom) Load Profile Development Process.

peak load analysis. While this is useful in determining whether a route is operating with high enough frequency in order to meet the level of travel demand along a route, it is difficult to compare average peak loads across routes. Individual passengers on two routes with similar average peak loads could face very different intensities and durations of crowding depending on the route characteristics mentioned above.

However, the combination of high trip coverage rates and the ability to develop passenger centric measures from ODX derived load and passenger flow estimates facilitate crowding comparisons between routes. There are four dimensions in which the use of these ODX estimates can ease comparisons compared to average load profile analysis, each are discussed in turn below:

1. Load variability's impact on passenger crowding
2. Duration of crowded situations
3. Number of unique passengers facing crowded conditions
4. Comparisons of routes with different frequencies

4.2.1 Load Variability

As Furth et.al. show and as discussed in Section 2.6, average load per vehicle trip calculations often mask the impact that load variability has on passenger crowding. Often when two buses are bunched, one is very crowded while the other is relatively empty. The average load of the two likely under-represents the crowding that passengers experience as many more ride in the very crowded vehicle. Since load and passenger flow estimates are available for nearly all operated trips, cumulative metrics, such as cumulative passenger crowding time, and the number of unique standing passengers, among others, can be used that consider the conditions experienced on each trip.

4.2.2 Crowding Duration

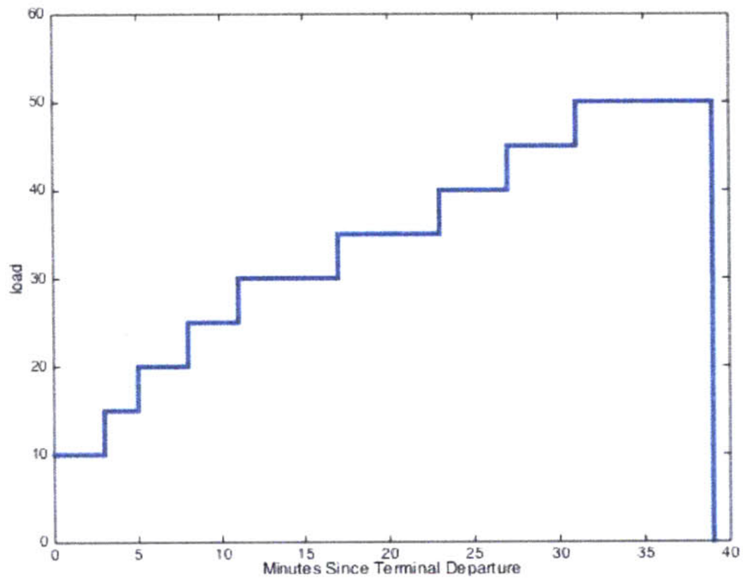
The amount of time that a passenger faces crowded conditions greatly affects his or her experience. Average peak load analysis only considers crowding intensity (i.e., vehicle load). It does not take into effect the duration of the peak or crowded portion of the route. This becomes especially important when comparing crowding conditions across routes that have similar peak loads but varying route lengths, run times, and ridership patterns.

Figure 4-2 shows two trips with similar peak loads but different peak load durations. The top is representative of typical inbound feeder route profile where loads increase throughout the trip as passengers board and do not alight until the final destination. The bottom is more typical of an express route where many people may board early in the route, then the vehicle may go a significant amount time without making a stop, sometimes traveling on a highway, with most people alighting at the final destination.

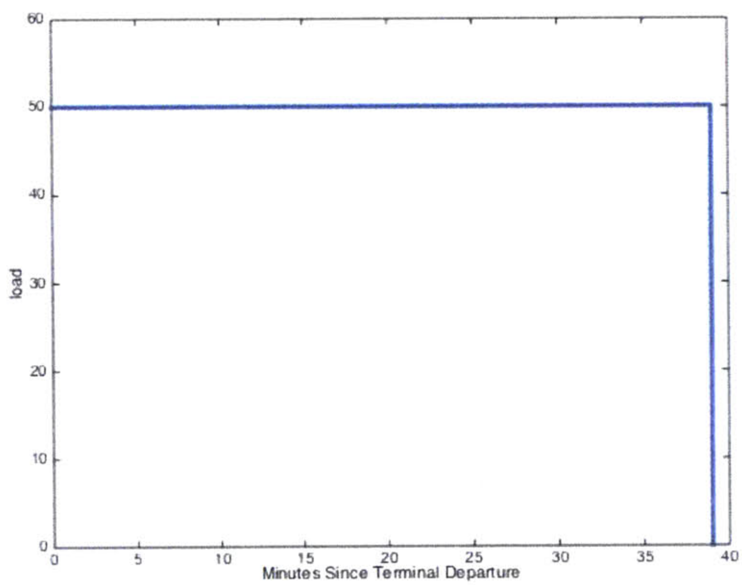
Both trips have a peak load of 50 passengers. However, many more passengers face conditions for a longer time on the bottom trip than the top. Passengers riding on the second (bottom) trip face peak load conditions for 38 minutes while passengers riding on the first trip only do for 8 minutes. Metrics that consider both crowding duration and intensity can be applied to better differentiate similar situations. ODX derived load profiles allow for this duration to be measured in two ways: time or distance.

4.2.3 Unique Passengers Facing Crowded Conditions

The turnover rate of passengers, i.e., the frequency that passengers board and alight during a trip, impacts the number of unique passengers who face crowded conditions. Given a constant load profile, the higher the turnover rate the more unique passengers who will face crowded conditions. An example is shown in Figure 4-3. Both trips have a load of 50 passengers for the entire 38 minute duration, however, the number of passengers experiencing this level of crowding differs significantly. In the top trip there is no passenger turnover. 50 passengers experience this level of crowding for the entire 38 minutes. In the bottom trip there is a significant amount of turnover. 130 passengers experience this level of crowding for an average duration of under 15 minutes. There is a difference between the number of passengers facing crowded conditions and the average duration for which these passengers spend in these conditions.



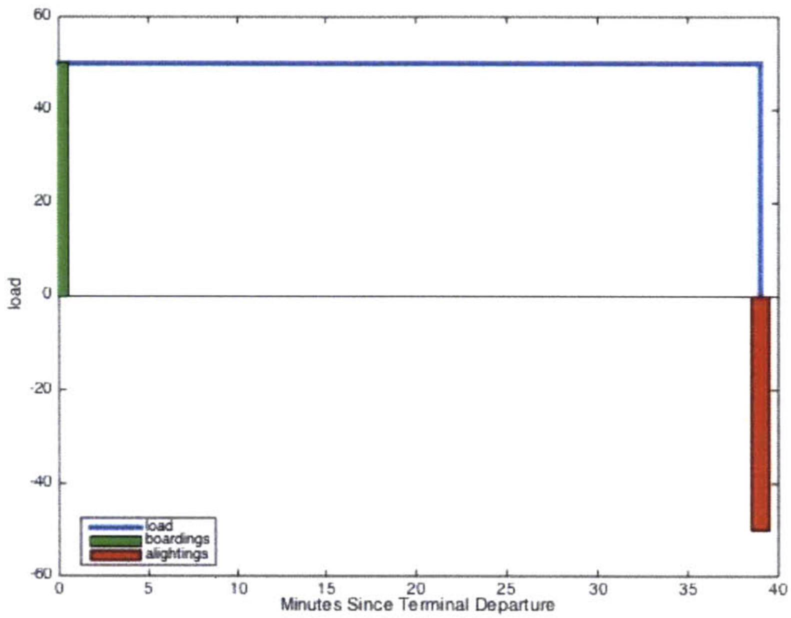
(a) Short Peak Load Duration



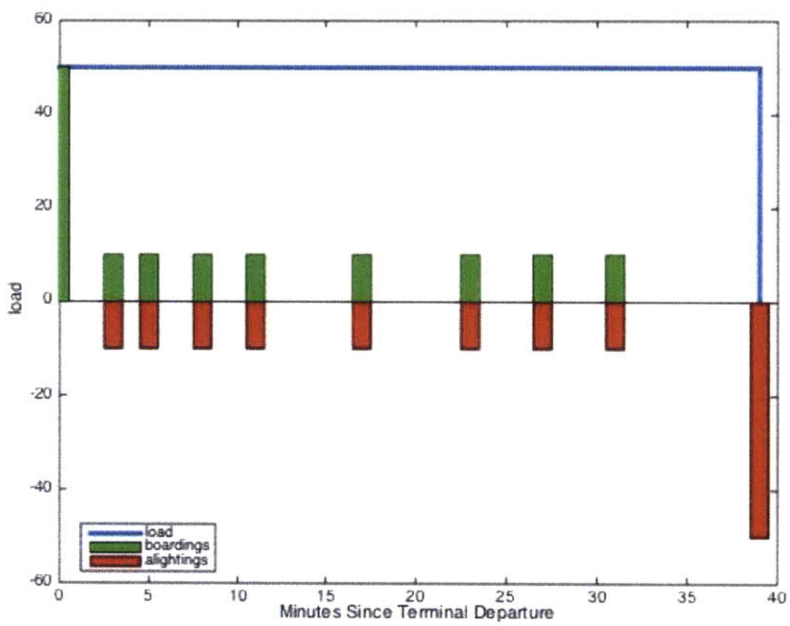
(b) Long Peak Load Duration

Figure 4-2: Comparison of Trips with Same Peak Load but Different Peak Load Durations

Turnover rate is not shown in peak load or even load profile analysis alone. The estimated passenger flows derived from the ODX based load profile development can be used to estimate the number of unique passengers who were on a vehicle during



(a) Trip with Low Turnover Rate



(b) Trip with High Turnover Rate

Figure 4-3: Comparison of Trips with Identical Load Profiles but Differing Turnover Rates any given section of a bus trip, identifying those who were onboard during crowded segments.

4.2.4 Comparisons Between Routes of Different Frequencies

Crowding can occur on routes of all frequencies. However, two routes with similar average load profiles and load variability but different frequency levels could have a significantly different number of passengers facing crowded conditions. The more frequent route will have have more passenger crowding since more bus trips will run over a given time period. For example, a route with an average of 10 minute headways will have 6 trips operate within a hour while a route with an average of 20 minute headways will only have 3 trips per hour. Since ODX based load estimates are available for almost every vehicle trip and crowding metrics can be passenger centric, frequency differences can be accounted for. Crowding metrics can be specified by the time period (e.g., crowding experienced on a given route for trips departing between 8:00-8:30AM). This recognizes that while crowding occurs on specific vehicles, a route operates as a system of multiple interrelated vehicles. Therefore, crowding analysis should be done on the route level over a selected period of time when possible.

4.3 Cumulative Passenger Crowding Time

Taking all of the issues discussed above into consideration, A recommended crowding metric, "Cumulative Passenger Crowding Time" (CPCT), was developed to utilize the additional information that ODX derived estimates provide. CPCT is the total amount of passenger time spent above a given crowding threshold. It is a cumulative metric which enables it to account for crowding caused by load variability by summing instead of averaging conditions on all trips under consideration. It also acknowledges the temporal effect of crowding by considering not only the intensity of crowding but also the duration. CPCT is passenger centric allowing for a more nuanced understanding of the passenger experience and enabling comparisons among routes.

Its formulation is described in Equation 4.1. For every trip segment, often between adjacent stops, the running time of that particular segment is multiplied by the number of passengers on board in excess of the crowding threshold. This is calculated for each segment of a vehicle trip and then summed to provide a cumulative passenger time above crowding threshold for that trip. This can be further aggregated to gain an understanding of crowding on a route or systemwide level.

$$CPCT = \sum_i R_i * max(0, L_i - T) \tag{4.1}$$

Where

- CPCT is the cumulative passenger time above crowding threshold for a given trip or set of trips
- i is a segment of a trip
- R_i is the running time for a given segment i
- L_i is the vehicle load over a given segment i
- T is the crowding threshold used in the calculation

An example application of this metric is shown in Figure 4-4. In this case, the crowding threshold is a load of 40 passengers, as marked by the green line. The red area is the passenger time for which loads exceed the threshold. For this trip, there are two instances of different duration and intensity in which loads exceeded the threshold. During the first instance, there are 5 passengers above the crowding threshold for 4 minutes for a total of 20 passenger minutes. Then during the second instance there are also 5 passengers above the crowding threshold for 10 minutes and then 10 passengers over the threshold for the final 5 minutes for total of 100 passenger minutes. Combined, there are two passenger hours of cumulative passenger time spent over the crowding threshold.

This metric estimates the intensity of crowding on each trip and weighs it by the duration. Therefore, a trip with intense crowding for a short duration could have

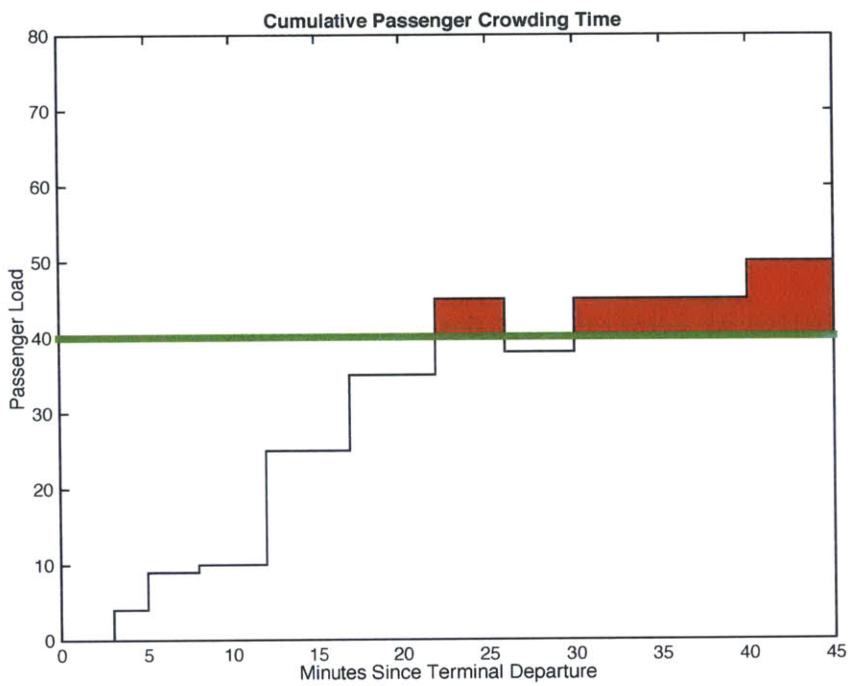


Figure 4-4: Example Calculation of Cumulative Passenger Crowding Time

the same CPCT as a trip with less intense crowding for a longer duration. While more passengers may have experienced intense crowding conditions on the first trip, the passengers experiencing crowding on the second may have experienced worse conditions since they needed to stand or remain in an uncomfortable position for a longer period of time.

The crowding threshold can vary depending on the circumstance. During relatively low demand periods, this could be the seated capacity of a vehicle while during high demand periods, this could be raised to some multiple of seated capacity like many agencies currently do for average peak load analysis, to identify situations with high intensity crowding.

4.4 Systemwide Crowding Analysis

CPCT of individual trips can be aggregated together in many different ways in order to do systemwide analysis. Although loads and therefore CPCT are not able to be calculated for every trip, since trips without calculations appear to be approximately uniformly distributed among routes and time periods CPCT can be used to make relative comparisons. (i.e., There is more crowding on Route A than Route B.) An analysis of crowding on the weekday MBTA bus network during the fall of 2015, September-November, was done in this manner.

4.4.1 Temporal Distribution of Crowding

A temporal analysis was done to show the distribution of crowding throughout a typical weekday. CPCT was calculated for each weekday trip for which an ODX derived load profile was created. This includes approximately 90% of trips operated on routes for which the scaling process is effective. See section 3.3.3 for a description of excluded routes and Appendix A for a complete list. The seated capacity of the vehicle in use on the specific route was used as the crowding threshold. Each trip was then grouped into 30-minute-time-period clusters depending on its scheduled departure time and CPCT was summed for each cluster. This distribution is shown in Figure 4-5.

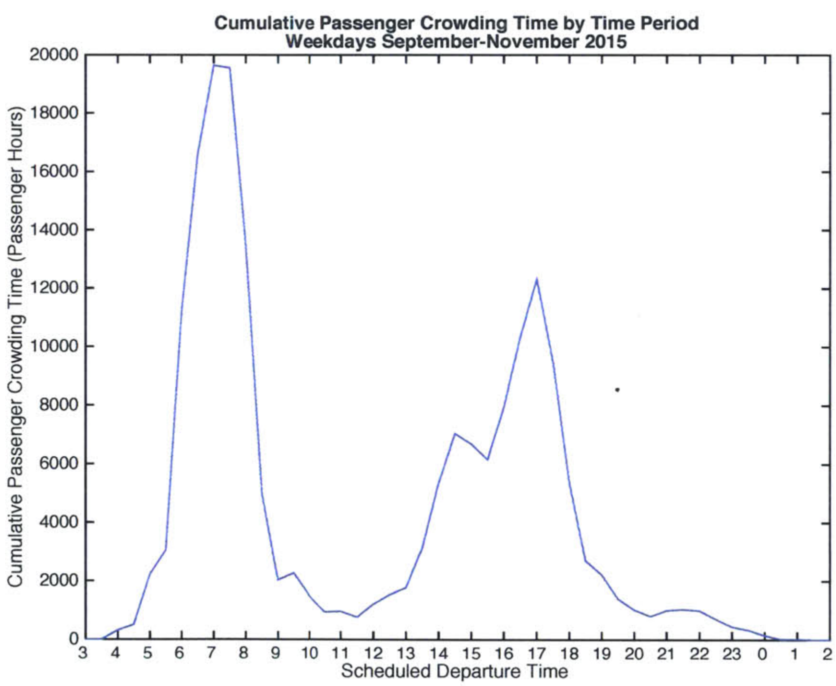


Figure 4-5: Cumulative Passenger Crowding Time by Scheduled Departure Time with a Crowding Threshold of Seating Capacity. Weekdays September-November 2015

It appears that most passenger crowding occurs during the peak periods. This occurs for three reasons:

1. In order to meet the needs of the entire bus network during the high demand peak periods the crowding standard that the MBTA uses during the peaks is relaxed to 140% of seating capacity (See Section 2.8.2 for a more detailed discussion) Therefore route schedules are built to allow for higher intensity crowding in the peak periods than during the off-peak periods.

2. Overall travel demand also peaks during these periods. Therefore, road congestion is generally higher as well. This causes running times to increase requiring passengers to remain in crowded conditions longer.

3. To meet this higher travel demand, frequencies are increased on most routes during the peak periods. Therefore, not only are average loads higher and running times longer but there are more trips operated resulting in more passengers experiencing these crowded conditions.

Crowding during the morning peak is more concentrated in trips that depart during the seven o'clock hour while crowding during the afternoon is less intense but occurs for a longer time frame. This appears to follow general travel patterns. Many of these passengers are likely to be commuters. In the morning, passengers are likely required to be at their place of work during the 8 o'clock hour. Since most passengers will likely head directly to their place of work, travel demand is concentrated within a narrow time frame. This peaking of travel demand is difficult for a transit agency to effectively manage without large amounts of resources sitting idle for the rest of the day. Therefore crowded conditions occur.

During the afternoon, passengers have more staggered times for which they depart for home. This is due to differing work hours and/or after work activities. Therefore, afternoon travel demand is much less concentrated than the morning and should be easier for the transit agency to effectively accommodate although the afternoon peak is often plagued by greater levels of traffic congestion that tend to increase the bus resources required to meet a similar demand.

Although Figure 4-5 appears to show that there is little crowding outside of the peak periods, it should not be interpreted as there are no crowded trips during the off peaks. There may be crowded trips, however frequencies are lower and travel times faster meaning fewer passengers are likely facing crowded conditions for shorter lengths of time. Further analysis of individual routes should be done to identify where a high percentage of passengers are facing uncomfortable conditions.

This metric is designed as a prioritization tool to identify the highest need routes and times in terms of passengers experiencing a given level of crowding weighed by their crowding duration. A crowding "standard" discussed in Section 4.7 that is designed to determine if a route is providing adequate levels of comfort to its passengers regardless of the number of customers served.

4.4.2 Distribution of Crowding by Route

Next, the distribution of CPCT among routes was explored. Trips were grouped by route and CPCT was summed. The cumulative distribution function of this distribution ordered by route ranking of CPCT is shown in Figure 4-6. Crowding thresholds of seated capacity and 140% of seating capacity were used.

The figure shows that much of the systemwide CPCT comes from a small portion of the bus network. Approximately 60% of seated capacity CPCT occurs in the top 20 most crowded routes, with almost 66% of 140% of seated capacity CPCT occurring in the top 20 routes. On many routes there are relatively small amounts of CPCT.

If route distribution is combined with temporal distribution a picture of even more crowding concentration is developed. Figure 4-7 shows this distribution for all route/30-minute-time-period combinations. 75% of seated capacity CPCT occurs in approximately 10% of route/time period combinations. 140% of seated capacity CPCT is even more concentrated with 75% occurring in roughly 5% of route-time period combinations. This is encouraging from an agency perspective as the passenger crowding experience systemwide can be improved with concentrated interventions.

Half hour time periods were used. This duration is short enough to gain a nuanced understanding of how crowding conditions may change through out the day while long

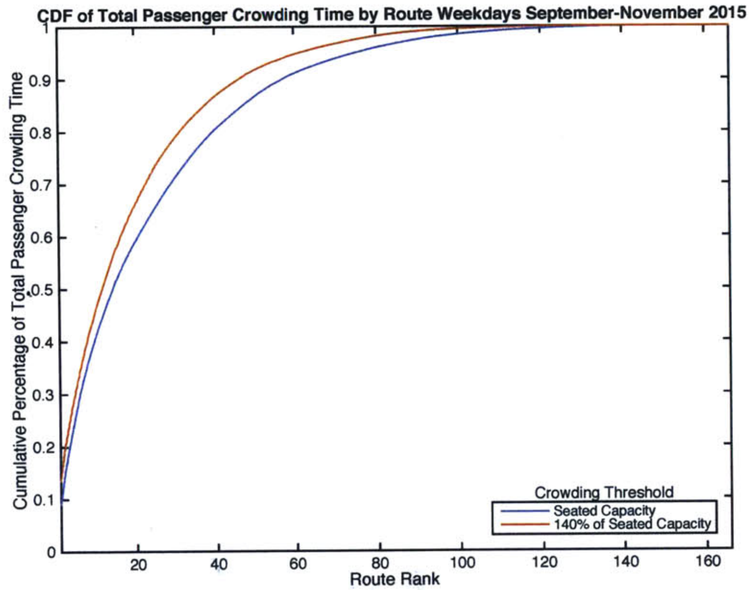


Figure 4-6: Cumulative Distribution Function of Cumulative Passenger Crowding Time Ordered by Route Ranking of CPCT. Weekdays September-November 2015

enough to ensure that most routes have at least one scheduled trip in each period within their span of service. Some routes which only operate during the peak periods or other select times may only operate in a few time periods while other routes that operate for the entire MBTA span of service may operate trips in as many as 46 different periods during an average weekday.

4.5 Prioritization

Comparative metrics also enable a ranking and prioritization of routes and time periods. It makes possible the identification of areas where the worst crowding occurs. This can be done for entire routes to determine where broad changes might be necessary such as changes in route design or transit priority measures. It can also be done on a route/30-minute-time-period level to identify where more narrow time specific changes might be warranted such as increased scheduled frequency, temporary bus lanes, or more intense dispatch control at terminals.

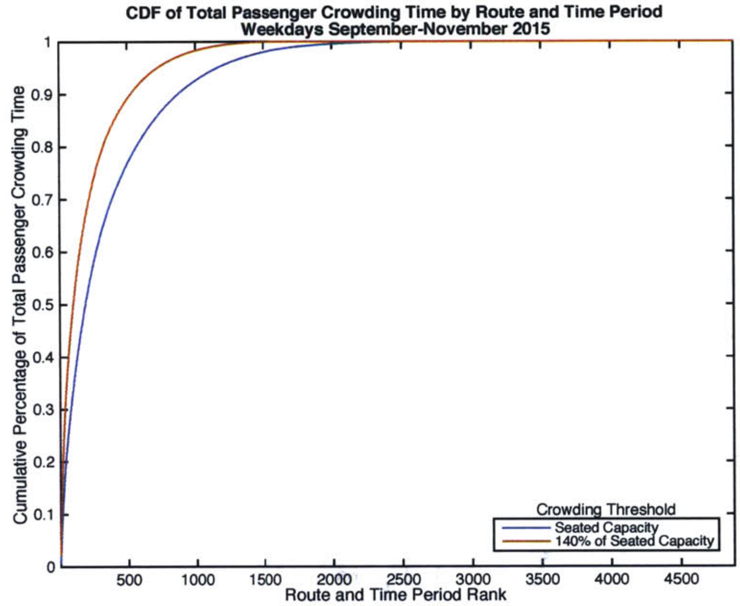


Figure 4-7: Cumulative Distribution Function of Cumulative Passenger Crowding Time Order Route and 30-Minute-Time-Period of CPCT. Weekdays September-November 2015

4.5.1 Route CPCT

Rankings were first completed by route. As shown in Section 4.4.2 most CPCT occurs in a relatively small selection of bus routes within the MBTA bus network. A list of the 20 routes with the most CPCT using seated capacity as the crowding threshold is shown in Table 4.1.

Rank	Route	CPCT (Passenger Hours)	Rank	Route	CPCT (Passenger Hours)
1	111	17329	11	7	4615
2	66	10579	12	77	4563
3	57	8265	13	39	4411
4	23	7351	14	47	4026
5	1	7319	15	16	3930
6	28	6951	16	22	3744
7	9	6696	17	34E	3251
8	32	5812	18	65	3191
9	86	5276	19	57A	2783
10	70	5039	20	93	2560

Table 4.1: MBTA Bus Routes with Most CPCT with Crowding Threshold of Seated Capacity Weekdays September-November 2015. Key Routes are Listed in Orange.

50% of the routes listed are considered key routes (in orange) under the MBTA's Service Delivery Policy. These routes meet higher frequency and span of service standards than the rest of the bus network. They are in turn also many of highest ridership routes in the network.

The vast majority are also high frequency urban routes. The only exception is the the 34E which is long local feeder route that runs from suburbs southwest of Boston to Forest Hills, the terminal station of one of the heavy rail lines. Although it only operates at 20 minute headways during the morning peak period, it has a scheduled run time of approximately an hour meaning that while fewer passengers likely experience crowded conditions, when they do they are likely to experience them for a significant amount of time.

Passengers on the Route 111 experience significantly more crowding than passengers on any other route. This is a high demand route operating with 4 minute headways during the morning peak. The crowding that occurs also occurs for a long duration. The 111 runs from Chelsea, an inner suburb, across Boston Harbor to Haymarket Station in Downtown Boston. Along the way it crosses the Tobin Bridge, one of the main connections to Boston from the North Shore suburbs. There are no stops between Chelsea and Boston meaning that when crowded conditions arise, passengers remain in these conditions for extended periods of time, especially when travel speeds over the Tobin Bridge slow due to congestion.

When a crowding threshold of 140% of seated capacity is used to isolate only high intensity crowding situations, as shown in Table 4.2, most of the routes from the seated capacity CPCT list remain. Only two are replaced: 22 and 34E which move to positions just outside of the top 20. The order of routes changes somewhat with non-key routes rising in rank (notably 7, 9, and 47). These appear to be routes with high peak period but low off-peak travel demand meaning that they are likely to have high intensity crowding occurring during the peak periods when schedules are built around 140% of seated capacity peak loads and less low intensity crowding during the off-peaks.

The 111 continues to have significantly higher CPCT than any other route. It has

Rank	Route	CPCT (Passenger Hours)	Rank	Route	CPCT (Passenger Hours)
1	111	4062	11	23	763
2	1	1649	12	32	727
3	9	1629	13	39	727
4	7	1329	14	16	656
5	66	1295	15	70	612
6	57	1217	16	65	512
7	86	1039	17	57A	498
8	47	974	18	117	456
9	28	893	19	109	456
10	77	819	20	93	440

Table 4.2: MBTA Bus Routes with Most CPCT with Crowding Threshold of 140% of Seated Capacity Weekdays September-November 2015. Key Routes are Listed in Orange.

almost 2.5 times more CPCT than the second highest route.

4.5.2 Route/Time-Period CPCT

In addition to understanding on which routes the most passenger crowding occurs, it is also important to understand when it occurs as this will inform schedule adjustments. As done in creating Figure 4-7, trips were grouped into 30 minute clusters based on scheduled departure time and route. Table 4.3 shows the route/time-periods with the most CPCT with seated capacity as the crowding threshold.

The spike in crowding during the peak periods shown in Figure 4-5 is highlighted in this table as well. All but 5 of the top 50 route/time-periods combinations occur during the peak periods. There are also more morning periods at the top of the list while afternoon periods are further down. Route 47 saw a significant amount of crowding on trips scheduled to depart between 16:30-17:00 which is likely due to high levels of congestion over the BU Bridge.

It is evident that for a number of the most crowded routes passengers are likely to experience crowded conditions if they board anytime within a large time frame. For example, Route 111 has four of the top twenty route periods and nine of the top fifty. They include four consecutive periods (6:00-8:00). Therefore a passenger boarding during these periods is likely to experience a significant amount of crowding. This

Rank	Route	Period	CPCT (Passenger Hours)
1	9	7:30-8:00	2002
2	111	6:30-7:00	1478
3	111	7:00-7:30	1471
4	9	8:00-8:30	1387
5	28	6:30-7:00	1345
6	23	6:30-7:00	1298
7	66	7:30-8:00	1247
8	111	6:00-6:30	1234
9	66	7:00-7:30	1220
10	7	8:00-8:30	1210
11	7	7:30-8:00	1146
12	28	6:00-6:30	1077
13	70	7:00-7:30	1036
14	23	6:00-6:30	980
15	1	9:30-10:00	970
16	57	7:30-8:00	968
17	111	7:30-8:00	941
18	86	7:30-8:00	928
19	47	16:30-17:00	890
20	86	8:00-8:30	875
21	65	7:30-8:00	857
22	86	7:00-7:30	844
23	111	14:00-14:30	841
24	9	7:00-7:30	834
25	1	16:30-17:00	833
26	111	14:30-15:00	796
27	16	6:30-7:00	787
28	1	17:00-17:30	781
29	57	7:00-7:30	754
30	66	16:30-17:00	750
31	66	6:30-7:00	725
32	77	7:30-8:00	724
33	32	7:00-7:30	702
34	66	17:00-17:30	667
35	1	8:00-8:30	664
36	111	5:00-5:30	659
37	77	7:00-7:30	656
38	16	6:00-6:30	642
39	47	16:00-16:30	619
40	28	15:00-15:30	613
41	32	6:30-7:00	611
42	111	15:00-15:30	592
43	86	16:30-17:00	578
44	57	8:00-8:30	574
45	66	17:30-18:00	562
46	9	17:00-17:30	554
47	66	14:00-14:30	551
48	23	14:30-15:00	545
49	66	8:00-8:30	543
50	111	17:00-17:30	536

Table 4.3: MBTA Bus Route/Time-Periods with Most CPCT with Crowding Threshold of Seated Capacity Weekdays September-November 2015. Key Routes are Listed in Orange.

is likely the cause of Route 111 having significantly more crowding than any other routes though many other routes have multiple periods listed as well.

Table 4.4 presents 140% of seated capacity CPCT by route/time-period. It appears to show similar patterns as seated capacity. The vast majority of route/time-period combinations occur during the peak periods and all occur on urban routes. As in Table 4-2, non-key routes also rise in rankings for CPCT of 140% seated capacity with significantly more non-key routes in the top 50 compared to seated capacity CPCT.

It now possible to compare temporal distributions of crowding among routes. For the use of many resources there is a temporal component (i.e., when is it being used) as well as spatial (i.e., where it is being used). Therefore, a resource could be used in one location for a period of time and then another for a different period of time.

Rank	Route	Period	CPCT (Passenger Hours)	Rank	Route	Period	CPCT (Passenger Hours)
1	9	7:30-8:00	698	26	64	7:00-7:30	176
2	1	9:30-10:00	686	27	23	6:30-7:00	175
3	111	7:00-7:30	455	28	1	16:30-17:00	174
4	9	8:00-8:30	431	29	28	6:30-7:00	172
5	111	6:00-6:30	406	30	65	7:30-8:00	166
6	7	8:00-8:30	404	31	47	16:00-16:30	164
7	7	7:30-8:00	392	32	16	6:30-7:00	158
8	111	6:30-7:00	381	33	11	8:00-8:30	148
9	47	16:30-15:00	322	34	116	6:30-7:00	145
10	86	7:30-8:00	249	35	77	7:00-7:30	145
11	66	7:30-8:00	249	36	57	7:00-7:30	144
12	111	14:00-14:30	242	37	1	8:00-8:30	142
13	86	7:00-7:30	235	38	7	8:30-9:00	141
14	28	6:00-6:30	221	39	47	8:30-9:00	141
15	66	7:00-7:30	211	40	1	17:00-17:30	134
16	111	14:30-15:00	210	41	111	21:30-22:00	132
17	70	7:00-7:30	205	42	111	10:00-10:30	130
18	9	7:00-7:30	204	43	111	15:00-15:30	119
19	23	6:00-6:30	202	44	28	15:00-15:30	119
20	57	7:30-8:00	200	45	86	16:30-17:00	118
21	111	7:30-8:00	195	46	39	16:30-17:00	116
22	77	7:30-8:00	193	47	64	8:00-8:30	114
23	16	6:00-6:30	193	48	19	6:00-6:30	112
24	86	8:00-8:30	190	49	11	7:30-8:00	110
25	111	5:00-5:30	186	50	57A	8:00-8:30	110

Table 4.4: MBTA Bus Route/Time-Periods with Most CPCT with Crowding Threshold 140% of Seated Capacity Weekdays September- November 2015. Key Routes are Listed in Orange.

Figure 4-8 shows how this comparison can be done between routes with an example comparing the temporal distribution of 140% of seated capacity CPCT on Routes 7 and 111.

The figure shows that the temporal distribution of CPCT can vary between routes. Although both routes have morning and afternoon peaks they occur at different times. The Route 111 peaks generally occur earlier than Route 7. This may be related to different running times (Route 111 scheduled run times are approximately three times longer than Route 7) as passengers may need to board earlier trips to make it to their final destination on time on the Route 111.

It may also be that passengers of Route 111 have different travel schedules than those of Route 7. The 111 connects working class neighborhoods in Chelsea, an inner suburb, to downtown Boston. A larger percentage of its passengers may need to be

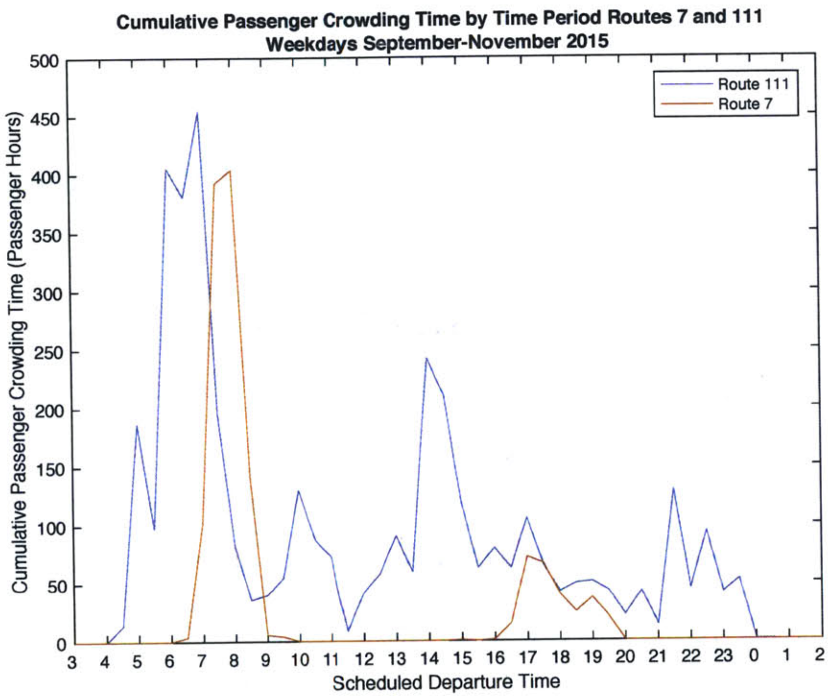


Figure 4-8: Comparison of Temporal Distribution of CPCT- Routes 7 and 111. Crowding Threshold of 140% of Seated Capacity. Weekdays September-November 2015.

downtown earlier in the morning and therefore may finish work earlier as well. Route 7 connects the increasingly affluent South Boston neighborhood to the Financial District in which more passengers may be working traditional 9-5 work hours.

Overall distributions vary significantly between the two routes. Almost all of Route 7's CPCT occurs during the peak periods while on Route 111 crowding occurs throughout the day. This indicates that broad interventions might be appropriate for the 111.

4.6 New Crowding Metrics

While cumulative passenger crowding time attempts to consolidate all crowding aspects into a single metric, it is also important to evaluate individual components to better understand the nature of each crowding situation. Four components are examined: the number of unique passenger trips experiencing crowded conditions, number of unique standees, number of passengers experiencing crowded conditions for extended periods of time, and average standing time.

4.6.1 Number of Unique Passenger Trips Experiencing Crowded Conditions

The number of unique passenger trips experiencing crowded conditions was calculated for routes with the most 140% of seated capacity CPCT. Experiencing crowded conditions is defined as being on a vehicle while loads exceed 140% of seated capacity for some portion of a passenger's trip. The number of vehicle trips on which passengers experience these conditions is also calculated. The results are shown in table 4-5.

Rank	Route	Passenger Trips Affected	Crowded Vehicle Trips	Rank	Route	Passenger Trips Affected	Crowded Vehicle Trips
1	111	117,491	1754	11	23	52, 619	777
2	1	61,686	860	12	32	80,954	1328
3	9	60,000	811	13	39	45,941	452
4	7	55,913	838	14	16	34,398	508
5	66	79,551	1075	15	70	34,240	472
6	57	87,236	1294	16	65	27,940	422
7	86	41,103	576	17	57A	32,149	488
8	47	30,197	434	18	117	32,466	462
9	28	55,161	559	19	109	25,154	388
10	77	36,198	523	20	93	25,661	393

Table 4.5: Passengers Experiencing Crowded Conditions (loads over 140% of Seated Capacity) for Routes with Most 140% of Seated Capacity CPCT. Weekdays September-November 2015

Although the number of passengers experiencing crowded conditions is generally correlated with CPCT, specific route characteristics cause relative fluctuations. The number of vehicle trips in which loads exceed 140% of seated capacity greatly affects the number of passengers who experience crowded conditions as everyone onboard experiences it no matter the intensity or duration of the crowding situation, meaning that routes on which many trips may barely exceed 140% of seated capacity will have more passengers experiencing crowded conditions compared to routes on which only a few trips greatly exceed 140% of seated capacity.

Route 111 has substantially more vehicle trips with loads over 140% of seated capacity than any other route, subsequently causing almost twice as many passengers to experience crowded conditions. Routes 66, 57, and 32 also had more trips exceed the crowding threshold than routes with similar CPCT.

The seated capacity of the vehicle in use also affects the number of passengers experiencing crowded conditions. Routes 28 and 39 operate with 60 foot articulated buses with a seated capacity of 57 which is significantly higher than the 39 seated capacity of the 40 foot vehicles used on the vast majority of routes. Both of these routes have more passengers experiencing crowded conditions on a similar number of crowded vehicle trips relative to routes with comparable CPCT totals.

4.6.2 Crowding Duration

The ridership patterns, route length and travel speeds of a route affect the amount of time passengers spend in crowded conditions. As shown in Section 4.2.2, passengers on certain routes are more likely to remain in crowded conditions for extended periods of time. An analysis of the distribution of passenger crowding time on routes with the most 140% of seated capacity CPCT is shown in Table 4.6.

There is significant variation between routes. Some have a high percentage of passengers who experience crowded conditions experiencing them for long durations. More than half of these passengers experience crowded conditions for over 10 minutes on Routes 111, 47, and 86. On other routes, very few passengers experience long duration crowding. Only 15% of passengers experiencing crowded conditions on Routes 32 and 117 experience them for longer than 10 minutes.

There is also a temporal component to a passenger's crowding duration. During the peak periods, travel speeds are slower due to more vehicles on the road causing congestion. Loads are also generally higher due to the relaxed crowding threshold built into the schedule causing crowding conditions to occur on longer spatial segments. Therefore, the amount of time passengers spend in crowding conditions also increases. Figure 4-9 shows how the average amount of time a passenger spends in crowded conditions changes throughout the day on Route 111.

There is a significant increase in the average crowding duration during the morning peak. Average durations increase from approximately 12 minutes for most periods

		Crowding Duration			
Rank	Route	>5minutes	>10 minutes	>20 minutes	>30 minutes
1	111	92	68	16	4
2	1	66	32	4	1
3	9	77	49	17	1
4	7	88	36	1	0
5	66	68	39	11	2
6	57	66	32	4	0
7	86	78	50	15	2
8	47	81	63	18	4
9	28	70	35	9	1
10	77	71	43	13	3
11	23	70	34	8	2
12	32	44	15	0	0
13	39	63	33	3	0
14	16	73	36	4	1
15	70	67	47	12	4
16	65	74	49	8	1
17	57A	74	37	3	0
18	117	42	16	3	1
19	109	79	29	1	0
20	93	78	35	2	0

Table 4.6: Duration Distribution of Passengers' Crowded Experience on Routes with most CPCT. Weekdays September-November 2015. Red Cell $\geq$75% Orange Cell= 50-74% Yellow Cell=25-50% Green Cell $\leq$ 25%

to over 24 for vehicle trips leaving during the 7 o'clock hour. This is likely due to slow travel speeds over the Tobin Bridge and crowding conditions occurring earlier in the route. There is also a midday lull likely caused by the reverse effect. Afternoon peak travel demand is more spread reducing congestion's effect on running times and therefore crowding duration.

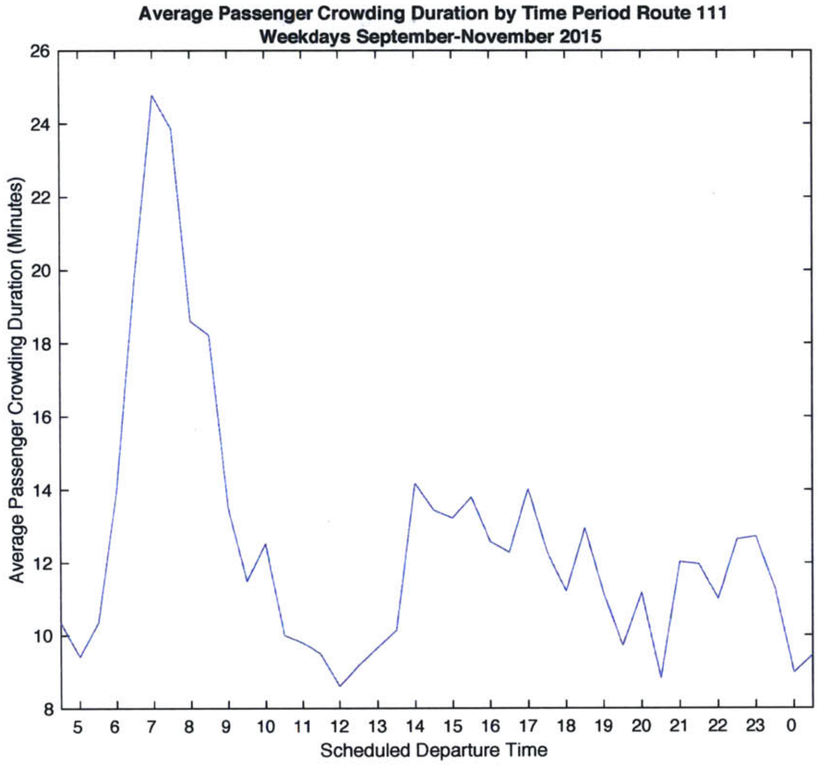


Figure 4-9: Average Passenger Crowding Duration by Time Period Route 111 Weekdays September-November 2015

4.6.3 Number of Standees

To isolate the passengers who are most affected by crowded conditions, the unique number of passenger trips on which a passenger had to stand for some portion of their trip can be identified. While this is similar to the number of unique passenger trips experiencing crowded conditions, it considers all intensities of crowding. Table 4.7 shows the number of standees for routes with the most seated capacity CPCT.

It is assumed that passengers sit when a seat is available and seats are filled on a first-in, first-out sequence. This is required since a passenger's actual actions are unknown. A passenger is only assumed to have stood if they board while a vehicle has a load exceeding its seated capacity.

This reveals a route's turnover rate. The comparison between Routes 66 and 111 highlight this. Route 66 is a crosstown route with many midpoint destinations while the 111 has a single primary destination. More passengers were required to stand on

Rank	Route	Standees	Vehicle Trips Exceeding Seated Capacity	Rank	Route	Standees	Vehicle Trips Exceeding Seated Capacity
1	111	81,867	4,982	11	7	30,789	1,691
2	66	91,212	3,600	12	77	31,735	1,728
3	57	76,927	3,640	13	39	43,974	1,609
4	23	63,791	3,354	14	47	22,726	1,200
5	1	61,113	3,075	15	16	34,582	1,670
6	28	56,809	2,073	16	22	38,494	2,268
7	9	44,074	1,945	17	34E	23,776	1,522
8	32	59,827	4,468	18	65	19,804	1,065
9	86	35,887	1,816	19	57A	21,671	1,035
10	70	38,669	1,781	20	93	19,390	1,244

Table 4.7: Number of Unique Standing Passenger Trips and Vehicles Trips Exceeding Seated Capacity for Routes with Most Seated Capacity CPCT. Weekdays September-November 2015.

Route 66 though this occurred on significantly fewer vehicle trips than Route 111. For every Route 66 trip that exceeded seated capacity, many more passengers were required to stand due to higher turnover rates.

4.6.4 Average Standing Time

It is also possible to calculate the average amount of time that a passenger is required to stand when a seat is not available. This was calculated for the routes in the rank order included in Table 4.7.

Rank	Route	Average Standing Time (Minutes: Seconds)	Rank	Route	Average Standing Time (Minutes: Seconds)
1	111	12:01	11	7	8:57
2	66	6:54	12	77	8:32
3	57	6:54	13	39	6:01
4	23	6:52	14	47	10:37
5	1	7:00	15	16	6:47
6	28	7:16	16	22	5:43
7	9	9:06	17	34E	7:56
8	32	5:49	18	65	9:39
9	86	8:48	19	57A	7:42
10	70	8:32	20	93	7:52

Table 4.8: Average Standing Time on Routes with Most Seated Capacity CPCT. Weekdays September-November 2015.

There is a wide range of average standing times. Route 111 has the longest standing time and it is almost double the duration of the route with shortest, Route 22. The range is likely caused by two factors: turnover rate and duration of crowded segments. Routes on which crowded segments are relatively short will only require passengers to stand for limited amounts of time regardless of turnover rate. This appears to be the case for Routes 22 and 32.

High turnover rates can shorten the average amount of time that each passenger spends standing. This likely causes Routes 1 and 66 to have relatively short average standing times compared to Routes 47 and 111 where turnover rates are lower.

4.7 Proposed Crowding Standard for the MBTA

A standard was also developed as a diagnostic tool to enable the identification of routes on which a significant percentage of passenger travel time is spent in uncomfortable conditions. It states that over a given time period, a set percentage of passenger time must be spent in comfortable conditions for a given route.

Comfortable conditions for passengers are defined only for those who are seated when vehicle loads are less than 140% of seated capacity. If vehicle loads exceed 140% of seated capacity, all passengers are said to be traveling in uncomfortable conditions. Figure 4-10 shows an example calculation. The black line represents a seated capacity of 39 while the blue line represents 140% of seated capacity. The red areas are passenger time spent in comfortable conditions. In this example, about 33% of passenger time is spent in comfortable conditions.

Cumulative total passenger time and comfortable passenger time can summed across trips departing during a given time period for a extended time frame (i.e., Route 1 Inbound Trips Departing Between 8:00-8:30AM during July 2015). Then a period percentage can be calculated.

The comfortable percentage thresholds can be adjusted for different time periods to account for different acceptable crowding conditions. Thresholds can be lower during the peak periods when more crowding is acceptable and higher during the off

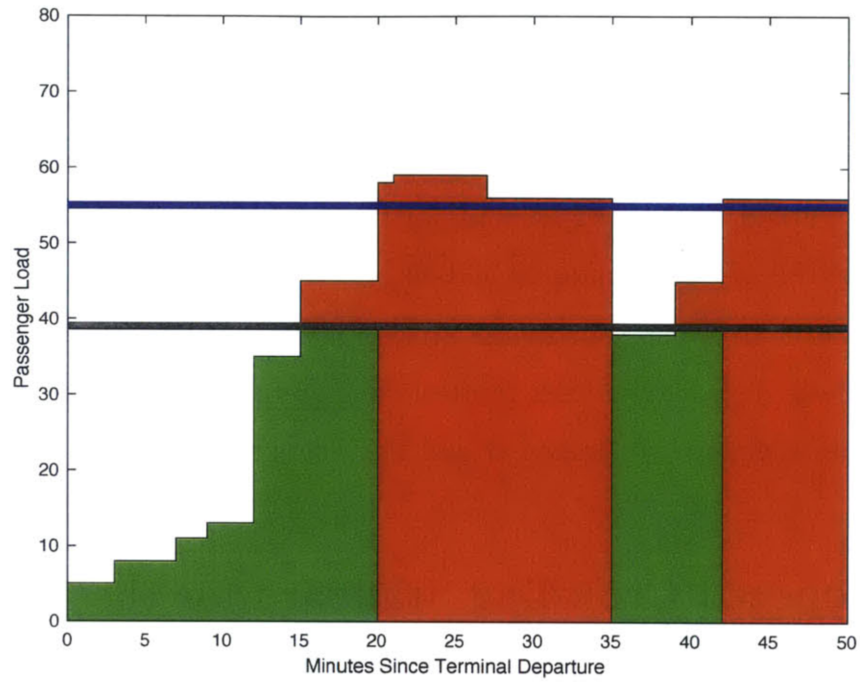


Figure 4-10: Example Calculation of Percentage of Passenger Time Spent in Comfortable Conditions. Green areas are comfortable passenger time while red areas are uncomfortable passenger time. The black line is seated capacity while the blue is 140% of seated capacity.

peaks when passengers often expect to have a seat available.

The standard allows for the passenger experience be evaluated regardless of the frequency of service provided. Where the prioritization calculations using the previously calculated metrics consider the number of passengers affected by crowded conditions to enable comparisons, the proposed standard allows a route to be evaluated in isolation of frequency to determine whether a route/time-period requires an intervention.

4.8 Summary of Metrics

Each of the metrics discussed in the previous sections have unique advantages and limitations. These are outlined in Table 4.9. To make the most use of these varied metrics, it is recommended that they are used together in some combination to gain the best understanding of crowding in a network. An example application could entail using a threshold percentage of passenger time spent in comfortable conditions as a standard to identify "crowded" routes and time periods. Then CPCT can be used to prioritize resource use among these "crowded" routes and time periods. Finally, the distribution of passenger crowding duration can be used to identify routes and time periods on which crowding intensity should be kept at a minimum as passengers are more significantly impacted when crowding does occur.

Metric	Advantages	Limitations
Cumulative Passenger Crowding Time (CPCT)	Considers multiple aspects of crowding and enables prioritization of routes and time periods	Difficult to determine appropriate level of crowding
Unique Crowded Passengers	Identifies the number of passengers affected by crowding	No information on magnitude or distribution of impacts for each passenger
Crowding Duration	Identifies where crowding is likely to have largest impact on affected customers	Does not measure how many passengers are affected
Standees	Identifies passengers most seriously affected by crowding	No information on magnitude of impact for each passenger
Average Standing Time	Identifies where crowding is likely to have largest impact on the most affected customers	Does not measure how many passengers are affected
Percentage of Passenger Time Spent in Comfortable Conditions	Enables the identification of route and time periods on which crowding exceeds an acceptable level regardless of frequency	Does not easily allow prioritization of interventions

Table 4.9: Summary of Metric Advantages and Limitations

4.9 Conclusions

This chapter discusses how ODX derived trip level load and passenger flow estimates can be used to gain a more nuanced understanding of passenger crowding throughout a bus network. These estimates can also be used to develop passenger centric crowding metrics such as CPCT. Other aspects of crowding beyond intensity can be evaluated including crowding duration and the number of unique passengers who are affected by crowding either by being in a crowded vehicle or having to stand for a portion of his/her trip. These metrics also allow for comparisons between routes and time periods enabling transit agencies to identify where passengers are most effected by crowding.

In the case of the MBTA bus network, CPCT reveals that during the fall of 2015, most crowding occurs in a relatively concentrated portion of the bus network. The vast majority of crowding occurs during the peak periods as the crowding standard is relaxed and congestion causes running times to increase. It is also concentrated mostly on high frequency urban routes, with approximately 66% of 140% seated capacity CPCT occurring in the top twenty most crowded routes and approximately 7% of unlinked passenger trips experiencing loads over 140% of seated for some portion of their trip, since schedules are set based on average peak load analysis.

Within the group of most crowded routes as identified with CPCT, the nature of the crowding situations varied. Some routes had many short duration crowded vehicle trips such as the 32 and 117 while others had long duration crowding such as 111, 86, and 47.

Although most crowding occurs during the peak periods, routes have different crowding peaks within the overall peak periods. Comparing the temporal distribution of crowding of Routes 111 and 7, for example show that for both the morning and afternoon peaks, Route 111 passengers experience crowded conditions earlier than Route 7.

A standard is also discussed that can be used to evaluate crowding conditions of a route in isolation of the rest of the network. The percentage of passenger time spent in comfortable conditions considers both the intensity and duration of crowded situations. It is also differentiates between low intensity crowding where standees are the passengers most affected by crowding and high intensity crowding where all passengers are affected.

While this research focuses on the use of automated farecard and vehicle location data, full implementation of APC systems on 100% an agency's bus fleet also allows for a more granular estimation of passenger loads and therefore a better understanding of the aggregate passenger crowding experience than is currently available using a sample of trips. There is a lot to gained by using accumulated metrics instead of averages. Many, but not all, of the alternative metrics discussed in this chapter could be calculated with complete coverage of APC load estimates. However, additional estimation procedures such as iterative proportional fitting (IPF) are required to estimate how individual passengers are affected when only using APC load estimates, as only boardings and alightings are estimated. The IPF estimation is similarly imprecise as is the ODX procedures discussed herein. And it should be noted that the passenger flow estimates provided directly from ODX provide a better understanding of the crowding impact on individual passengers without further estimation.

For example, even a metric like CPCT, which could be estimated with full coverage of APC systems groups all passengers together to create a total passenger crowding time. However, as can be seen with the distributions of passenger crowding durations in Table 4.6, each passenger's experience of crowding can be different, even within a given vehicle trip where there can be a wide range of experiences. ODX provides both the aggregate and more fine-grained individual passenger crowding information to not only identify where the most passenger crowding occurs but also when and where an individual is mostly likely to be negatively affected by crowding.

Thus, while transit agencies may have previously believed that the "gold standard" method to measure and evaluate passenger crowding on buses was to equip every bus with APC systems, this research provides quite compelling arguments that the joint use of AFC and AVL data with the ODX algorithm along with a sample of APC data can provide an additional base of comparison of crowded conditions throughout a high demand urban bus system.

Chapter 5

Crowding Source Determination

The wealth of trip level load and passenger flow information available from ODX derived estimates also enables further investigation of the causes of crowding for each route/direction/time-period combination. This helps inform a transit agency's selection of the most appropriate passenger experience improvement strategy for a given situation. A more nuanced crowding reduction program can increase the degree to which the overall passenger experience can be improved given finite resources.

A hierarchical classification tree method is proposed to determine the portion of observed crowding, in terms of seated capacity CPCT, that can be attributed to one of four categories. These categories include: scheduled frequency, dropped trips, daily fluctuations in passenger demand, and within-period load variability factors. Although there are many assumptions made in this analysis, it provides a method to estimate the magnitude of effect for each potential "reason" for crowded conditions.

5.1 Sources of Crowding

There are many are many factors that cause crowding situations to arise on fixed route bus service. A summary is shown in Figure 5-1. The base factors are the fixed schedule not being able to meet fluctuations in demand and inadequate supply. The fixed schedule's inability to meet normal passenger travel demand is crowding that occurs when there is peaking in passenger travel demand (e.g., adequate capacity over the course of a scheduling period but not enough to meet demand for one or more specific vehicle trips on a specific day). This variability in demand can occur on large scale such as between days or smaller scales such as between vehicle trips within a specific small time period. Since these fluctuations are generally not routine, they are not incorporated into the schedule.

On the supply side, some crowding is inherent in the scheduled frequency. This may be intended during the peak periods when resources need to be spread across the entire network or unintended if schedules have not been adjusted over time to meet higher levels of average demand.

Other reasons for crowding can be attributed to the provided service not operating as scheduled. There are two primary areas in which provided service can fall short of scheduled service: dropped trips and headway variability. Drop trips are when a scheduled vehicle trip is not operated for any reason. This lowers average provided frequency as well as increases headway variability, both requiring passengers to be distributed among fewer vehicle trips as well as resulting in passengers being less evenly distributed among these trips. Headway variability causes uneven vehicle loading if passenger arrival rates are constant. It can be either caused by endogenous factors such as poor terminal departure performance or more exogenous factors such as general traffic congestion.

5.2 Methodology

A four stage classification tree is used in the source estimation process as shown in Figure 5-2. At each junction a calculation is made to estimate the portion of crowding to be attributed to each sequential branch. This requires controlling for both duration and intensity components of crowding at different points throughout the procedure.

The classification process starts by clustering the Fall of 2015 non-school-holiday weekday vehicle trips on routes which ODX and the scaling process are effective by route, direction and 30-minute-time-periods during the day determined by scheduled departure time. Ideally, actual departure time would be used, however there

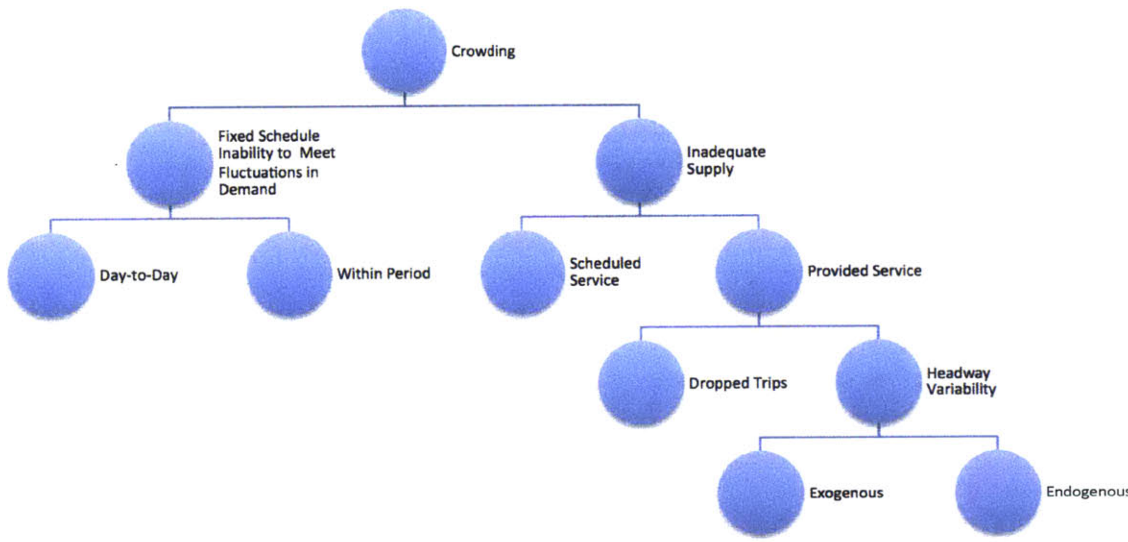


Figure 5-1: Sources of Crowding

were many operated trips in which actual departure times are not available. Observed seated capacity CPCT using scheduled run times is then summed for each route/direction/30-minute-time-period combination. This controls for any effect that deviations from scheduled run times might have on crowding duration. Seated capacity is used as the crowding threshold instead of 140% of seated capacity for all periods to capture crowding of all intensities. This becomes the base measure on which all other sequential calculations are compared.

5.2.1 Scheduled Frequency and Variability Factors

The "scheduled frequency" factor estimates the amount of crowding that would be expected if all scheduled service was delivered perfectly and passenger arrival rates and alighting distributions were constant. It attempts to evaluate the degree to which the scheduled level of service is able to meet average passenger travel demand controlling for day-to-day variations in demand and service performance.

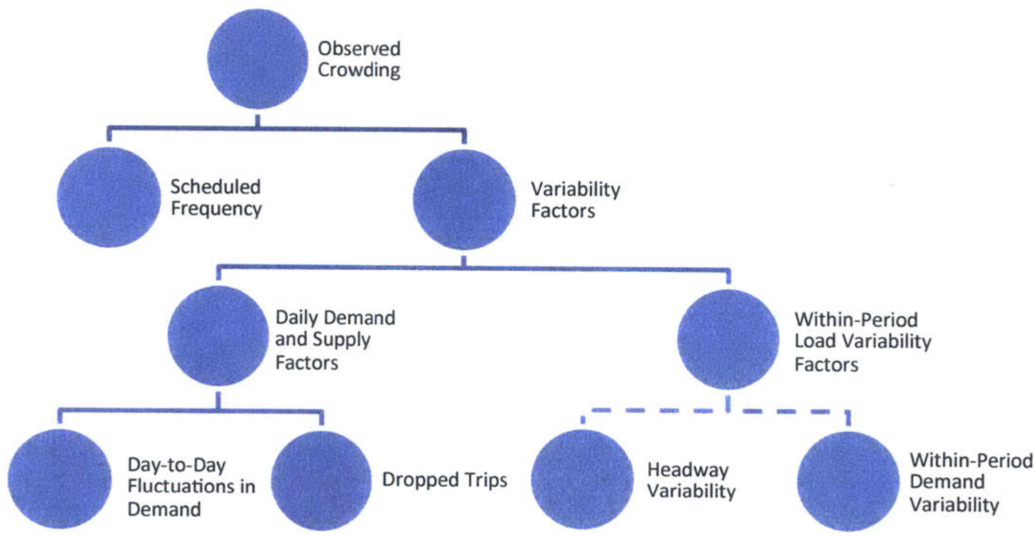


Figure 5-2: Classification Tree Used in Crowding Source Methodology

Variability factors are the antithetic components to provided frequency. They are components that cause vehicle trips within a route/direction/30-minute-time-period to have differing levels of CPCT. These could be either intensity or duration factors. Variability factors included in this analysis are day-to-day fluctuations in demand and factors that cause load variation between trips within a given day/route/direction/period combination that is usually caused by uneven headways.

This attribution split is accomplished by determining the level of crowding, in terms of seated capacity CPCT, that would be expected if passengers were equally split between all scheduled vehicle trips and trips operated with their scheduled departure and running times. First, average load profiles are constructed for each route variation within a route/direction/time-period combination. Then scheduled running times are assigned to each vehicle trip to calculate CPCT. This is divided by the observed seated capacity CPCT to derive an attribution percentage for each route/direction/time-period combination as shown in Equation 5.1.

Figure 5-3 provides an example calculation. Here, there are four vehicle trips with two dropped trips for a total of 6 scheduled trips. Trip 1 has 100 passengers minutes of CPCT. Trip 2 has 75 passenger minutes while loads on trips 3 and 4 remain under the crowding threshold for the entire duration. In total there are 175 passenger minutes of CPCT experienced on the four trips. The scheduled average load profile of the four trips is represented by the black line. This remains under the crowding threshold for the entire duration. Therefore, if every scheduled trip was operated and passengers were equally distributed among trips with the same boarding and alighting distributions we would not expect any crowding.

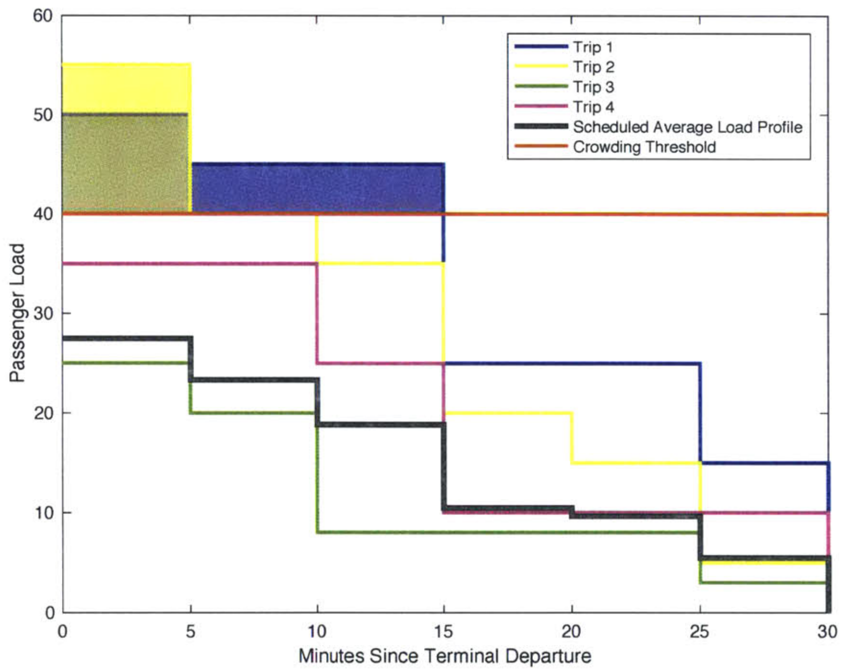


Figure 5-3: Example Calculation of Scheduled Frequency Contribution

The remaining CPCT that cannot be attributed to scheduled frequency is attributed to variability factors (In the example case, all 175 passenger minutes of CPCT). This is also divided by the observed CPCT to attain an attribution percentage as shown in Equation 5.2.

$$SF = \frac{CPCT_{SS}}{CPCT_{OS}} \tag{5.1}$$

$$VF = \frac{CPCT_{OS} - CPCT_{SS}}{CPCT_{OS}} \tag{5.2}$$

Where

- SF is the percent of observed seated capacity CPCT attributed to scheduled frequency
- VF is the percent of observed seated capacity CPCT attributed to variability factors
- $CPCT_{SS}$ is the CPCT calculated using period scheduled average load profiles and scheduled run times
- $CPCT_{OS}$ is the CPCT calculated using observed loads and scheduled run times

There is a possibility that there is some passenger demand peaking within a period. (e.g., within the 8:00-8:30 AM period the 8:15AM trip is always more crowded than the 8:25AM) This might cause more crowding to be attributed to variability factors that should be in reality. Time periods are set to a relatively short 30 minutes in order to minimize this possibility.

There is also a degree of load variability inherent in the ODX scaling process due to the use of average scaling factors. This may cause the scheduled frequency component contribution to be slightly under-attributed on some routes.

5.2.2 Daily Demand and Supply Factors and Within-Period Load Variability Factors

The amount of crowding that can be attributed to day-to-day fluctuations in demand as well as the number of vehicle trips actually provided on any given day can be defined as the portion of crowding that results from the actual provided frequency levels not being able to meet passenger travel demand on any given day. Passenger travel demand is likely to vary between days for a number of reasons. A few examples include: individuals may decide to ride the bus for a trip they usually make on a bike or by walking if the weather is bad; there may be an event that draws people somewhere on a particular day increasing travel demand along a route or existing public transit users might even be attracted to a particular route on a given day if there is a disruption on another portion of the system. Actual provided frequency also varies on a daily basis if any vehicle trips are dropped on a given day.

This daily versus within-period variation split is attempting to determine what portion of crowding observed can be attributed to the provided frequency levels not being able to meet passenger travel demand on any given day compared with crowding brought on by factors that lead to variations in CPCT between adjacent vehicle trips within a given day including uneven headways, and within-period variation in passenger arrival rates and boarding/alighting distributions.

This is estimated by first calculating the daily average load for each route/direction combination for each period and day it is in operation. Then using scheduled run times, CPCT is calculated for every trip and summed for all trips within a route/direction/time-period combination. This sum becomes the portion of crowding attributed to daily demand and supply factors, as shown in Equation 5.3.

The remaining CPCT is attributed to factors causing CPCT to vary between adjacent trips, as shown in Equation 5.4. For low frequency periods where only a single trip is scheduled this is always zero. This reflects the high likelihood that passengers during these periods select a specific vehicle trip to take each day and therefore crowding is less likely to be affected by headway variation.

$$DPV = \frac{CPCT_{DAS}}{CPCT_{OS}} \tag{5.3}$$

$$LV = \frac{CPCT_{OS} - CPCT_{DAS}}{CPCT_{OS}} \tag{5.4}$$

Where:

- DPV is the percent of observed seated capacity CPCT attributed to daily provided frequency
- LV is the percent of observed seated capacity CPCT attributed to factors causing loads to vary between adjacent vehicle trips
- $CPCT_{DAS}$ is the CPCT calculated using the daily average load profile and scheduled run times
- $CPCT_{OS}$ is the CPCT calculated using observed loads and scheduled run times

An analysis of the example trips used in Figure 5-3 can be expanded with this methodology. If trips 1 and 3 occurred on day 1 and trips 2 and 4 occurred on day 2 then the daily average load profiles shown in Figure 5-4 would result. On day 1 we would not expect any crowding if passengers were evenly distributed between the two trips. On day 2 we would expect 25 passenger minutes of CPCT per trip for a total of 50 passenger minutes.

Therefore, 50 passenger minutes is attributed to daily demand and supply factors. The remainder of CPCT is attributed to within-period load variation factors, in this case 125 passenger minutes.

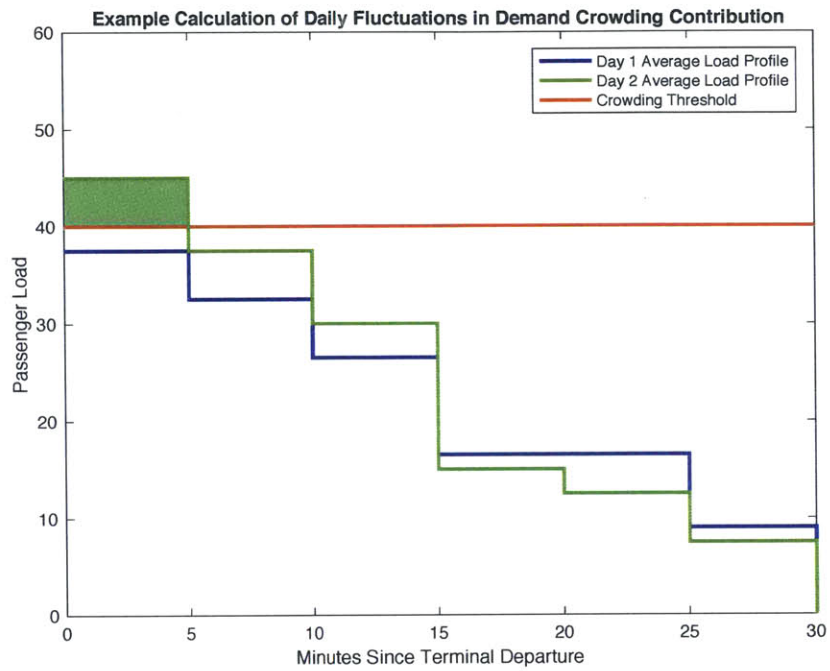


Figure 5-4: Example Calculation of Daily Demand and Supply Factors Crowding Contribution

5.2.3 Headway Variability and Within-Period Demand Variability

Further analysis may be necessary to definitively separate within-period load variation factors into service delivery factors (such as uneven headways) and aspects that a transit agency has little control over such as fluctuating passenger arrival rates and boarding/alighting distributions. Using the passenger arrival rates indicated by the number of passengers boarding each vehicle at each stop as well the alighting distributions of these passengers, loads could be estimated under a perfectly delivered service scenario. The additional crowding estimated with operated service delivered perfectly compared to the crowding attributed to daily demand and supply factors would be the amount of crowding attributed to within-period demand variability (varying passenger arrival rates and boarding/alighting distributions) as shown in Equation 5.5. The remaining CPCT is attributed to uneven headways as in Equation 5.6. This process assumes that there were no left behinds and that passengers arrived at a constant rate between actual vehicle arrivals. Though this partition was not included in the overall model, initial results suggest that the vast majority of crowding due to within-period load variability factors is due to uneven headways. An example application is shown for Inbound Routes 1 and 9 in Tables 5.2 and 5.3 respectively. More work should be done on this topic to expand this analysis systemwide.

$$WPDV = max(\frac{CPCT_{PS} - CPCT_{DAS}}{CPCT_{OS}}, 0) \tag{5.5}$$

$$HV = min(\frac{CPCT_{OS} - CPCT_{PS}}{CPCT_{OS}}, \frac{CPCT_{OS} - CPCT_{DAS}}{CPCT_{OS}}) \tag{5.6}$$

Where:

- $WPDV$ is the percent of observed seated capacity CPCT attributed to within-period demand variability.
- HV is the percent of observed seated capacity CPCT attributed to headway variability.
- $CPCT_{PS}$ is the CPCT estimated if operated vehicle trips run perfectly on schedule.
- $CPCT_{DAS}$ is the CPCT calculated using the daily average load profile and scheduled run times.
- $CPCT_{OS}$ is the CPCT calculated using observed loads and scheduled run times.

5.2.4 Daily Factors: Day-to-Day Fluctuations in Demand and Dropped Trips

The day-to-day fluctuations in demand factor is the amount of crowding that would be expected in addition to overall scheduled frequency if all scheduled service was delivered perfectly and passenger arrival rates and boarding/alighting distributions were constant, due to surges in demand on particular days. It is attempting to determine how effectively the scheduled level of service was able to meet passenger travel demand on each day controlling for within-period load variability factors. While overall scheduled frequency captures the average amount of crowding expected each day from the schedule, day-to-day fluctuations in demand captures the impact that fluctuations in passenger demand have on the ability of the schedule to meet demand on a specific day.

The dropped trips component is the additional crowding expected from the lowering of average frequency due to vehicle trips being dropped. It only considers the impact that a dropped trip has on the average frequency provided. If trips are dropped without making adjustments to the schedule, it very likely that there will be increases in headway variability. This effect is captured in the within-period load variability component.

If all operated vehicle trips had ODX derived estimates then the average load could be calculated in a similar process as described in Section 5.1.1 using scheduled trips as the denominator. However, since ODX derived load profiles are only able to be calculated on approximately 90% of operated vehicle trips, a different method needed to be developed.

Instead the percentage of trips dropped for each route/direction/period/day was estimated. Then this percentage was added to the number of trips observed. Average loads were calculated by aggregating passenger flows from all trips within a particular period on a specific day with load estimates and dividing these aggregate flows by the adjusted trip total.

CPCT is then calculated using scheduled run times and multiplied by the dropped trips adjustment factor. The CPCT calculated from overall scheduled frequency is then subtracted from this sum to provide the amount of crowding attributed to day-to-day fluctuations in demand, shown in Equation 5.7. The difference between the amount of crowding attributed to daily demand and supply factors and the amount of crowding attributed to the scheduled frequency not able to meet demand on a specific day is attributed to dropped trips, as in Equation 5.8.

$$DD = \frac{CPCT_{DSS} - CPCT_{SS}}{CPCT_{OS}} \tag{5.7}$$

$$DT = \frac{CPCT_{DAS} - CPCT_{DSS}}{CPCT_{OS}} \tag{5.8}$$

Where:

- DD is the percent of observed seated capacity CPCT attributed to day-to-day fluctuations in demand
- DT is the percent of observed seated capacity CPCT attributed to dropped trips
- $CPCT_{SS}$ is the CPCT calculated using an overall scheduled proxy average load and scheduled run times
- $CPCT_{DSS}$ is the CPCT calculated using a daily scheduled proxy average load and scheduled run times
- $CPCT_{DAS}$ is the CPCT calculated using daily average load profiles and scheduled run times
- $CPCT_{OS}$ is the CPCT calculated using observed loads and scheduled run times

For example, if there were six scheduled trips during the period depicted in Figure 5-3, meaning that two trips were dropped, one on each day, the schedule adjusted average load profile would remain under the crowding threshold for the entire length of the route. This is shown in Figure 5-5. Therefore, no crowding is attributed to day-to-day fluctuations in demand. All of the crowding attributed to daily demand and supply factors is due to the decrease in frequency resulting from trips being dropped.

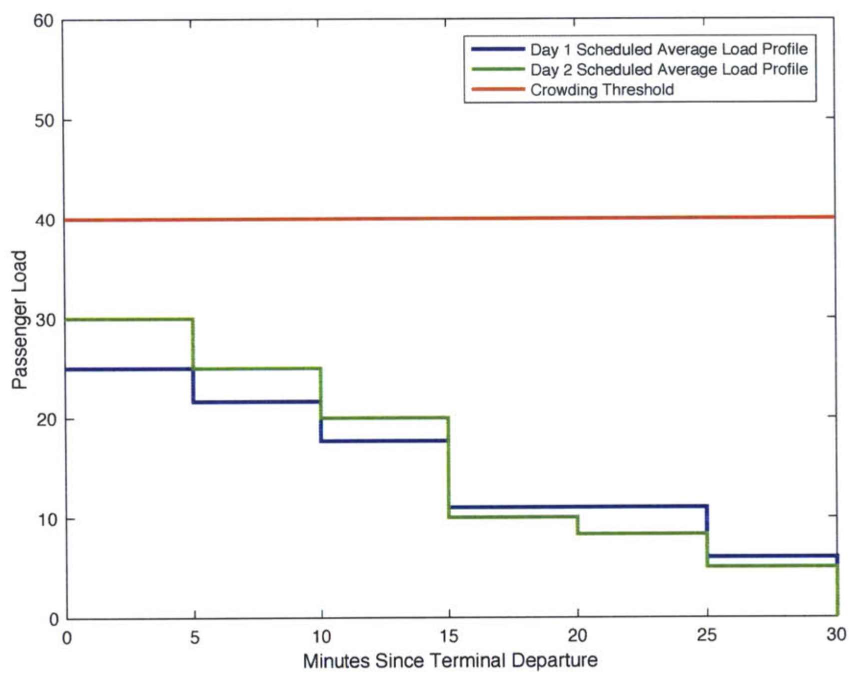


Figure 5-5: Example Calculation of Day-to-Day Fluctuations in Demand Contribution

Therefore, for this example, 0% of the observed crowding is attributed to scheduled frequency, 29% is attributed to dropped trips, 0% is attributed to day-to-day fluctuations in demand, and 71% is attributed to within-period factors including trip-to-trip load variations caused primarily by headway variation.

Total CPCT Passenger Minutes	175	100%
Scheduled Frequency	0	0%
Day-to-Day Fluctuations in Demand	0	0%
Dropped Trips	50	29%
Within-Period Load Variability Factors	125	71%

Table 5.1: Example Period Component Analysis from Figures 5-3 Through 5-5

5.3 Example Application

The procedure can be used to explore how the causes of crowding change through the day along a route or between routes. Tables 5.2 and 5.3 depict the attribution percentage and Figures 5-6 and 5-7 depict the absolute attribution of each component for every weekday period in which trips were scheduled for inbound Routes 1 and 9 respectively during the Fall of 2015 to highlight differences between routes.

On Route 1, there was observed crowding during most operating periods with higher amounts during the peak periods. Only during the early morning and overnight periods was no crowding observed. For the periods in which crowding was observed, it appears to be primarily initiated by factors causing within-period load variability, of which the vast majority is caused by uneven headways. For a number of time periods, within-period demand variability has some impact. Since there are many popular boarding and alighting locations along the route, passenger boarding/alighting distributions are somewhat more likely to vary. Passengers are also likely to arrive in bunches at the many significant transfer points along route causing passenger arrival rates to vary. For vehicle trips scheduled to depart between 7:30 and 8:30 there was some crowding due to scheduled frequency which accounts for less than 25 percent of observed CPCT during those periods.

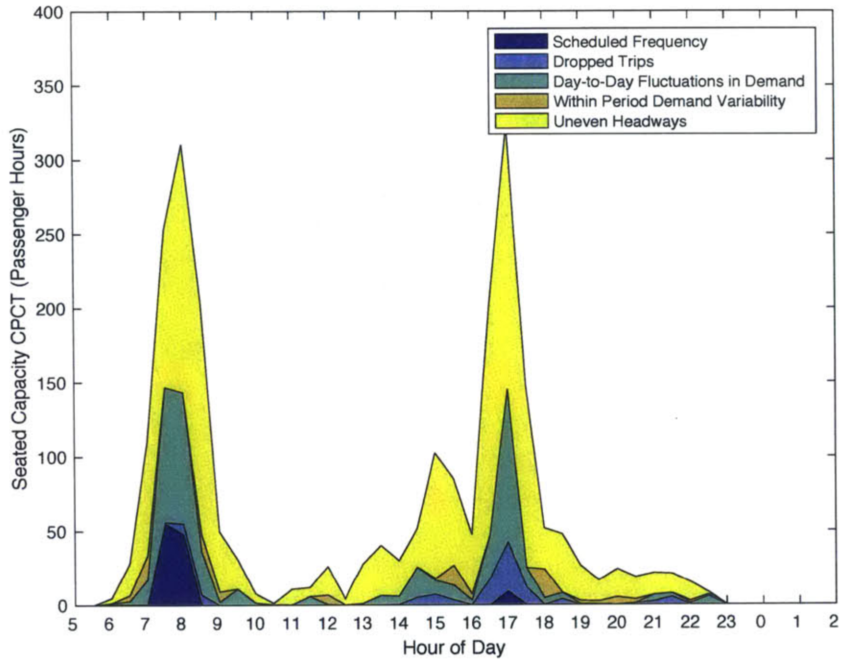


Figure 5-6: Absolute Crowding Source Contribution Inbound Route 1

During many periods, day-to-day fluctuations in demand have a non trivial impact though it rarely accounts for more than 50% of observed crowding.

Time Period	Total CPCT (Passenger Hours)	Scheduled Frequency	Dropped Trips	Day-to-Day Fluctuations in Demand	Headway Variability	Within-Period Demand Variability
5:00	0	N\A	N\A	N\A	N\A	N\A
5:30	0	N\A	N\A	N\A	N\A	N\A
6:00	5	0	20	9	71	0
6:30	28	0	1	8	76	14
7:00	109	0	1	15	69	15
7:30	253	22	0	36	42	0
8:00	310	15	2	28	54	0
8:30	204	0	3	14	77	6
9:00	50	0	0	4	82	14
9:30	31	0	0	35	65	0
10:00	8	0	0	19	81	0
10:30	1	0	0	17	77	6
11:00	11	0	0	1	99	0
11:30	12	0	0	48	52	0
12:00	26	0	0	1	74	25
12:30	4	0	0	0	100	0
13:00	28	0	0	4	96	0
13:30	40	0	0	16	84	0
14:00	30	0	1	19	80	0
14:30	51	0	10	39	50	0
15:00	102	0	7	9	83	1
15:30	85	0	4	12	69	16
16:00	47	0	0	7	85	8
16:30	201	0	9	15	77	0
17:00	323	3	10	32	55	0
17:30	148	0	8	9	83	0
18:00	52	0	0	9	54	37
18:30	48	0	8	8	83	0
19:00	26	0	0	2	90	8
19:30	17	0	0	2	86	12
20:00	24	0	0	0	79	21
20:30	19	0	6	0	81	13
21:00	21	0	11	20	69	0
21:30	21	0	26	11	63	0
22:00	15	0	0	7	93	10
22:30	8	0	0	68	32	14
23:00 and after	0	N\A	N\A	N\A	N\A	N\A

Table 5.2: Percentage Attribution of Crowding Sources for Weekday Inbound Route 1 by Period. September-November 2015.

Time Period	Total CPCT (Passenger Hours)	Scheduled Frequency	Dropped Trips	Day-to-Day Fluctuations in Demand	Headway Variability	Within-Period Demand Variability
5:00	0	N\A	N\A	N\A	N\A	
5:30	4	0	0	12	88	0
6:00	13	0	0	2	98	0
6:30	332	56	0	20	23	0
7:00	677	30	0	18	52	0
7:30	1534	52	3	23	22	0
8:00	1181	37	0	19	44	0
8:30	151	0	0	1	99	0
9:00	12	0	0	20	80	0
9:30	6	0	0	9	91	0
10:00	0	N\A	N\A	N\A	N\A	N\A
10:30	0	N\A	N\A	N\A	N\A	N\A
11:00	0	N\A	N\A	N\A	N\A	N\A
11:30	0	N\A	N\A	N\A	N\A	N\A
12:00	2	0	0	100	0	0
12:30	0	N\A	N\A	N\A	N\A	N\A
13:00	0	N\A	N\A	N\A	N\A	N\A
13:30	24	0	0	12	88	0
14:00	14	0	0	46	54	0
14:30	2	0	0	0	100	0
15:00	3	0	0	86	14	0
15:30	6	0	0	0	100	0
16:00	0	N\A	N\A	N\A	N\A	N\A
16:30	<1	0	0	0	100	0
17:00	1	0	0	0	100	0
17:30	0	N\A	N\A	N\A	N\A	N\A
18:00	0	N\A	N\A	N\A	N\A	N\A
18:30	0	N\A	N\A	N\A	N\A	N\A
19:00	<1	0	0	100	0	0
19:30	0	N\A	N\A	N\A	N\A	N\A
20:00 and after	0	N\A	N\A	N\A	N\A	N\A

Table 5.3: Percentage Attribution of Crowding Sources for Weekday Inbound Route 9 by Period. September-November 2015.

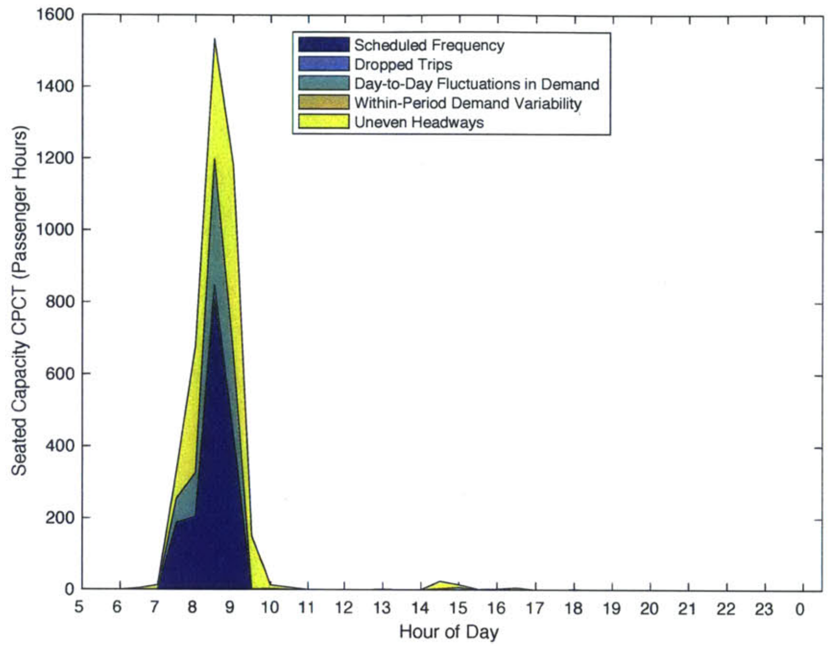


Figure 5-7: Absolute Crowding Source Contribution Inbound Route 9

Much of the crowding observed on Inbound Route 9 occurs during the morning peak period, in which significantly more crowding occurs than Route 1. (Please note the difference in scales between the two figures) In terms of crowding sources, while there are some periods in which a large percentage of crowding can be attributed to factors causing within-period load variability, these are periods with relatively small amounts of crowding. During the morning peak, scheduled frequency accounts for approximately between a quarter and a half of the crowding observed. All of the within-period load variability factor contribution is estimated to be caused by uneven headways. Compared to Route 1 there are fewer destinations along Route 9. Therefore, passenger boarding/alighting distributions are more likely to stay constant.

5.4 Systemwide Analysis

Although each situation should be evaluated individually to determine the appropriate intervention, these component estimates can also be used for systemwide analysis to gain a better understanding of the primary sources of crowding throughout the network. The distribution of percentage contribution across all periods in which passengers experienced some amount of crowding was explored for each component.

5.4.1 Scheduled Frequency

Figure 5-8 shows the cumulative distribution function of scheduled frequency component contributions for periods with observed CPCT. Although these schedule attribution estimates are likely to be slight underestimates due to the factors mentioned in Section 5.1.1, it appears that for most route/direction/time-period combinations, a relatively small portion if any crowding can be attributed to the schedule alone. During approximately 90% of the route/direction/30-minute-time-period combinations in which crowding was observed, ***no*** portion of the crowding was estimated to be caused by the schedule, meaning that average peak loads were less than the seated capacity of the vehicle in use on the particular route. This percentage is slightly higher for off-peak periods when the crowding standard is the seated capacity compared to the peak and extended peak periods when the crowding standard is relaxed to 140% of seated capacity and some crowding is expected. There are very few periods in which the schedule accounts for the majority of passenger crowding.

5.4.2 Dropped Trips

A CDF of dropped trip component percentages is shown in Figure 5-9. It follows similar trends to the scheduled frequency distribution though slightly more periods have a portion of crowding attributed to dropped trips. For approximately 75% of route/direction/time-period combinations with observed CPCT, drop trips are not estimated to be responsible for passenger crowding. There are two possible explanations. A period might not have a trip dropped through the entire three month analysis duration. Therefore, the daily scheduled frequency is identical to the daily provided frequency. If trips are dropped during a period, it is possible that the dropped trip component accounts for zero percent of crowding if no crowding is expected with the provided frequency. In this case, daily average loads would increase when accounting

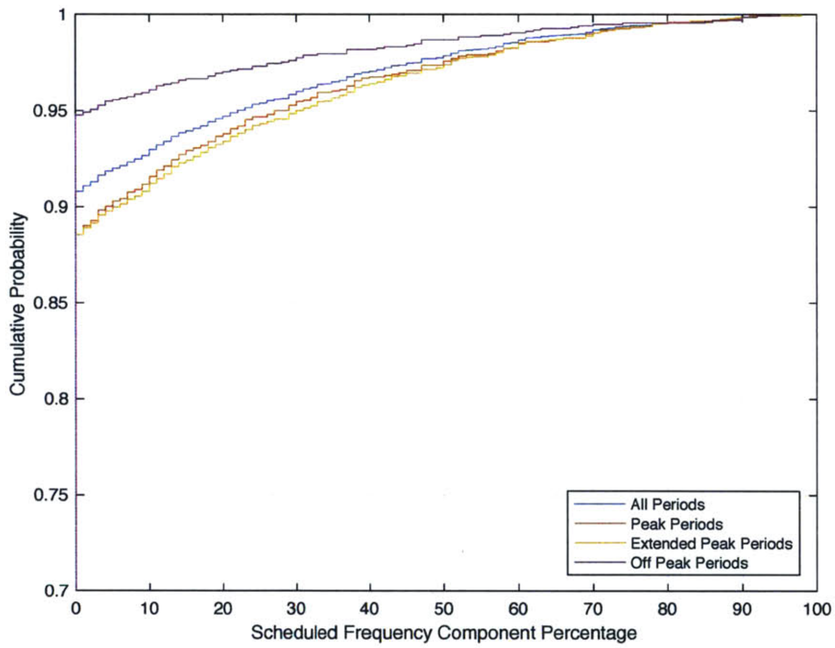


Figure 5-8: CDF of Scheduled Frequency Component Percentages for Route/Direction/Time-Period Combinations with Observed CPCT. Weekdays September-November 2015.

for the dropped trips; however daily average peak loads remain under seated capacity. In such cases, both scheduled frequency and dropped trips would have no affect on creating crowding conditions. For the majority of route/direction/30-minute-time-period combinations for which dropped trips contribute to some crowding, it contributes less than 15%.

A higher percentage of route/direction/time-period combinations that occur during the peak periods appear to be affected by dropped trips compared to combinations occurring during the off-peak periods. This is likely due to higher dropped trip rates as the system is functioning near its operational capacity during the peaks when additional resources (either operator or vehicle) may not be available. It may also be caused by peak period schedules allowing for more crowding therefore fewer situations are likely to have provided frequency able to maintain average peak loads under seated capacity.

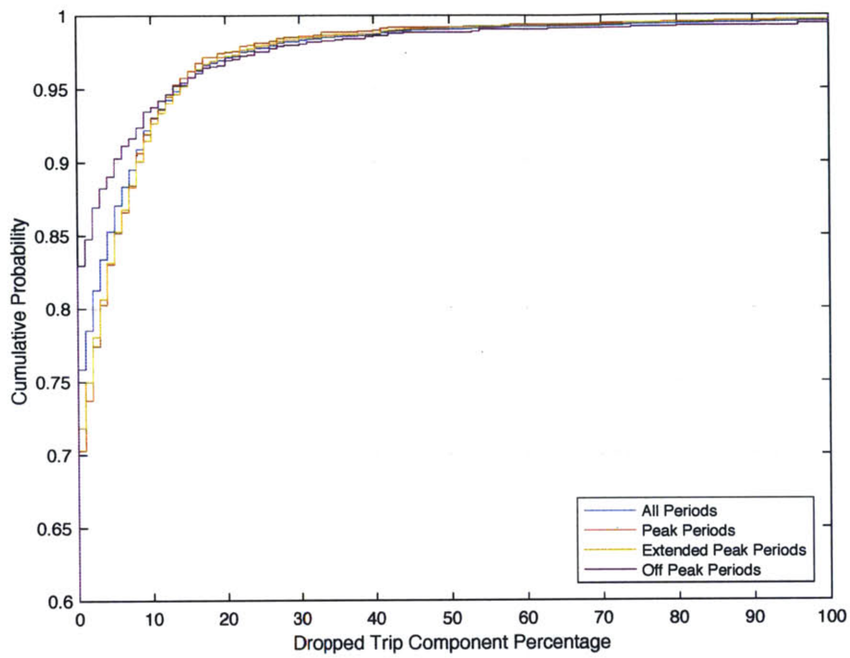


Figure 5-9: CDF of Dropped Trips Component Percentages for route direction time period combinations with Observed CPCT. Weekdays September-November 2015.

5.4.3 Day-to-Day Fluctuations in Demand

The distributions of day-to-day fluctuations in demand component percentages are shown in Figure 5-10. For the majority of route/direction/time-period combinations, daily fluctuations in demand are estimated to have either no effect or cause 100% of the crowding observed. Route/direction/time-period combinations with no estimated crowding contribution are periods for which overall and daily average peak loads remain below seated capacity. The passenger crowding experienced is due to within-period load variability factors or dropped trips. There likely are fluctuations in demand during these periods, however the fluctuations are not large enough to cause the provided frequency to be inadequate to meet travel demand on any given day.

For the majority of periods with crowding attributed to day-to-day fluctuations in demand, it accounts for 100% of the crowding observed. These are route/direction/time-period combinations with only one scheduled trip and for which average loads remain under seated capacity. Crowding occurs when an increase in passenger travel demand

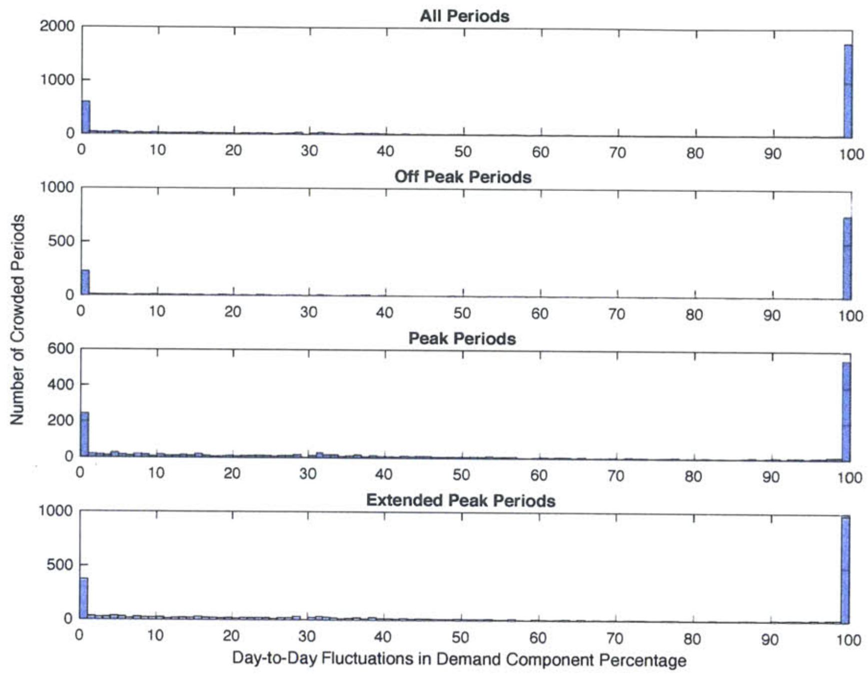


Figure 5-10: Distribution of Day-to-Day Fluctuations in Demand Percentages for Route/Direction/Time-Period Combinations with Observed CPCT. Weekdays September-November 2015.

causes loads on a specific day to exceed seated capacity. These are low frequency routes for which passengers are much more likely to select a specific vehicle trip to ride and therefore headway variations will have little effect.

There does not appear to be much difference between time periods. Peak, extended peak (all time periods where the crowding threshold is 140% of seated capacity), and off-peak periods all show similar distributions for crowded route/direction/time-period combinations.

It is likely that some of the crowding attributed to day-to-day fluctuations in demand during high frequency periods is in reality caused by within-period load variability factors. Daily average loads can be affected if vehicles trips bunch with vehicle trips in adjacent periods. The period in which the more heavily loaded vehicle trip of the pair occurs will likely have a higher average load than its level of passenger travel demand would indicate while the period with the less loaded vehicle trip will likely have lower daily average loads. This causes some days to appear to have higher levels of average demand than actual average levels of demand would indicate.

5.4.4 Within-Period Load Variability Factors

The distribution of within-period load variability factors percentages for route/direction/time-period combinations with more than one scheduled trip is presented in Figure 5-11. Periods with only a single scheduled trip are excluded as the component percentage will always be zero.

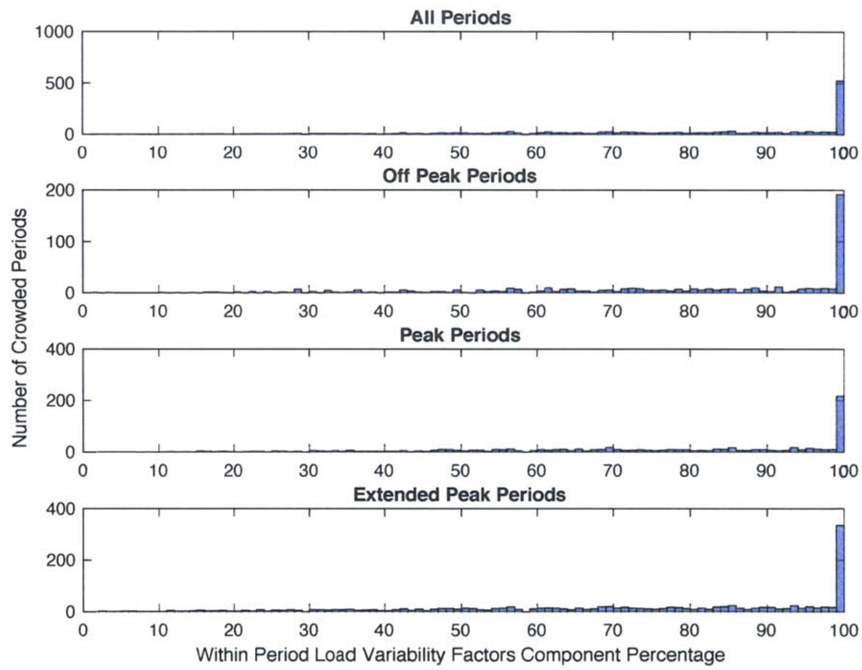


Figure 5-11: Distribution of Within-Period Load Variability Factors Percentages for Route/Direction/Time-Period Combinations with More than one Scheduled Trip and Observed CPCT. Weekdays September-November 2015

This shows that for the majority of periods, within-period load variability factors have a large impact on creating crowded conditions. For many higher frequency route/direction/time-period combinations, within-period load variability factors account for 100% of the crowding observed.

As mentioned in Section 5.1.1, it is likely the load estimation and scaling process over estimates the amount of load variation. Therefore, it is also likely that the contribution of overall variability factors (both day-to-day demand fluctuations and within-period load variability factors) is slightly over emphasized in this analysis. Although even with this slight bias it still appears that within-period load variability factors have contributed significantly to creating many crowded situations.

More research is needed to further the understanding of how within-period load variability factors are partitioned into the uneven headway and within-period demand variability categories though the Route 1 and 9 examples suggest that uneven headways may account for the vast majority of this crowding.

5.4.5 Overall Crowding

This methodology can also be applied across routes and time periods to the entire bus network (excluding the routes mentioned in Table A.1). The total amount of passenger time attributed to each "reason" can be summed and divided by the total amount of seated capacity CPCT observed to calculate systemwide contribution percentages. It is estimated that approximately 17% of the CPCT calculated with scheduled run times for September-November 2015 weekday bus trips could be accounted for by the scheduled frequency; 4% is estimated to be caused by the decrease in frequency resulting from cancelled vehicle trips; 36% is estimated to arise from the fixed schedule not being able to account for day-to-day fluctuations in demand and 44% is estimated to result from factors that cause loads to vary within a route/direction combination on a given day such as uneven headways or varying passenger arrival rates and boarding/alighting distributions.

This was also calculated systemwide for individual time periods. Figure 5-12 highlights how this composition changes throughout the day. Day-to-day demand fluctuations and within-period load variability factors remain the dominant sources for much of the day. Only during the early morning hours is scheduled frequency the primary factor. The relatively high contribution of scheduled frequency during the midday period is due to Route 111. It accounts for a significant portion of total crowding during this period, much of which is caused by scheduled frequency. Figure 5-13 shows these contributions in absolute terms for each 30-minute-time-period.

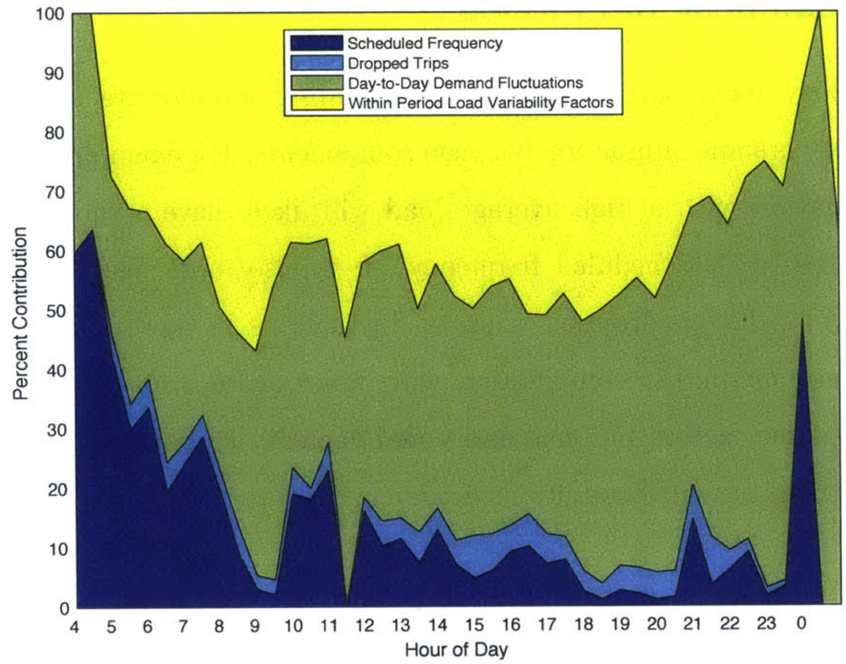


Figure 5-12: Crowding Component Contribution Systemwide by Time Period. Weekdays September-November 2015

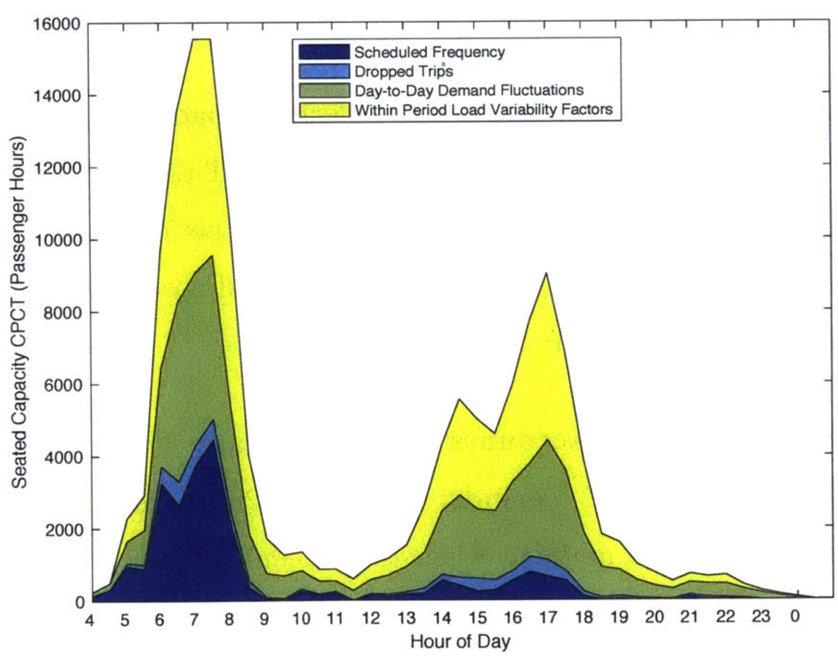


Figure 5-13: Absolute Crowding Component Contribution Systemwide by Time Period. Weekdays September-November 2015

5.4.6 Component Interaction

The methodology above partitions passenger crowding into discrete categories. However, there likely is some interaction between components. For example, a route/direction/time-period combination with a high average load will likely have a significant amount of crowding due to its scheduled frequency. It is also more likely to present additional crowding due to dropped trips or fluctuations in day-to-day demand. A route/direction/time-period combination with lower average loads may still be affected by the same factors but has additional capacity to be able to absorb more passengers without crowded conditions arising.

5.5 Highest Priority Routes

Routes can be classified by their crowding characteristics. This was done for the ten routes with the most 140% of seated capacity CPCT, which are listed in Table 4.2. Four categories emerged:

- **Most Crowding Caused by Within-Period Load Variability Factors:** Routes 1 and 77 have only limited amounts of CPCT that can be attributed to scheduled frequency even during the extended peak periods. Most is attributed to within-period load variability factors with a small portion attributed to day-to-day fluctuations in demand.

- **Scheduled Frequency Accounts for a Significant Amount of Crowding during Distinct "Peak" Periods:** Routes 7, 9, 28, 47, and 66 all have distinct "peak" periods in the morning when the scheduled frequency accounts for a significant amount of crowding. The exact timing of these periods vary between routes. It occurs from 5:30-7:00 for Route 28 while it occurs between 7:30-8:30 for Route 7. Some also have an afternoon peak period. For example, scheduled frequency accounts for a significant amount of crowding on Route 47 between 15:30-17:00.

- **Scheduled Frequency Accounts for a Significant Amount of Morning "Peak" Crowding then Scattered Amounts during the Rest of Day:** Routes 57 and 86 have a distinct morning "peak" period like the routes mentioned in the previous category however; they have scattered periods during the afternoon and evening when the schedule accounts for a significant amount of crowding as well. These afternoon periods are spread out intermittently over multiple hours.

- **Scheduled Frequency Accounts for a Significant Amount of Crowding During the Off-Peak Periods:** Route 111 has a crowding pattern distinct from the other routes. It is estimated that other than between 19:30- 20:00 scheduled frequency has no crowding contribution for outbound vehicle trips. Most crowding is split between day-to-day demand fluctuations and within-period load variability factors. Inbound, however, scheduled frequency accounts for much of the observed crowding for the early morning and midday periods. This contribution is concentrated on the longer of the two variations run during the off-peak periods. During the peak periods, it is estimated that scheduled frequency does not cause any crowding.

5.6 Conclusions

This chapter discusses a methodology to estimate the contribution that different potential sources have in creating crowding conditions. This estimation was done for route/direction/30-minute-time-period combinations enabling an understanding of how the causes of crowding on a route can change throughout the day. Four sources were examined: scheduled frequency, the reduction in frequency due to dropped trips, the fixed schedule not being able to account for day-to-day fluctuations in demand, and factors that cause loads to vary within a period on a given day.

Overall, it was estimated that day-to-day fluctuations in demand and factors causing loads to varying within a period on a given day account for the majority of crowding experienced. These two factors combine for approximately 80% of CPCT observed while scheduled frequency was only estimated to account for 17%. The variability factors contribution is likely a slight over-estimate of their impact as their is some inherent variability in the ODX derived load scaling process. The sourcing methodology may also attribute some crowding that is caused by within-period load variability factors to day-to-day demand fluctuations on high frequency routes if trips scheduled near the boundaries of a period bunch with vehicle trips on adjacent periods. This results in more fluctuations in daily average loads than occur in reality.

These contributions vary between time periods. For many routes, the contributions from different sources change throughout the day. Scheduled frequency is more likely to have larger contributions during the extended peak periods, particularly the morning peak period, on high demand routes as the crowding standards are relaxed.

However, these patterns vary between routes as well. Some routes like Routes 1 and 77 have significant amount of crowding, though almost all of it can be attributed to within-period load variability factors while others have certain periods where the schedule accounts for a significant amount of crowding.

Classification of routes by these patterns can help decision makers determine which interventions are appropriate for each route.

Chapter 6

Crowding Reduction Strategies and Recommended Programs

Identifying the most effective crowding reduction strategy for each situation is important for transit agencies to make the most effective use of available resources. This chapter outlines potential strategies and discusses the impact that each has on the entire passenger experience along with the situations in which they are most appropriate. Then, using the findings of this thesis, crowding reduction programs are suggested for the MBTA given different resource scenarios.

6.1 Potential Crowding Reduction Strategies

There are many ways to improve the passenger crowding experience. Some involve increasing the amount of seated capacity throughput while others make more effective use of capacity in operation. Six categories of interventions are explored. These categories include improving reliability in the form of headway maintenance, improved passenger in-vehicle distribution, increasing scheduled frequency, increasing travel speeds, network redesign, and increasing the capacity of vehicles in use.

The impact on the entire passenger experience is considered when evaluating potential strategies as an intervention that improves the crowding aspect of a passenger's experience may worsen another aspect to the point where the intervention decreases a passenger's overall riding experience. For example, mid point holding of vehicles may reduce the amount of crowding that certain passengers face but increase others' travel time to the point where the overall experience is diminished.

6.1.1 Improving Service Reliability

For cases in which within-period load variability factors have significant impacts, the most effective intervention is likely to be to improve reliability. Maintaining even headways enables more effective use of existing capacity. Average loads will remain the same but there will be much less load variability between adjacent vehicle trips meaning that high intensity crowding conditions will occur on fewer vehicle trips. Much of this unreliability can be attributed to poor on-time (either headway or schedule-based) terminal departure performance. Maltzan (2015) [33] showed that much of the unreliability of two MBTA key routes (Routes 1 and 28) could be attributed to poor terminal departure headway adherence.

Depending on the situation, this may or may not require additional resources. There may be cases when stricter dispatching protocols at terminals may be all that is required. These are cases where the scheduled vehicle "cycle time" is long enough for most operators to leave on schedule but due to various reasons they do not depart on time.

There may also be cases where a large percentage of vehicle trips are not able to depart on time due to insufficient half cycle times in which case the optimal solution may be to lengthen cycle times. This would result in more vehicle hours required to maintain the current frequency. However, even in cases where all buses are late and additional vehicle resources are unavailable, an even headway terminal dispatching strategy would be largely effective.

A final case may be that the half cycle times are long enough to complete the previous vehicle trip but operators do not have enough time to get ready for the successive trip. Here an additional operator could be hired so that each operator could "drop back" after they finish a vehicle trip. The additional operator waiting at the terminal could operate the successive trip scheduled for that vehicle. The first driver would then wait at the terminal for the next vehicle to arrival and would operate its successive trip. This would promote on time terminal departure while still ensuring that operators have rest time in between trips. It would not require additional vehicles, only additional operator hours.

Improving service reliability also has the added benefit of reducing a passenger's expected wait time. Since more passengers are likely to board vehicle trips with longer headways maintaining more even headways also decreases the expected passenger wait time.

6.1.2 Passenger In-Vehicle Distribution

Although most of the analysis in the thesis focuses on using vehicle loads as a measure of crowding intensity, passenger in-vehicle distribution greatly effects how passengers perceive crowded conditions. Passengers are likely to feel more crowded on vehicles where riders are clustered together compared to instances where passengers are uniformly distributed throughout the vehicle. Therefore passengers could perceive very different levels of crowding given the same vehicle load depending how passengers are distributed.

Often times this distribution can be improved by "nudging" passengers to change their behavior. Some transit agencies have produced marketing campaigns to make passengers more aware of the behaviors that decrease effective use of capacity. Figures 6-1 and 6-2 highlight examples from Translink in Vancouver and the Metropolitan Transportation Authority in New York. Making passengers more aware of the impact that certain behaviors have on other riders may encourage behavior change and reduce subjective passenger crowding all without making any operational changes. This may also increase the effective capacity of vehicles.

Figure 6-1: Example of Rider Etiquette Advertisement on Translink Buses in Vancouver, Canada that encourages passengers to move towards the back of the bus. [35]

Figure 6-2: Example of Rider Etiquette Advertisement on Metropolitan Transportation Authority in New York that persuades passengers to be mindful of how much space they are taking up while seated. [36]

6.1.3 Scheduled Frequency

In all situations an increase in frequency will reduce passenger crowding by increasing the seated capacity throughput of a route. If passenger arrival rates are relatively constant throughout a period, this will enable customers to be distributed among more vehicles and therefore decrease average loads. It also benefits passengers by reducing the average amount of time required to wait for a vehicle.

This strategy does require more resources if all else is equal. Unless there is significant slack in the schedule, additional vehicle hours will be required to provide a higher level of service.

Increasing frequencies may help reduce load variability as well by decreasing the impact that fluctuations in passenger arrival rates and alighting distributions have on run times (e.g., a route with an average peak load of seated capacity may be better able to maintain headways given fluctuating passenger arrival rates compared to a route with an average peak load of 140% of seated capacity.). This could reduce the downstream effect of such fluctuations. However, on some routes headways may fluctuate for reasons other than varying loads (i.e., many traffic lights, traffic congestion, low on-time terminal departure rates). In these cases, increasing frequencies will likely not improve the effective use of capacity. The majority of passengers may still face crowded conditions if even headways are unable to be maintained.

6.1.4 Speed

There are also instances where improving travel speeds can significantly reduce the length of time in which passengers spend in crowded conditions. All else equal (i.e., frequencies and headway variability remain the same) this will not decrease the intensity of crowding but will improve the passenger crowding experience. However, increasing travel speeds will likely enable a reduction in cycle times. This can reduce the number of vehicles required to provide the current level of service. This savings could theoretically be used to increase frequencies on either the current route or others within the system, lowering average loads. This strategy's effectiveness will often depend on saving enough cycle time to save a whole bus either on a single route or through an interlining schedule strategy with one or more routes.

This could be achieved through either transit-supported roadway strategies or operational efficiencies. Roadway strategies are measures that increase travel speeds between stops by modifying roadway infrastructure. Possible interventions include bus lanes on routes with high levels of traffic congestion or transit signal priority on routes with many signaled intersections. The impact on all road users should be considered. While street right-of-ways are usually not directly controlled by transit agencies, cooperation between municipalities and transit agencies can foster effective use of street space.

Operational efficiency interventions are strategies that reduce dwell times. This includes offboard fare payment and reduction of "slow transaction" fare media usage. Offboard fare payment requires passengers to pay their fare before boarding the vehicle either at a faregate or through a ticket validator. Then when a vehicle arrives, passengers can board through all doors without interacting with a fare collection device. This could greatly improve dwell times especially on high volume stops where passengers are forced to wait in line to board a vehicle and for which the boarding process can take a significant amount of time.

The type of fare media used also makes a significant difference in amount of time required for passengers to board. Cash and magnetic strip card transactions take significantly more time to process than contactless smart cards. Fare policies that encourage greater smartcard use while still ensuring access for all could significantly improve dwell times especially in locations with a high percentage of "slow" transactions.

These interventions will not only shorten the duration of crowding that passengers experience, it will also shorten their overall expected travel time while decreasing their travel time variability.

6.1.5 Network Design

It is also possible to reduce crowding by making changes to the bus route network. Through the development of new routes or increased frequencies on existing routes, passengers may be encouraged to switch from more crowded routes. For example, passengers traveling from one area on the periphery of a transit agency's service area to another may have to transfer in the central business district (CBD). This may worsen crowding by adding to ridership that is destined directly for the CBD. These passengers may be better served by a crosstown route that bypasses downtown all together. In some cases, these crosstown routes do not exist and a new route would need to be established. In others, they operate at insufficient frequencies to draw passengers from the arterial routes so that an increase in crosstown route frequency could be effective to not only increase productivity on the crosstown route but also to reduce crowding on the arterials.

Since ODX allows for the analysis of passenger linked trips, opportunities for network redesign can be identified. Vanderwaart (2016) [32] develops a framework to use ODX to analyze the impact of bus service changes in the MBTA context.

These interventions would not only reduce crowding on the urban core of a network but could also greatly improve passenger travel times and access across the region.

6.1.6 Increase Vehicle Capacity

Increasing the capacity of vehicles that operate on a route increases the seated capacity throughput of a route without any changes to the schedule. For example, a typical 40 foot bus contains approximately 40 seats while a 60 foot articulated bus contains approximately 60 seats. (These can vary depending on the interior configuration.) Introducing articulated buses on a route that currently operates with 40 foot buses can increase seated capacity throughput by almost 50%. There is also a standing capacity increase as there is more floor space on the larger vehicles.

This is an effective strategy for routes with high demand for the entire day as there would be no low demand period where significantly more capacity is provided than is needed as service delivery policies require minimum frequencies.

It is a cost effective strategy for transit agencies as more passengers can be served without dramatically increasing operating costs. If frequencies remain constant, no additional operators would be needed. There is likely a slight increase in the marginal vehicle operating costs of a 60 foot bus compared to a 40 foot bus but this is likely significantly lower than the combined costs of increasing the number of operators and vehicles required in increasing frequencies of 40 foot buses.

With larger vehicles, passengers are likely to experience less crowding, however, they will face longer wait times compared to the same route with increased frequencies using smaller vehicles.

6.2 Potential MBTA Crowding Reduction Scenarios

The following section proposes recommendations for crowding reduction on the MBTA bus system given different levels of resource availability including no additional operating budget, no additional peak vehicles but with increases in operating budget, minimal additional vehicles, and sufficient vehicles to reduce the majority of crowding systemwide.

6.2.1 No Additional Operating Budget

If there are no additional operating resources available, recommended strategies focus on improving effective use of existing capacity without adding any new operating costs. These include:

- **Implement strategies to improve on-time terminal departure performance especially on high frequency routes.** Within-period load variability factors account for a significant amount of crowding on high frequency routes and much of this is likely to come from uneven headways. Decreasing headway variability will likely reduce crowding significantly on many route/direction/time-period combinations. In many cases this may be achieved by improved dispatching at terminal locations.

- **Create a passenger etiquette marketing campaign to encourage more effective use of in-vehicle space.** A marketing campaign similar to Courtesy Critters could be introduced to raise awareness of passenger behaviors that limit effective use of space within vehicles while promoting desired behaviors. Topics could include moving away from the doors when standing, making open seats available to others, placing backpacks between legs when vehicles become crowded, among others. These could be used on both the rail and bus networks as similar issues arise in both. While this would not necessarily decrease vehicle loads, it may significantly improve the subjective crowding that passengers experience and increase effective capacity.

- **Work with municipalities to identify situations where transit-supported roadway strategies could be implemented.** Projects similar to the transit-signal priority measures on Commonwealth Avenue and bus lanes on Washington Street in Boston's South End could be implemented throughout the network in places where many routes operate.The BU Bridge, Longwood Medical Area, Malcolm X Boulevard and Columbus Avenue between Dudley and Ruggles Stations and Washington Street between Forest Hills Station and Roslindale Square may be effective locations where further corridor analysis should be conducted. By targeting a small number of specific problems, it may be possible to get more cooperation from municipal traffic departments.

- **Implement operational efficiency measures to reduce dwell times.** Promoting greater use of CharlieCards instead of cash or CharlieTicket transactions should help reduce the duration for which passengers experience crowded conditions. This could also improve farebox interaction rates.

  Offboard fare payment could also be promoted especially at high volume mid-point stations such as Dudley and Ruggles Stations, among others, by increasing validation use. This will greatly improve dwell times at these locations by enabling all door boardings and eliminating slow onboard fare payment. These operational efficiencies may even facilitate higher frequencies by lowering the cycle times required to operate a given route.

- **Look for opportunities to improve crowding through network redesign.** Opportunities where changes in network design could improve crowding conditions could be explored. Vanderwaart (2016) [32] provides a framework that uses ODX to evaluate the effect that changes in network design have on individual routes. For example, If route A is modified and improved, X number of passengers are likely to switch from using Routes B and C since Route A provides a better service for traveling between their desired origins and destinations. Therefore crowding conditions are likely to improve on Routes B and C.) She uses this process to describe an extension of Route 29 to Ruggles Station, Longwood Medical Area, and Brookline Village as a intervention to improve access to Longwood and decrease crowding on the D Branch of the Green Line and Route 28. The impact that service changes have on crowding of specific routes could be evaluated. A memo in Appendix B shows another example application of ODX in this manner. It estimates the effect that service changes in the Route 47 corridor would have on existing riders.

6.2.2 No Additional Peak Vehicles but with Increases in Operating Budget

If there are no additional peak vehicles available for service but small increases in the operating budget are possible, strategies that make more effective use of existing capacity during the peak periods as well as increasing capacity during the off-peaks should be utilized. The recommendations from the no additional operating budget scenario should be pursued in addition to the following:

- **Implement "drop back" operator scheduling strategies to improve on-time terminal departure performance, especially on high frequency routes for which operators routinely do not have enough time to depart on-time.** Within-period load variability factors account for a significant amount of crowding on high frequency routes and much of this is likely to come from uneven headways. Decreasing headway variability will likely reduce crowding significantly on many route/direction/time-period combinations. If operators routinely do not have enough time to depart on-time for their subsequent trip, a "drop back" operator scheduling strategy (where each subsequent operator at a given terminal gets off the arriving bus and the previous arriving operator takes over the following trip with that vehicle, thus giving all operators a headway long break and keeping vehicles in service (theoretically) on a continuous basis.) can be introduced with no additional vehicle costs.

- **Increase scheduled frequency for off-peak route/time-periods with high amounts of 140% seated capacity CPCT and on which scheduled**

frequency accounts for a significant portion of crowding. During the off-peak periods bus fleet size constraints are less prevalent. Therefore, scheduled frequency can be increased without introducing additional vehicles into the network. Route/time-periods with a significant amount of crowding due to scheduled frequency will see the most benefit from increases in frequency. Example routes include the 23, 28, and 111 during the early morning period and the 111 during the 2 PM hour.

6.2.3 Minimal Additional Vehicles

If some additional vehicles are available to be introduced into the bus network more interventions could be implemented including expanding throughput capacity by increasing scheduled frequency and lengthening cycle times to improve reliability. Using the data presented previously, it is roughly estimated that to reduce 30-minute average peak loads to under 140% of seated capacity systemwide approximately 10 additional vehicles would be needed during the morning peak period. To eliminate all scheduled frequency crowding (bring average peak loads under seated capacity) on the top ten routes with the most overall 140% of seated capacity CPCT during the 7:30-8:00 time period, approximately 25 additional vehicles would be needed. The recommendations from the previous scenarios would still apply along with the following new recommendations:

- **Increase scheduled frequency for route/time-period combinations with high amounts of 140% seated capacity CPCT and on which scheduled frequency accounts for a significant portion of crowding.** These are situations in which increases in scheduled frequency are likely to have the largest impact. The highest priority combinations include Routes 7, 9, 28, 66, and 86 during the morning peak, 47 during the afternoon peak and 111 during the 2 PM hour. The exact timing within each period varies by route with Routes 28, 66, and 86 requiring additional vehicles on earlier morning trips than Routes 7 and 9.

- **Increase scheduled frequency on routes for which passengers are likely to have long durations of crowding.** These are routes for which ridership patterns and route design result in passengers experiencing onerous crowding when conditions arise. Therefore extra attention should be made to ensure that crowding conditions do not arise. Example routes include the 47 and 111.

- **Increase scheduled cycle times on routes for which they are currently inadequate.** For routes on which on-time terminal departure performance is hampered by vehicles not arriving before their next scheduled trip, better dispatching may have limited impact. In these cases, lengthening scheduled cycle times could significantly improve on-time terminal departure performance and reduce crowding assuming improved on-time departure dispatching procedures and performance.

6.2.4 Numerous Additional Vehicles

If a significant number of peak and off-peak vehicles are available, then increases in scheduled frequency can be implemented on routes for which there will be lesser impacts on crowding but which may also serve to attract new riders. This includes both routes on which lesser amounts of total CPCT was observed and routes for which there is significant crowding but most of it is attributed to either day-to-day fluctuations in demand or within-period load variability factors. The recommendations from the previous sections would still apply in addition to the following:

- **Increase scheduled frequency on highly crowded route/time-period combinations for which variability factors account for the majority of crowding.** The lower the average peak loads the less likely crowded conditions are to arise even on high variability routes. Therefore, if there is still significant crowding due to variability factors after implementing the recommendations from Sections 6.2.1 and 6.2.2 , increasing scheduled frequency will likely improve conditions though perhaps not as efficiently if reliability issues remain. This includes Route 111 during the morning peak and Route 1 during both peak periods.

- **Increase scheduled frequency on lightly crowded route/time-period combinations for which scheduled frequency accounts for a significant amount of crowding.** These are routes for which either relatively few passengers experience crowded conditions and/or they experience these conditions for limited amounts of time. An increase in frequency will likely have a significant impact for these riders though they account for small percentage of systemwide crowding. Examples include Routes 238 and 430 during the morning peak period.

6.3 Summary for Highest Priority Routes

Recommended crowding reduction programs for the routes with the highest amounts of 140% of seated capacity CPCT along with other select routes are shown in Table 6.1. These only include increases in reliability and scheduled frequency recommendations. Other strategies listed in Section 6.1 should also be implemented though additional analysis is needed to determine the appropriate situations and in which form they should be implemented.

If there is no additional operating budget available, measures that would increase reliability are recommended for all routes except Route 430 as it is low frequency the entire day. If there is additional operating budget but no additional peak period vehicles available, increasing frequency for crowded off-peak periods is recommended for Routes 23, 28, 47, and 111. If there are minimal additional vehicles available, reliability measures are recommended in addition to increases in scheduled frequency during the highest impact periods where crowding is severe and a significant portion of crowding can be attributed to scheduled frequency. For most routes, this occurs during the morning peak period. In the many additional vehicles available scenario, the same recommendations from the previous scenarios apply although additional periods where scheduled frequencies should be increased are also included for certain routes.

Route	No Additional Operating Budget	Additional Actions with No Additional Peak Vehicles but with Increases in Operating Budget	Additional Actions with Minimal Additional Vehicles	Additional Actions with Numerous Additional Vehicles
111	Reliability Measures	Increase Frequency During Early Morning and 2-3PM Periods	Increase Frequency During Morning Peak	
1	Reliability Measures			Increase Frequency During Both Peak Periods
9	Reliability Measures		Increase Frequency During Morning Peak	
7	Reliability Measures		Increase Frequency During Morning Peak	
66	Reliability Measures		Increase Frequency During Morning Peak	
57	Reliability Measures		Increase Frequency During Morning Peak	Increase Frequency During Afternoon Peak Period
86	Reliability Measures		Increase Frequency During Morning Peak	Increase Frequency During Afternoon Peak Period
47 [1]	Reliability Measures	Increase Frequency During Early Afternoon Period	Increase Frequency During Morning Peak	
28	Reliability Measures	Increase Frequency During Early Morning Period		
77	Reliability Measures			Increase Frequency During Morning Peak
23	Reliability Measures	Increase Frequency During Early Morning Period		
238	Reliability Measures			Increase Frequency During Morning Peak
430				Increase Frequency During Morning Peak

Table 6.1: Recommended Programs for Routes with Most 140% Seated Capacity CPCT and Other Select Routes

[1] Also see a short-turn network design strategy that is detailed in Appendix B.

6.4 Conclusions

Different situations call for different crowding reduction interventions. This chapter outlines potential strategies and discusses the situations for which they are best suited. Some strategies make more effective use of existing seated capacity throughput while others increase capacity.

These strategies are then applied to the MBTA bus network under four scenarios: no additional operating budget available, no additional peak vehicles available but with increases in operating budget, limited number of vehicles available and many additional vehicles available. In the no additional operating budget available scenario, strategies focus on making more effective use of existing capacity using no additional budget resources. In the increase in operating budget but no additional vehicles scenario, reliability measures are emphasized as well as increasing capacity during the off-peak periods when fleet size constraints are less restrictive. In the limited number of vehicles available scenario, making effective use of existing capacity is still emphasized though capacity increases in high impact situations is also proposed. Finally, in the many additional vehicles scenario, the recommendations from the previous two scenarios continue though resources are now available to address lower impact situations. These recommendations recognize that ideal strategies are contingent upon both the individual micro situations (e.g., route/time-period combinations) and the overall context in which they arise (e.g., amount of resources available systemwide).

Since these techniques make it feasible to measure crowding more effectively and at a reasonable cost, it is recommended that crowding be routinely monitored to understand how ridership growth may be affecting crowding and conversely how excessive crowding may be discouraging ridership growth.

Chapter 7

Summary and Recommendations for Future Research

This chapter summarizes the overall findings of this thesis. Then recommendations for the transit industry are proposed based on these findings. Finally, areas of future research in the area of crowding analysis are suggested.

7.1 Summary of Findings

This thesis has four primary areas of focus: adapting and implementing a scaling process to estimate trip level passenger flows and vehicle loads through inferred passenger origins and destinations using the MBTA bus network as a case study; creating passenger centric crowding metrics to enable prioritization of resources to times and routes on which the most crowding occurs; developing a methodology to estimate the contribution that different potential sources have in creating crowded situations; and identifying ways to use ODX based analysis to design effective strategies to reduce crowding under different resource scenarios. The primary findings for each area are discussed in the following sections.

7.1.1 Trip Level ODX Scaling

A two step scaling process was adapted from Wang (2010) [25] to scale ODX inferred passenger flows. These flows were used to create trip level load profiles and origin destination matrices. Comparisons to APC derived load profiles indicate relatively accurate estimates though the scaling process does introduce some additional load variability.

Although there are routes on which this process does not work well (e.g., routes on which fares are paid at fare gates instead of onboard vehicles) this process enables transit agencies with high penetration rates of reusable individually marked fare media and granular AVL information in combination with a sample of vehicle trips with APC load information to develop more complete coverage of load profile information than can be accomplished with APC sample trips alone. In the MBTA case, on routes which ODX and the scaling process works well, load profile coverage expanded from 13% of weekday vehicle trips operated from September-November 2015 with APC estimated loads to almost 90% through the ODX scaling process. The operated trips without load profiles appear to be relatively uniformly distributed though there were lower coverage rates during the overnight period and on some very low frequency routes.

7.1.2 Passenger Centric Crowding Estimation

The load and passenger flow estimates were then used to create metrics to evaluate different aspects of crowding from the passenger perspective. These metrics consider the intensity of crowding, its duration, and the number of unique passenger trips on which crowding conditions were experienced.

Cumulative Passenger Crowding Time (CPCT) was developed to attempt to consolidate all of these aspects into a single metric. It was used to evaluate crowding in the MBTA bus network and revealed that much of the crowding observed occurs in a concentrated portion of the network both temporally and spatially. Most occurs on high frequency routes during the peak periods. This appears to be a byproduct of both creating schedules based on a average peak loads and reduced travel speeds due to traffic congestion.

Analysis of individual aspects of crowding was also completed which revealed that passengers experience crowding in different ways on different routes. It was found that passengers on certain routes are more likely to experience crowded conditions for extended durations due to route design and ridership patterns. This duration also varies throughout the day as loads and travel speeds fluctuate.

The number of unique passenger trips on which crowded conditions were experienced was strongly correlated with the number of vehicle trips on which the crowding threshold was reached though there was some deviation when vehicle seated capacity and passenger turnover rates were considered.

7.1.3 Crowding Source Methodology

CPCT was then used to estimate the contribution of various factors towards creating the crowding conditions observed. Four potential factors were considered and analyzed. All MBTA bus trips were clustered by route, direction, and 30 minute scheduled start period to enable both comparisons across time periods on a given route and between routes.

Although it is important to evaluate each route individually, this analysis was used to evaluate the sources of crowding systemwide. It was estimated that for most routes, the schedule frequency accounts for only a small portion of the crowding observed. Even during peak periods, when the crowding standard allows for some crowding to be built into the schedule, almost 90% of route/direction/30-minute-time-period combinations would not be expected to have any crowding if scheduled headways were perfectly maintained and passenger arrival rates and boarding/alighting distributions were constant across the 30-minute-period.

The decrease in frequency resulting from dropped trips appears to have a minor effect and for many combinations the effect is estimated to be zero. This does not account for increases in headway variability that might result from cancelled trips.

For many route/direction/30-minute-time-period combinations, day-to-day fluctuations in demand either accounted for none of the crowding observed or almost all of the crowding observed. On low frequency periods where neither scheduled frequency or dropped trips contribute to crowding and only a single vehicle trip is scheduled, day-to-day fluctuations in demand are estimated to account for 100% of crowding as passengers are likely to select a specific vehicle trip to ride and therefore within-period load variability factors are likely to have limited effects. Therefore, any crowding that occurs is due to an infrequent surge in passenger travel demand on a particular day. Periods in which daily average peak loads never exceed seated capacity always have high enough scheduled frequency to meet demand on any specific day. Therefore, within-period load variability factors cause all crowding conditions while day-to-day fluctuations in demand are not estimated to have any contribution.

Within-period load variability factors were estimated to account for a significant amount of crowding for higher frequency routes, of which the vast majority appears to be caused by uneven headways although more research on this topic is recommended (See Sections 5.2.3 and 7.3.3). It appears that if even headways were able to be maintained much of the observed crowding would be eliminated. For many periods, it accounts for nearly all of the crowding observed. For low frequency route/direction/30-minute-time-period combinations on which only a single vehicle trip is scheduled, this "reason" is estimated to be zero as passengers are more likely to select a particular vehicle trip to board and for which varying headways are less likely to have a significant impact on creating crowded situations. All crowding not attributed to scheduled frequency or dropped trips for these combinations is considered to be caused by the fixed schedule not accounting for day-to-day fluctuations in demand. It is likely that boundary issues cause some crowding attributed to day-to-day fluctuations in demand in high frequency periods to be in reality caused by within-period load variability factors (i.e., if there is bunching between vehicle trips in successive periods, with one bus scheduled to depart in one period and the other in the successive period, daily average loads for a specific 30-minute period could fluctuate more than overall day-to-day fluctuations in passenger travel demand for hourly or longer periods).

Overall, across all combinations, scheduled frequency was estimated to cause approximately 17% of seated capacity CPCT calculated with scheduled run times, the decrease in frequency resulting from dropped trips accounts for approximately 4%, the fixed schedule not accounting for day-to-day fluctuations in demand approximately 36% and within-period variability factors approximately 44%. While the ODX scaling process is likely to introduce a small level of additional variability into these estimates, it is likely that fluctuations in demand and within-period variability factors account for the vast majority of crowding observed.

7.1.4 Crowding Reduction Strategies

Finally, using both the passenger centric crowding metrics and the crowding source estimates, crowding reduction strategies were proposed for the MBTA bus network systemwide under different resource scenarios. The most effective mitigation strategies for each type of crowding situation were identified.

Four scenarios were analyzed: No additional operating budget, no additional peak vehicles available but with increases in operating budget, minimal additional vehicles available and many additional vehicles available. No additional operating budget emphasizes no additional operating costs interventions that improve the effective use of existing seated capacity throughput, such as improving headway maintenance and passenger in-vehicle distribution. For the no additional peak vehicles available but with increases in operating budget scenario, strategies focus on increasing the effective use of existing seated capacity throughput and increasing capacity in high impact off-peak route/time-periods. For the minimal vehicles available scenario, the recommendations from the previous scenario can be augmented by increasing seated capacity throughput in high impact situations. In the many additional vehicles available scenario, increasing seated capacity throughput in lower impact situations are added to the recommendations from the previous scenarios.

7.2 Recommendations for the Transit Industry

Three primary recommendations for the transit industry are proposed as a result of the findings of this research: increase the ability to granularly measure passenger crowding, increase the use of cumulative passenger centric crowding metrics, and explore additional crowding reduction strategies besides increasing scheduled frequency.

7.2.1 Granular Crowding Measurement

Transit agencies should make efforts to expand their capacity to granularly measure passenger crowding. There is a lot to be gained by regularly monitoring crowding and carefully estimating its source. Crowding caused by uneven headways may be under detected using the traditional vehicle load metrics calculated from a sample of vehicle trips.

APC systems should be specified on all new bus orders when possible. This will eventually lead to full implementation throughout a bus fleet without costly retrofits. APC systems do not only provide load estimates but also stop level run and dwell times.

Agencies that have high penetration rates of reusable identified fare media and granular AVL information should also implement the ODX inference algorithm for their system. If they have less than complete coverage of APC equipped vehicles, it can be used to increase the number of vehicle trips with reliable load estimates and provide trip level passenger flow estimates.

Agencies that do not have high penetration rates of reusable identified fare media should consider implementing such systems because of the improvement in both the passenger experience and the passenger information they provide.

Even if agencies have 100% coverage with APC equipped buses ODX can provide more nuanced crowding monitoring by enabling crowding measurement to be done on the passenger level (e.g., estimating the distribution of passenger crowding durations on a specific route). It also facilitates passenger linked trip analysis to identify network design changes that can reduce crowding conditions and improve access for its customers.

7.2.2 Cumulative Passenger Centric Crowding Metrics

With this complete coverage of trip load and passenger OD estimates, transit agencies should move towards passenger based cumulative metrics. Average metrics such as average peak load can mask the true amount of crowding that passengers experience. Cumulative metrics consider the experience of every instance measured (e.g., vehicle or passenger trip) and therefore can provide a better understanding of a crowding situation where is significant amount of variation.

Passenger centric metrics can better represent the impact that crowding has on passengers. Vehicle based measures (e.g, percentage of vehicle trips with peak loads over a crowding threshold) can mask the passenger effect since many more passengers travel in crowded vehicles compared to less crowded vehicles. A higher percentage of passengers are likely to experience crowded conditions than the percentage of vehicle trips on which crowded conditions arise.

7.2.3 Additional Crowding Reduction Strategies

This research shows that much of the crowding experienced on the MBTA bus network is not due directly to scheduled frequency. Therefore, increasing scheduled frequency may not always be the most effective crowding reduction solution. Transit agencies should use all available interventions when introducing crowding reduction programs. Maintaining even headways on unreliable routes may greatly reduce crowding and improve the passenger experience without increasing scheduled frequency or operating resources. While increased frequency in addition to maintaining even headways should be pursued as a goal to stay ahead of ridership growth and encourage additional ridership, agencies that are resource constrained can still effectively reduce crowding levels and improve passenger comfort using the strategies out lined in Chapter 6.

7.3 Future Research

This thesis focuses on route level in-vehicle passenger crowding. There are additional aspects of passenger crowding that should be studied as well. These include analyzing crowding on a corridor or OD perspective as well exploring how passengers experience crowding before they even board a vehicle.

7.3.1 Corridor and Segment Analysis

In dense bus networks, multiple routes often operate on a shared corridor. Many passengers boarding along these stretches have multiple options of routes to take. Therefore crowding conditions across all routes of a corridor combined could be evaluated instead of isolating each route. This can help inform resource allocation decisions across all routes within a corridor to provide the best level of service for all passengers of the associated routes.

This type of analysis often requires a spatial analysis of where the most passenger crowding time is spent. This could help agencies determine the most effective locations for transit-supported roadway improvements or short-turn route variations to be introduced.

7.3.2 Left Behind Analysis

Passengers on crowded routes not only experience less comfortable conditions within vehicles but also face a higher probability of being denied boarding if an arriving vehicle does not have sufficient available space. Very high vehicle loads can be used as a proxy measure to determine when and where denied boardings may occur but it is difficult to determine exactly when an individual is unable to board. With advances in software to analyze video and increasing use of surveillance cameras (including video of the boarding and alighting areas outside of buses), a systemwide analysis could be done to identify when and where passengers are denied boarding. This issue cannot be analyzed through ACDS developed vehicle loads alone. In addition, video analysis may also help determine the effective capacity of vehicles and factors that may cause it to vary.

7.3.3 Role of Varying Passenger Arrival Rates in Crowding

For this analysis, all factors that cause within-period load variation are generally considered at the same time. These factors include varying passenger arrival rates and boarding/alighting distributions as well as the perhaps more prevalent condition of varying headways. Transit agencies have limited control over the randomness of passenger travel behavior. Therefore, even if service is perfectly delivered, loads may still vary between vehicle trips. A simulation may help to gain a better understanding of how much load variability is caused by uneven headways compared to variations in passenger travel behaviors. ODX derived load and passenger flow estimates provide both passenger arrival rates between vehicle arrivals and alighting distributions. Simulation analysis can then reassign passengers to vehicles based on the scheduled arrival times. Crowding metrics could be calculated and compared to the crowding expected with daily average loads to estimate how much crowding is due to uneven headways and how much is due to variations in passenger travel behavior.

7.3.4 Effect of Crowding on Route Choice

Combining the passenger linked trip information that ODX provides with crowding information allows for a more refined analysis of passenger path choice. For passengers with multiple possible routes and/or paths on these routes, it be desirable to measure the relative "disutilities" of different components (e.g, overall travel time, number of transfers, expected level of crowding, wait time, among others) of the riding experience in path selection. This would allow transit agencies to identify components that are most important to their customers, which in turn enables better prioritization of service changes that improve the characteristics of their transit trip.

Appendix A

Tables

Routes Excluded From Crowding Analysis

Route	Reason
71	Speciality Vehicle (trolley bus)
73	Speciality Vehicle (trolley bus)
431	Low Ridership (Unable to Construct Seed Matrices)
9701	School Specific Route
9702	School Specific Route
9703	School Specific Route
Slver Line 1 (SL1)	Offboard Fare Payment
Silver Line 2 (SL2)	Offboard Fare Payment
Silver Line Shuttle (SLW)	Offboard Fare Payment
Government Center Shuttle (608)	Free Service
Work As Directed	Free Service
Run As Directed	Free Service

Table A.1: Routes Excluded From Crowding Analysis

Appendix B

Route 47 Memo

To: MBTA Staff
From: Chris Southwick
Date: December 8, 2015
Re: ODx Evaluation of Route 47

Route 47 is a popular crosstown route serving Central Square, Cambridge; Longwood and Boston Medical Centers; and Dudley, Ruggles, and Broadway Stations. It is the 24th most popular weekday bus route with an average weekday ridership of 5,036 according to the 2014 edition of the Blue Book. ODx (origin destination inference) was used to analyze both current (Spring 2014) use and the impact of potential route modifications.

Current Use

Average load profiles and origin destination matrices were derived from scaled up ODx for each weekday period and direction for March, April, and May 2014 to gain a better understanding of current ridership and crowding patterns.

There are likely many trips with significant crowding between Central Square and Longwood Medical Area. Most crowding occurs on inbound AM Peak and outbound PM Peak trips with average peak loads of 46 and 43 occurring as buses cross the BU Bridge. A trip level load profile methodology is currently in development that could better describe the nature of crowding on these trips.

There are also many passengers boarding at the Massachusetts Avenue @ Pearl Street stop especially on inbound AM Peak trips with an average of 34 passenger boardings there per trip.

Finally, there appears to be differing levels of demand between the Central Square-Ruggles and Ruggles-Broadway segments. For example, inbound AM Peak trip average loads never exceed 10 passengers after the vehicle passes through Ruggles Station. On outbound AM Peak trips, running at headways of 21 minutes (more than double that of inbound trips), average peak load does not exceed 20 passengers.

Below are average load profiles of both inbound and outbound trips during the AM Peak.

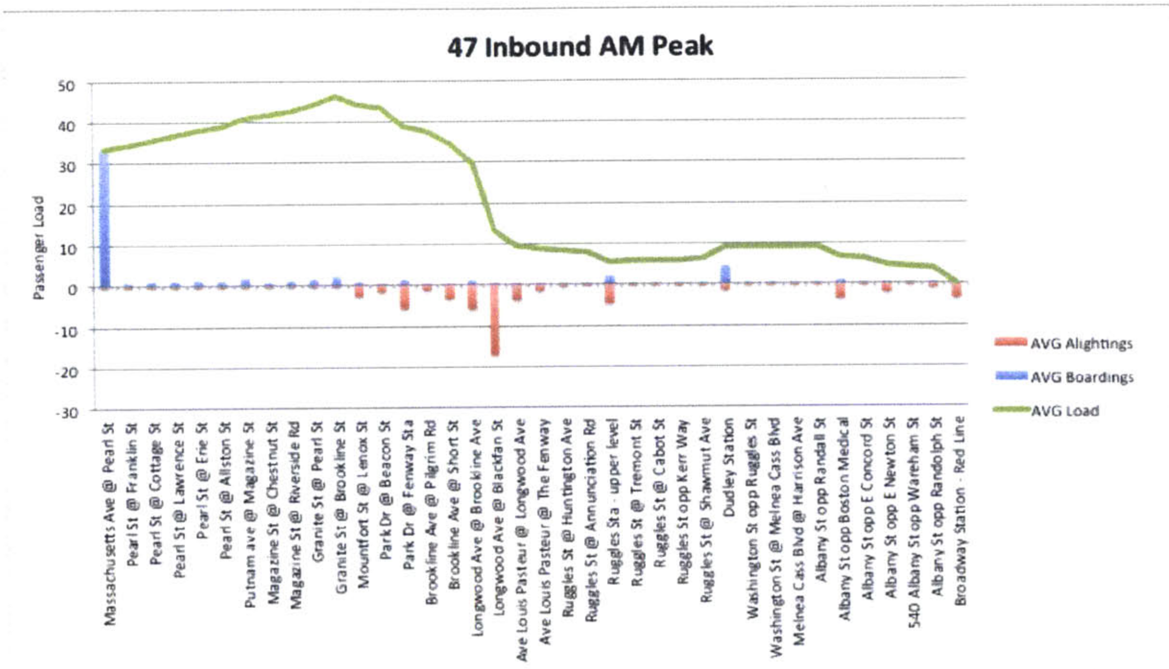


Figure 1: Average Loads for AM Peak inbound 47 trips during March, April, and May 2014.

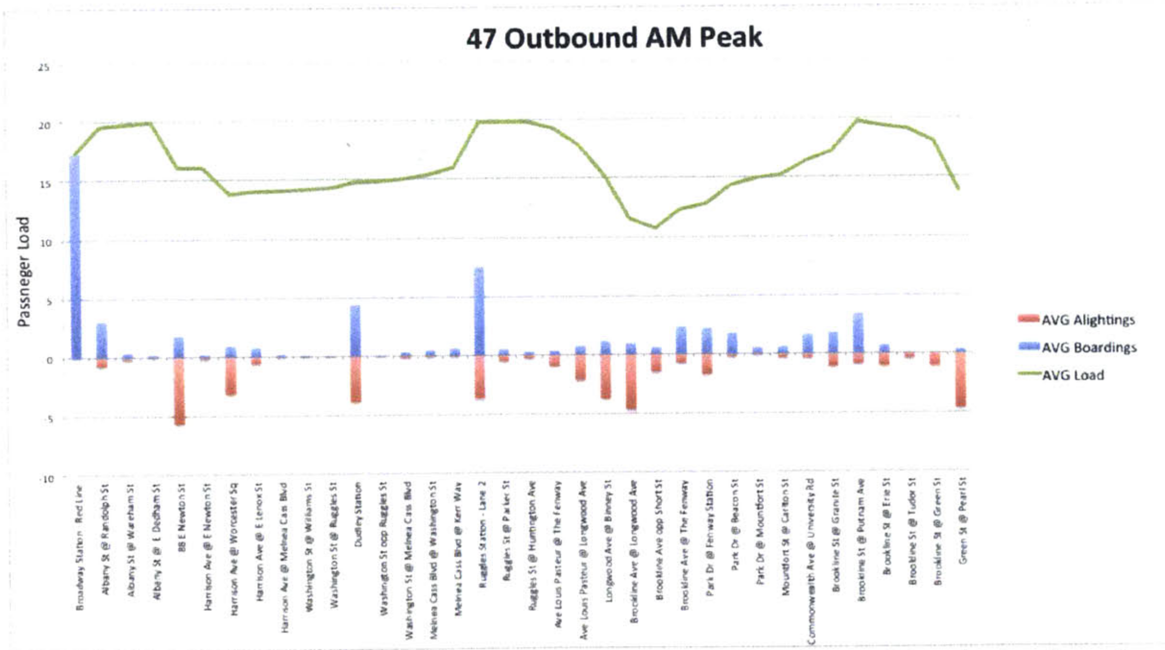


Figure 2: Average Loads for AM Peak outbound 47 trips during March, April, and May 2014.

Next, ODx was used to develop origin destination matrices for each period and direction to understand passenger travel patterns. It showed very few people riding from end to end and indicated that there are two separate groups of passengers who use the route: passengers travelling between Central Square and Ruggles Station; and passengers travelling between Longwood Medical Area and Broadway Station.

47 Inbound AM Peak — Origin \ Destination	Central Square	Upper Cambridgeport	Lower Cambridgeport	Kenmore	Longwood	Ruggles Station	Ruggles Street	Dudley Station	Washington St/Melnea Cass Blvd	Albany Street	Broadway Station	Totals
Central Square		0.4	0.5	5.8	25.0	1.4	0.0	0.1	0.0	0.3	0.0	33.6
Upper Cambridgeport	0.0	0.1	0.1	1.0	3.6	0.7	0.0	0.1	0.0	0.2	0.0	5.9
Lower Cambridgeport			0.1	3.0	3.7	0.7	0.0	0.1	0.0	0.1	0.0	7.8
Kenmore				0.3	1.3	0.6	0.0	0.0	0.0	0.5	0.0	2.8
Longwood					0.3	1.2	0.1	0.6	0.1	0.8	0.3	3.3
Ruggles Station	0.0			0.0	0.0		0.0	0.4	0.0	1.5	0.2	2.2
Ruggles St							0.0	0.3	0.0	0.2	0.2	0.7
Dudley Station					0.0				0.2	3.4	0.9	4.5
Washington St/Melnea Cass Blvd										0.2	0.3	0.5
Albany Street										0.5	1.7	2.2
Totals	0.0	0.5	0.7	10.2	34.0	4.7	0.2	1.6	0.4	7.8	3.6	63.6

Figure 3: Origin destination matrix of average passenger flows between grouped stops for AM Peak inbound 47 trips during March, April, May 2014.

47 Outbound AM Peak — Origin \ Destination	Albany St/Harrison Ave	Eastern Washington St/ Melnea Cass Blvd	Dudley Station	Western Washington St/ Melnea Cass Blvd	Ruggles Station	Longwood	Kenmore	Lower Cambridgeport	Upper Cambridgeport	Central Square	Totals
Broadway Station	8.9	0.1	2.5	0.2	0.9	4.4	0.2		0.1	0.0	17.3
Albany St/Harrison Ave	1.8	0.0	1.4	0.1	1.1	2.2	0.4	0.0	0.2	0.2	7.4
Eastern Washington St/ Melnea Cass Blvd			0.0		0.1	0.2	0.1	0.0	0.0		0.4
Dudley Station				0.0	0.9	2.2	0.4	0.1	0.4	0.3	4.4
Western Washington St/ Melnea Cass Blvd				0.0	0.6	0.6	0.2	0.1	0.1	0.0	1.6
Ruggles Station						4.6	0.9	0.8	0.9	0.4	7.6
Longwood						0.5	0.5	0.5	3.1	2.5	7.2
Kenmore							0.0	0.4	2.1	4.3	6.8
Lower Cambridgeport								0.0	0.3	4.9	5.2
Upper Cambridgeport									0.1	1.2	1.3
Totals	10.7	0.1	3.9	0.4	3.7	14.7	2.7	2.0	7.2	13.8	59.2

Figure 4: Origin destination matrix of average passenger flows between grouped stops for AM Peak outbound 47 trips during March, April, May 2014.

Potential Route Modifications

The observations from the current use analysis show that the 47 corridor could be better served by increasing the frequency of service between Central Square and Longwood Medical Area during the peak periods. This could be achieved by either adding a supplementary 47A, serving Central to Longwood during the peak periods, or breaking the current 47 into two routes. The two routes could consistent of one serving passengers between Central Square and Ruggles; and one serving passengers between Longwood and Broadway Station. Splitting would have many benefits including:

- Improving the ability to match frequency to demand on the two segments
- Increasing the ease of interlining by locating new route endpoints near the large terminals of Ruggles and Dudley Stations.
- Could be implemented without negatively affecting most current users

A split would likely require additional resources during the off peak periods in order to meet current service levels due to overlapping service coverage.

Creating a 47A route on top of the current 47 and adjusting service to maintain even 7.5 minute combined headways in the peak direction during the AM Peak and 10 minute in the PM Peak will increase service between Central Square and LMA while only a small portion of riders would see a slight decrease in frequency.

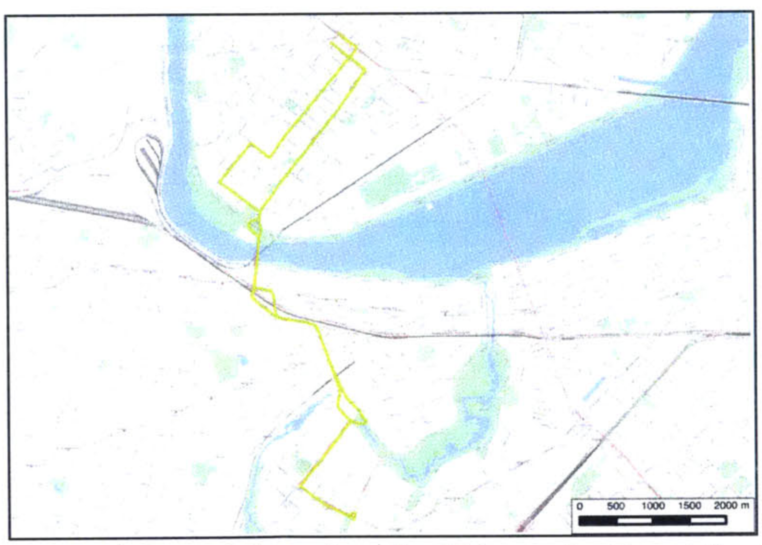


Figure 5: Potential routing of a 47A terminating at the intersection of Longwood Avenue and Louis Pasteur Avenue.

Route Consolidation

Consolidating the CT3 into the modified 47 was also explored since the routes are similar in many respects. They have a significant amount of overlap and both connect the Red Line, Boston Medical Center (BMC), Ruggles and Longwood.

While more people currently travel on the Andrew-BMC section of the CT3 (roughly 1000 passenger trips on the average weekday) than on the Broadway-BMC section of the 47 (approximately 750) the majority of outbound CT3 passengers transfer from the Red Line (63%+) on their way to BMC (52%) and Longwood (28%). The vast majority of these Red Line riders (91%) come from the south.

Because of the route similarities with the 47 most CT3 passenger journeys would only be slightly impacted, requiring a transfer from the Red Line at Broadway instead of Andrew (adding three additional minutes of travel time for those coming from the south). The 10 also serves passengers travelling between Andrew and BMC who are unable to transfer at Broadway with 20 minute headways during the AM Peak and 24 during the PM Peak.

Shifting the eastern endpoint of the 47 from Broadway Station to Andrew Station was also explored but was rejected, as this would eliminate all MBTA service along Albany Street north of BMC. With the amount of development occurring in the eastern South End it is important to maintain service along this corridor.

Consolidation would produce resource savings that could be used to either improve service on the modified 47 routes or other routes.

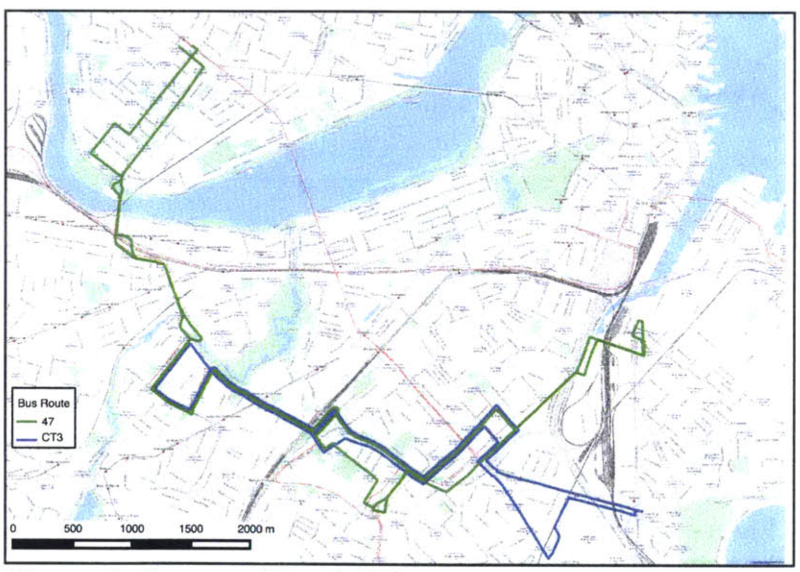


Figure 5: Current 47 and CT3 routes.

Route Variations

Finally, different route variations of the Broadway-Longwood segment were evaluated. The Central-Ruggles segment was straightforward with Ruggles as an endpoint because of its good terminal facilities and natural break in demand. However, there were two route components on the Broadway-Longwood segment that, based on service goals, could be adjusted. They include the location of the western endpoint and whether to bypass Dudley Station.

Western Endpoint Location

Extending the western endpoint of the Broadway segment to Longwood would drastically decrease the percentage of current passengers who would have to make an additional transfer with the new route configuration from about 15% to less than 5% for most time periods. Most of this improvement comes from giving Longwood commuters travelling from east of Ruggles a one-seat ride.

Period	Inbound (Broadway bound)		Outbound (Central Bound)	
	Ruggles	Longwood	Ruggles	Longwood
AM Peak	5.5%	2.7%	20.8%	4.6%
Midday	13.4%	4.1%	18.7%	5.8%
PM Peak	17.7%	3.7%	7.1%	3%
Evening	15.8%	5.8%	7.2%	3.1%

Figure 6: Percentage of current 47 riders who would have to make an additional transfer with the route modification for each direction, endpoint and time period.

Also, overlapping the western and eastern segments will significantly improve frequency between Ruggles and Longwood. This will improve access to Longwood for passengers transferring at Ruggles, which includes passengers coming from the Orange Line, Commuter Rail, and 7 bus routes (4 of which are key routes).

There would be some additional costs associated with running the eastern segment through Longwood. Extending the length of the route will increase cycle time, requiring more resources than the Ruggles endpoint variation to run the same frequency (see Operational Expenses section). There would also be fewer opportunities to interline vehicles since they would likely have to deadhead back to Ruggles or Dudley from Brookline Avenue.

Dudley Bypass

A Dudley bypass could have Broadway bound vehicles turn left on Washington Street from Ruggles Street instead of travelling down Shawmut Avenue and Central Square bound vehicles continue straight on Melnea Cass Boulevard without turning onto Washington St. This would reduce outbound running times on average 8 minutes and inbound by 5 minutes. It would also reduce congestion at Dudley Station where 17 routes currently stop.

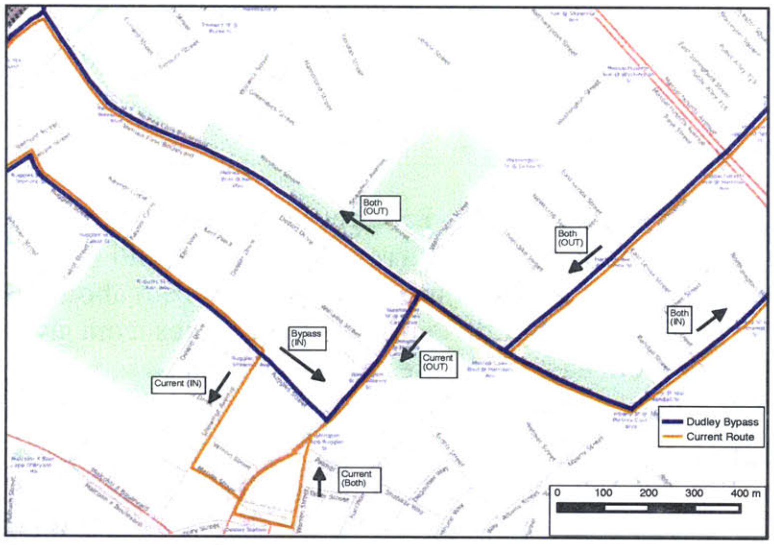


Figure 7. Map of Potential Dudley Station Bypass

However, the bypass would exclude the Dudley Station stop for Broadway bound trips at which 7% of current 47 riders board or alight and five stops for Central bound trips at which 16% of current 47 riders board or alight.

For Broadway bound riders, there would be two main options to complete their journey with the reconfigured route:

1. Walk 1/8 of a mile from Dudley to Ruggles St @ Washington St to board the 47
2. Wait for either the 1, 8, or Silver Line at Dudley for similar South End destinations

For Central bound riders, there are three options:

1. Transfer at Ruggles instead of Dudley (27%+ of current 47 riders who board at these stops transfer from routes that continue to Ruggles)
2. Take one of the Malcolm X Boulevard routes to Ruggles and then transfer
3. Take the 8 or 19 for a direct trip to Longwood.

Passenger Effects with the 47A Addition

In the PM Peak, in order to achieve 10 minute combined headways between Central Square and Longwood Medical Area both the 47 and 47A would need to operate with 20-minute headways. Current 47 headways vary between 15 and 25 minutes meaning the difference in headway passengers will face will vary based on the trip they ride though all passengers riding between Central and Longwood would see an increase in frequency. This amounts to over 600 passengers on an average weekday for a Ruggles turnaround configuration and 400 for a Louis Pasteur turnaround.

In the AM Peak, in order to achieve 7.5-minute combined 47-47A headways, inbound frequency on the 47 will need to be reduced (10 to 15 minute headways) and outbound frequencies increased (21 to 15 minute headways). All outbound passengers will benefit from this increased frequency.

21% of inbound passengers will see a reduction in frequency with a turnaround point at Ruggles and 32% with a turnaround point at the intersection of Longwood Avenue and Louis Pasteur Avenue. These are passengers who alight after the turnaround point.

The preponderance of riders would benefit from the increase in frequency, approximately 700 riders on an average weekday with a Ruggles turnaround configuration and 650 with a Louis Pasteur turnaround.

Operational Expenses

Split

Increasing the frequency of the Central-Ruggles route to 7.5 minute headways and Broadway-Longwood route to 15 minute headways during the AM Peak and 10 minute and 20 minute respectively during the PM Peak, same as adding a 47A, will help to better match supply to demand. There are some efficiency gains by consolidating the CT3 into the 47 though slow travel speeds over the BU Bridge and through Longwood make the increase in frequency expensive in terms of vehicle requirements especially during the PM Peak. Current conditions do not warrant a schedule adjustment during the off peaks.

The configuration of the new routes also affects the number of vehicles required. The longer the combined length of the two routes of a configuration the more cycle time and vehicles required. The following chart shows the change in vehicles required for each major time period and configuration from the current amount allocated between the 47 and CT3.

Configuration	AM Peak	Midday	PM Peak	Evening
Broadway-Ruggles (w/o Dudley)	0	-1	0	1
Broadway-Longwood (w/o Dudley)	2	1	1-2	2
Broadway-Longwood (with Dudley)	2-3	1	2	2

Figure 7. Increase in Vehicle Requirements from Current Combined CT3 and 47 Allocation by Configuration and Time Period

Adding a 47A

Adding a 47A, with the frequency changes mentioned above, will require 5 additional vehicles during the AM Peak and 4-5 in the PM Peak if operated between Central Square and Ruggles. Operating between Central Square and the intersection of Longwood Avenue and Louis Pasteur Avenue will require 3-4 vehicles in the AM Peak and 2-4in the PM Peak depending on scheduling aggression. With consolidation of the CT3 into the 47 these vehicle requirement costs would decrease by 4 for each period.

Interlining vehicles may reduce the costs listed above.

Transit Priority on BU Bridge

Buses often face congested road conditions as they cross the BU Bridge and adjacent rotary. This congestion is exacerbated during the peaks, with large increases both in average and variance of running time. Implementing transit priority measures in the area, even just during the PM Peak could produce significant running time decreases.

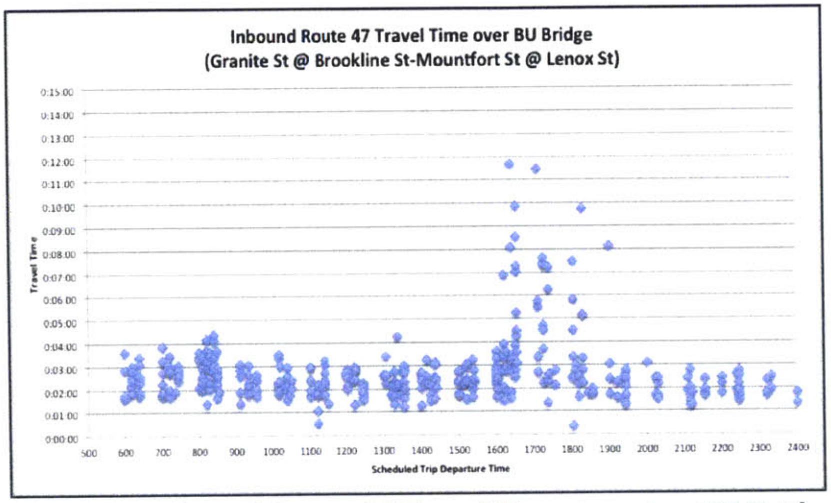


Figure 8. Inbound Route 47 Travel Time crossing BU Bridge by Scheduled Trip Departure Time

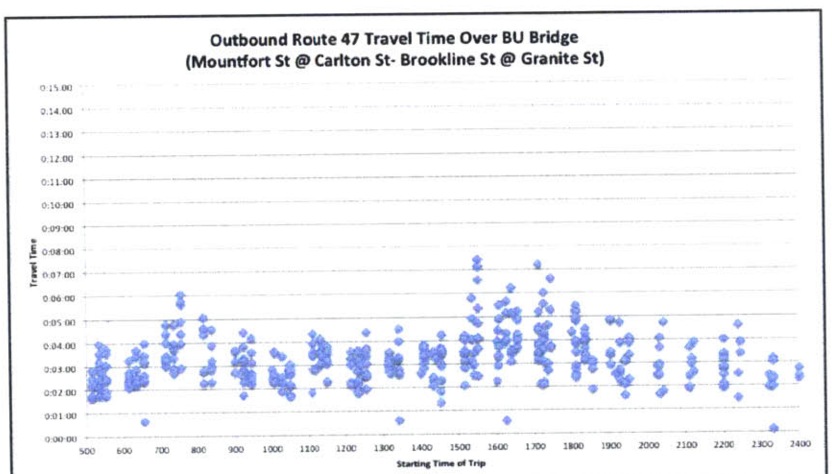


Figure 9. Outbound Route 47 Travel Time crossing BU Bridge by Scheduled Trip Departure Time

Recommendations

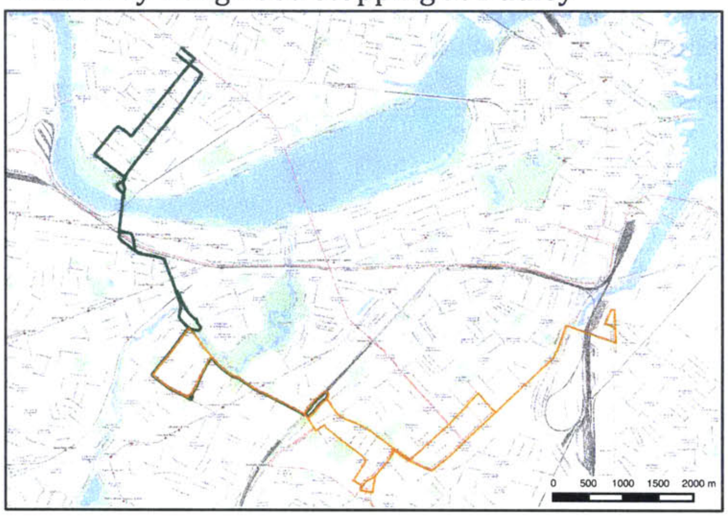
After the above analysis, the following recommendations for different service goals:

Split Proposal

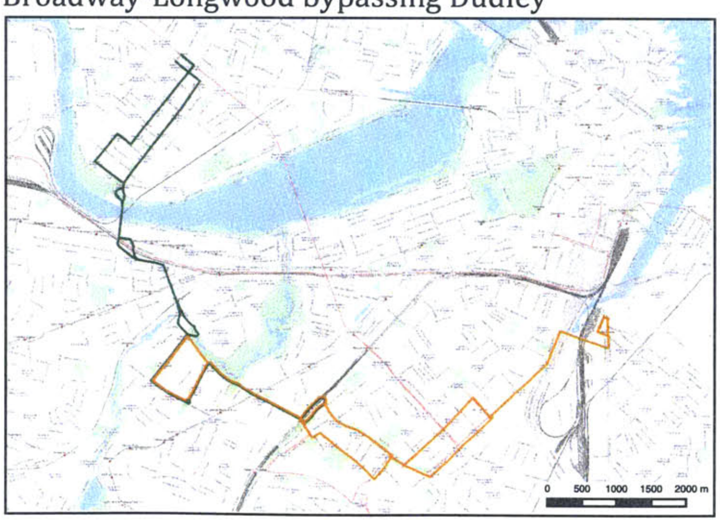
1. Consolidate the CT3 into the 47.

2. Split the 47 into 2 separate routes from one of the three following configurations:

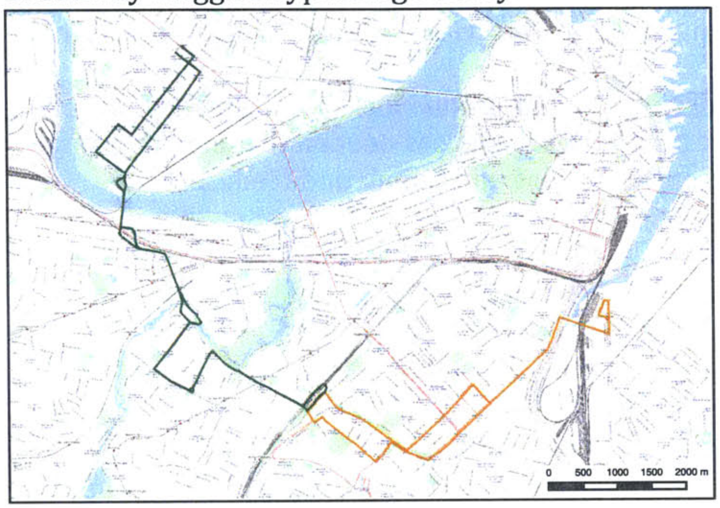
   i. Broadway-Longwood stopping at Dudley



   ii. Broadway-Longwood bypassing Dudley



iii. Broadway-Ruggles bypassing Dudley.



The Central-Ruggles route could operate with a 7.5 minute headways while the Broadway-Longwood route could operate with 15 minute headways during the AM peak. During the PM Peak, when travel demand is slightly less peaked, 10 minute and 20 minute respective headways could be run. During the off peaks, the routes could operate at the current headway of the 47.

The Broadway-Longwood variations are listed in order of decreasing service coverage and costs. A Broadway-Dudley-Longwood route would provide the most service coverage while being the most costly to run. The Broadway-Ruggles route would be the least costly to run but would directly serve the fewest people. The Broadway-Longwood route is a middle ground option.

47A Proposal

The 47A proposal would add an additional route (47A) on top of the 47 from Central Square to either the intersection of Longwood Ave and Louis Pasteur Ave or Ruggles Station during the peak periods. Both the 47 and the 47A would operate with 15 minute headways during the AM Peak and 20 minute headways during the PM Peak. The combined headway between Central and Longwood would be 7.5 minutes and 15 minutes for the entire 47 corridor during the AM Peak and 10 and 20 minutes respectively during the PM Peak.

The two proposals have similar costs when service coverage is taken into effect. A decision between the two rests mostly on implementation and flexibility preferences. The 47A proposal would be simpler to implement than the rerouting and elimination of routes as is suggested in the split proposal, as passengers would only see changes in frequency. However, it is not as flexible as the split proposal in terms of having the ability to change frequencies for the different segments in the future. The frequencies on the two split routes are largely independent and can be changed without drastically affecting the performance of the other route while the 47A proposal requires that both routes have the same frequency in order to maintain even headways over the critical Central-Longwood segment.

Overall, both of these recommendations increase frequency most where demand is highest limiting the amount of resources needed to meet the ever-growing travel demand between Central Square and Longwood Medical Area.

Bibliography

[1] American Public Transportation Association. Rising fuel costs: Impacts on transit ridership and agency operations. Technical report, American Public Transportation Association, 2008.

[2] Thompson, K, et. al. A socioeconomic study of carriage and platform crowding in the australian railway industry. Technical report, CRC for Rail Innovation, 2012.

[3] Lepore, S & Allen, K. Cross-cultural differences in tolerance for crowding: Fact or fiction? *Journal of Personality and Social Psychology*, 2000.

[4] Mahudin, M, et. al. Modelling the spillover effects of rail passenger crowding on individual well being and organisational behaviour. *WIT Transactions on The Built Environment*, 2011.

[5] Li, Z & Hensher, D. Crowding in public transit: A review of objective and subjective measures. *Journal of Public Transportation*, 2013.

[6] Tiranchini, A, et. al. Crowding in public transport systems: Effects on users operation and implications for the estimation of demand. *Transportation Research Part A: Policy and Practice*, 2013.

[7] Hirsch, L & Thompson, K. I can sit but i'd rather stand: Commuter's experience of crowdedness and fellow passenger behaviour in carriages on australian metropolitan trains. Australasian Transport Research Forum (ATRF), 34th, 2011, Adelaide, South Australia, Australia. Vol. 34 No. 0245.2011.

[8] Wardman, M & Whelan, G. Twenty years of rail crowding valuation studies: Evidence and lessons from british experience. *Transport Review*, 2010.

[9] Lin, T & Wilson, N. Dwell time relationships for light rail systems. *Transportation Research Record*, 1992.

[10] Senna, L. The influence of travel time variability on the value of time. *Transportation*, 1993.

[11] Uniman, D. Service reliability measurement framework using smart card data: Application to the london underground. Master's thesis, Massachusetts Institute of Technology, 2009.

[12] Adebisi, O. A mathematical model for headway variance of fixed-route buses. *Transportation Research Part B: Methodological*, 1986.

[13] Southeastern Pennsylvania Transportation Authority. Service standards and process. Technical report, Southeastern Pennsylvania Transportation Authority, 2014.

[14] Los Angeles County Metropolitan Transportation Authority. Transit service policy october 2015. Technical report, Los Angeles County Metropolitan Transportation Authority, 2015.

[15] San Francisco Municipal Transportation Authority. Municipal transportation quality review fiscal years 2013-2014. Technical report, San Francisco Municipal Transportation Authority, 2013.

[16] Regional Transportation District. Service standards. Technical report, Regional Transportation District, 2002.

[17] Kittelson & Associates et. al. *A Guidebook for Developing a Transit Performance-Measurement System - TCRP Report 88.* Transportation Research Board, 2003.

[18] Boyle, D. *Passenger Counting Systems- TCRP Synthesis 77.* Transportation Research Board, 2008.

[19] Furth, P, et. al. *Using Archived AVL-APC Data to Improve Transit Performance and Management - TCRP Report 113.* Transportation Research Board, 2006.

[20] Strathman, J, Kimpel, T , and Callas, S. Headway deviation effects on bus passenger loads: Analysis of tri-met's archived avl-apc data. Technical report, United States Department of Transportation, 2003.

[21] Milkovits, M. Simulating service reliability of a high frequency bus route using automatically collected data. Master's thesis, Massachusetts Institute of Technology, 2008.

[22] Zhao, J. The planning and analysis implications of automated data collections systems: Rail transit od matrix inference and path choice modeling examples. Master's thesis, Massachusetts Institute of Technology, 2004.

[23] Cui, A. Bus passenger origin-destination matrix estimation using automated data collection systems. Master's thesis, Massachusetts Institute of Technology, 2006.

[24] Gordon, J. Intermodal passenger flows on london's public transport network-automated inference of full passenger journeys using fare-transaction and vehicle-location data. Master's thesis, Massachusetts Institute of Technology, 2012.

[25] Wang, W. Bus passenger origin-destination estimation and travel behavior using automated data collection systems in london, uk. Master's thesis, Massachusetts Institute of Technology, 2010.

[26] Dumas, R. Analyzing transit equity using automatically collected data. Master's thesis, Massachusetts Institute of Technology, 2015.

[27] Viggiano, C. Bus user behavior in multi-route corridors. Master's thesis, Massachusetts Institute of Technology, 2013.

[28] Massachusetts Bay Transportation Authority. Ridership and service statistics-fourteenth edition 2014. Technical report, Massachusetts Bay Transportation Authority, 2014.

[29] Massachusetts Bay Transportation Authority. Service delivery policy 2010 update. Technical report, Massachusetts Bay Transportation Authority, 2010.

[30] Trepanier, M & Tranchant, N & Chapleau, R. Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems*, 2007.

[31] Hall, E. *The Hidden Dimension.* Anchor Books, 1966.

[32] Vanderwaart, C. Planning transit networks with origin, destination, and interchange inference. Master's thesis, Massachusetts Institute of Technology, 2016.

[33] Maltzan, D. Using real-time data to improve reliability on high-frequency transit services. Master's thesis, Massachusetts Institute of Technology, 2015.

[34] Gordon, M. Developing strategies for resource-constrained transit growth through increased private sector involvement. Master's thesis, Massachusetts Institute of Technology, 2015.

[35] Translink Rider Etiquette Guide: www.translink.ca/en/Rider-Guide/Etiquette-on-Transit.aspx.

[36] Metropolitan Transportation Authority's Courtesy Counts Campaign: http://web.mta.info/nyct/service/CourtesyCounts.htm.

[37] Center for Urban Transportation Research University of South Florida. Best practices in transit service planning. Technical report, Florida Department of Transportation Research Center, 2009.