20

REAL TIME DISPATCHING CONTROL IN TRANSIT SYSTEMS

by

Wei Song

B.E. Transportation Management Engineering
Northern Jiaotong University (1993)

M.S. System Engineering
Northern Jiaotong University (1996)

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 1998

Author ....
Department of Civil and Environmental Engineering
June 18, 1998

Certified by ....
Professor Nigel H. M. Wilson
Department of Civil and Environmental Engineering
Thesis Supervisor

Accepted by ....
Joseph M. Sussman
Chairman, Departmental Committee on Graduate Studies

Real Time Dispatching Control in Transit Systems

by

Wei Song

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

MASTER OF SCIENCE IN TRANSPORTATION

Abstract

Maintenance of reliable service is a goal of any rail transit agency. Reliability is difficult to maintain due to the perturbations that serve to disrupt headway sequences. These incidents that affect the service quality of transportation agencies can be categorized into two types, major disruptions and minor disturbances, based on their nature and causes.

To maximize the capacity of a rail transit line and avoid busing, single track operation is analyzed in this thesis to deal with major disruptions. Based on the tracks and crossover configuration of the Massachusetts Bay Transportation Authority Red Line, a full analysis of possible strategies is presented which might form the basis for a major disruption response system. This could take the form of pre-planned short term operation plans which would be geared to the type, location and time of day of the disruption.

The dispatching problem occurs around a terminal when a train is not expected to arrive at the terminal early enough to be dispatched on its next trip on schedule. This problem can be considered as a special case of minor disturbance. Its solution can also supply insight into the more general minor disturbance problem. We use holding and short turning as our control strategies to deal with the dispatching problem. Choosing minimizing passenger waiting time or the number of overcrowded trains as the objective, a heuristic dispatching control model is designed and evaluation and simulation models are used to estimate and compare the effectiveness of the current dispatching system and the heuristic dispatching control model. The results show that the heuristic dispatching control model could produce savings in average passenger waiting time of up to 14%, with the effectiveness increasing as the disruption becomes more severe.

As part of this research, a dwell time model is estimated for Red Line trains in order to predict the running time of a train to help select the appropriate control strategy.

Thesis Supervisor: Nigel H. M. Wilson

Title: Professor of Civil and Environmental Engineering

ACKNOWLEDGEMENTS

My deeply sincere gratitude goes to a large number of people who give me guidance and assistance, inspiration and patience.

First I would like to thank my advisor, Nigel Wilson, who give me the guidance, motivation and encouragement throughout the 2 years of my stay in MIT. He erects a statue of erudite literate in my heart. I have come to respect and admire Nigel for the high standards he has both for himself and for his students, his scrupulous attitude as a researcher, his humorous intelligence as a lecturer and his dedication both to education and to the filed of urban transit. My stay at MIT worked out better than I could have ever expected, and I owe much of it to Nigel. I hope one day I could be more like Nigel.

I would also like to thank Anne Herzenberg, Jeffrey A. Parker and Frank Feltes of MBTA who provide us both great suggestions for the project and important information that I need to complete it. In addition, they help me collect data from the OCS database and at Braintree Branch. Without their help and insights, my research will never be done.

Su Shen and Susan O'Dell help me get past many project problems. Susan's thesis develop the basis of my research, and Su bring out excellent contributions to the travel time estimation on the Braintree Branch.

Finally, I would like to thank my wife, Jiangmai, who support and encourage me through 12 years, who never loses confidence to our future, who always with me through every happiness and bitterness.

This thesis is dedicated to my parents. I hope they are proud of me forever.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

The urban high-frequency rail transit system has been playing an increasingly important role in urban transportation, because of concerns about the environmental, urban structure and social equity impacts of growing reliance on the car for urban mobility. The safety, speed, and capacity of such systems have improved with the implementation of advanced technologies and communications. However, transit service still suffers from minor disturbances especially during the peak period, which may in turn cause bigger delays, overcrowded trains, frustration and longer waiting time of passengers.

Transit agencies will usually employ several control strategies, such as holding, short turning and expressing, to deal with such disturbances. Unfortunately, it is quite difficult for dispatchers to select, within a short period of time, the best action from a system-wide perspective. However, new advanced technologies make it possible for us to develop heuristic real-time control models in order to help the dispatcher make these decisions more effectively.

Among all possible problems, dispatching at a terminus is perhaps the easiest to formulate and to implement the resulting control through a bell ring-off system. The strategies that are employed at a terminus can also be used to shed light on treatment of general disturbances at other points along the line. Dispatching control is a logical point of departure in developing a general real-time control system.

A real time dispatching control model to deal with disruptions is the focus of this thesis, and a specific dispatching control system (DCS) is developed for use as a module in the Massachusetts Bay Transport Authority (MBTA)'s new Operation Control System (OCS). This DCS will initially be applied on the Braintree branch of the MBTA system, but the design is general in nature so it could be adapted to control any terminal in any rail transit system.

1.1 Rail Transit Service

Rail transit service, one of the family of high-frequency transit services, has mean headways of less than 10 minutes. In long headway transit service, passengers tend to arrive at the stations based on the schedule and their expectations in order to minimize their waiting times[1]. For example, passengers may try to arrive at stations just before the scheduled trip departure time. However, in short headway transit service such as rail systems, passengers may not care about the schedule due to the high frequency nature of the service. Accordingly, passengers can be assumed to arrive at rail transit stations randomly. During a relatively short time period, say several headways, we also can assume that the passenger arrival rate is constant at a given station.

Due to the above assumption about the constant passenger arrival rate, we can conclude that the number of boarding passengers will be directly proportional to the preceding headway of that train. We also assume that dwell time is a function of the number of boarding and alighting passengers, and the load in the car[2]. Therefore, a train that already has a long preceding headway will have more passengers waiting at the following stations, and need more time for passengers to board and alight; thus its long headway will become longer and longer along the route. On the other hand, a train that already has a short headway will have shorter passenger waiting time at the following stations, and need less time for passengers to board and alight; thus its short headway will become shorter and shorter.

The above phenomenon can make two trains, in which the 1st train is behind schedule and the 2nd train is (initially) on time, have different travel times. The travel time of the 1st train will be larger than normal, while the travel time of the 2nd train will be smaller than normal. These different travel times, which will bunch these two trains together, may leave large gaps ahead of the 1st train, and cause another gap between the 2nd train and the 3rd train. These gaps will cause longer passenger waiting time at some stations, uneven loads on trains, and other bunching effects. Generally, this bunching phenomenon, also referred to as the pairing problem, is the reason that even minor disruptions may cause severe impacts on transit service quality, especially in peak periods. In the following example, we will look at how uneven headways affect the passenger waiting time. We assume that there is a normal situation and a bunching situation at the same station. The time period is 10 minutes, and the schedule headway is 5 minutes. The passenger arrival rate is constant during this period, say r passengers per minute.

Normal Situation (10 minute period)	Bunching situation (10 minute period)
The 1st and 2nd headway is 5 min.	1st and 2nd headway is 7.5 and 2.5 min.
Passenger waiting time: $1/2*r*5^2+1/2*r*5^2=25r$	Passenger waiting time: $1/2*r*7.5^2+1/2*r*2.5^2=31.25r$

The above results show that uneven headways cause larger passenger waiting time than even headways during the same time period.

Under such conditions, the on-time performance measure, which is critically important in long headway services, becomes secondary. Maintaining an even headway to minimize the probability of bunching is most important in terms of supplying good service quality. A good overall measure of service quality in short headway service is the expected passenger waiting time. Therefore, minimizing the total passenger waiting time is one reasonable objective of the dispatching control model, which will be discussed in this thesis. The expected passenger in-vehicle time is also an important measure of service quality. Passengers usually will value in-vehicle time and out-of-vehicle time differently, with in-vehicle time being less onerous than out-of-vehicle time. Typically, the value of out-of-vehicle time is 2-3 times the value of in-vehicle. Thus it would clearly not be appropriate to increase in-vehicle time by 10 minutes in order to save 1 minute of out-of-vehicle time.

The number of the passengers who are affected by the uneven headways is also important to the service quality assessment. We want to minimize the number of these affected passengers. Some transit agencies also consider the number of complaints they receive a good proxy of the service quality. Since the number of complaints is strongly related to the number of overcrowded trains or cars, we might have an objective to minimize the number of overcrowded trains. Moreover, we can minimize the numbers of passengers who have extra waiting time larger than a threshold value, based on an assumption that passenger will be less sensitive to shorter waiting time. To conclude, the objective of our dispatching control model is one or some combination of these goals, to be decided based on the real world situation.

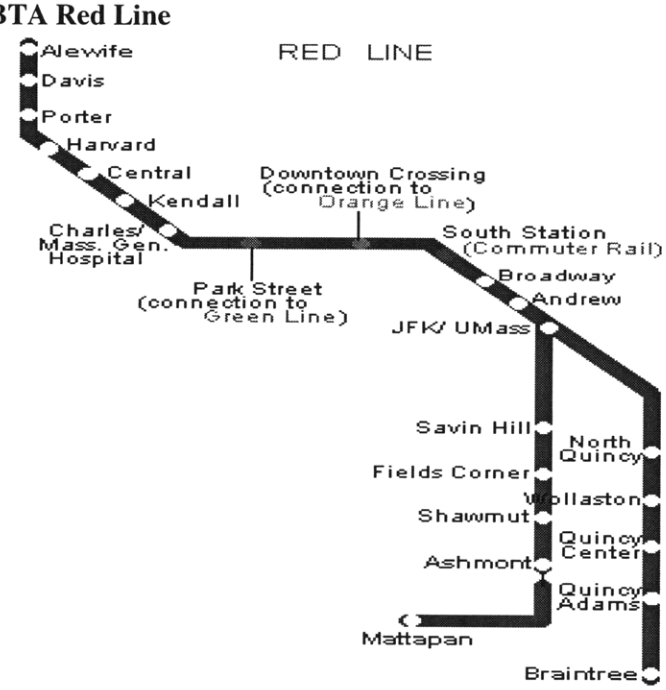
1.2 The MBTA Application Context

To test the theoretical advantages of real-time dispatching control, we developed a dispatching control system for the Braintree branch of the MBTA Red Line. Following is a brief introduction to the MBTA and Red Line Braintree branch, which is our case study.

The MBTA, which has been the dominant public transport operator in Boston since its creation in 1963, provides service on four major interconnecting rail lines, the Blue, Green, Orange, and Red Lines (see Figure 1-1). Among these four rail lines, the Red Line plays a critical role in the entire system (see Figure 1-2). It runs southeast from Alewife through downtown Boston before splitting with branches serving Braintree and Ashmont. It interconnects with two other major rail lines (Green and Orange Line) as well as commuter rail, intercity rail and bus service, and provides public transport access from the northwest and south to the Boston downtown, the financial and business center of Metropolitan Boston.

Figure 1- 1 MBTA Rail Lines

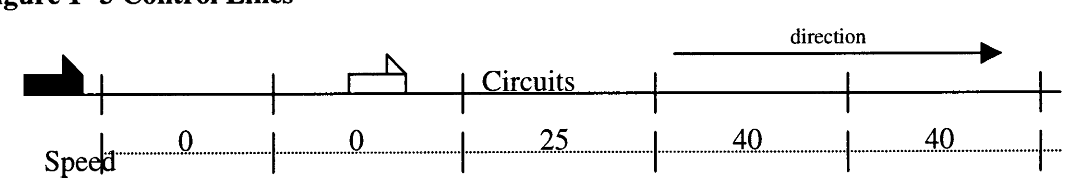
Figure 1- 2 MBTA Red Line



The Red Line has two branches, referred to as the Braintree and Ashmont branches, which merge at the JFK/UMass station. It is a high platform, third rail rapid transit system, which has relatively high travel speed and capacity. The MBTA defines its AM peak as 7:15 to 8:30 AM; and the PM peak as 4:30 to 6:30 PM. The average schedule headway is about 6-8 minutes on each branch with 3-4 minutes on the trunk portion during the peak period with 6-car trains. Currently, the MBTA uses 1500, 1700 and 1800 series cars. The newer Bombardier cars (1800 series) have more doors and are more advanced than the other series cars.

MBTA is using a specific circuit occupancy method, which is referred to as the "control lines," to determine the maximum permitted speed of a train, and to keep a safe stopping distance between two trains. In other words, the maximum speed of any train is determined by the location of the preceding train and the control lines. As the train approaches the preceding train, the train will receive a lower permitted speed or stop signal to ensure safe operations. While this method prevents trains from developing extreme bunching problems, we still find that bunching effects exist which cause significant impacts on the passenger waiting time, especially during peak periods. The following figure (Figure 1-3) is an illustration of the control lines. When there is a train on the extreme left circuit, the numbers on each other circuit are the maximum permitted speeds that this train would receive when its preceding train is on the indicated circuits. In Figure 1-3, the preceding train shown in outline only is on the third circuit. Therefore, the maximum speed for following train would be 0 mph: i.e., the train has to stop in order to keep a safe separation. After the preceding train moves to the 4th circuit, the maximum permitted speed for the following train will become 25 mph.

Figure 1- 3 Control Lines



Because the Red Line passes through the downtown area and has very heavy passenger flow during the peak period, it is highly susceptible to both major and minor disruption problems. Now, a new Operation Control System, which was fully accepted by the MBTA earlier this year, provides control information to the dispatchers to help resolve disturbance problems. Dispatchers, who used to make decisions based on years of experience, should be able to find better solutions. However, the large amount of data that the new OCS supplies may overwhelm the dispatchers. Moreover, it may be very difficult to choose correct strategies, especially some complex control actions such as short turning a train, within a short time period. Therefore, a heuristic real-time disruption control system needs to be added to the OCS to help develop fast responses to real-world problems.

A specific disruption problem that can be solved relatively easily is a disturbance around the terminus, and the rules or insights that we can get from this problem can then be used in the general disturbance control system.

In this thesis, a heuristic dispatching control system is developed to treat disturbances on the Braintree branch. In the future, it should be possible to extend this heuristic dispatching control system to the general heuristic disturbance control problem. The reason that we develop a heuristic system, rather than an optimization system, is that the optimization algorithm may not be feasible for real time control due to the complexity of the problem and the number of variables, as will be discussed in later chapters.

1.3 Prior Research

1.3.1 Bunching

The bunching problem in high frequency transit service has been noted since Welding and Day's[3] work in 1957. In their paper, they opened a new area of transportation study and defined various important concepts and relationships. They first discussed the contributors to headway variation; then, using the data from bus and train operation in London, they tried to find the cause of the irregular running time in general, which they believed might cause the pairing problem. They also estimated two linear functions for the two contributors of headway variation, running time between stations and dwell time. Finally, Welding and Day developed a simulation model for the Victoria Line in 1965[4]. In their simulation model, the position of a train at any moment is determined by the time at which it entered the system, its running speed between stations, the dwell time at the previous station and the position of the preceding train. The objective of the simulation model was to determine whether, under normal conditions and with the intended scheduled peak service of various trains per hour, an undesirable degree of irregularity would be likely to develop in the proposed Victoria Line. Welding and Day also presented several ideas for future application.

In 1969 Vuchic[5] developed an expression for the deterministic behavior of trains along a route, after one train was slightly delayed at a stop. His paper is the second one to address the bunching problem in rail systems.

Barnett[6] explained very clearly the departure and arrival pattern of trains in a rail system in his paper. He was the first researcher that tried to use mathematical tools to model the increasing irregularity of headway as a train moves along the line. In his paper, he presented an algorithmic solution that could be used to counteract the effects of random fluctuations in headways in rail systems. The core idea in his algorithm was to find an approximate optimal dispatching strategy, particularly holding, to smooth the operation. The objective of the algorithm was to minimize the average passenger waiting time and average delay for boarding passengers. This algorithm was tested on actual operation data from the MBTA Red Line, which was a single route without any branch at that time.

All this research gave us basic understanding and insight into the bunching problem of rail transit systems.

1.3.2 Holding

The holding strategy is the simplest and easiest control action to implement. Therefore, it is also the first strategy that has been analyzed. The following papers are notable: Osuna and Newell[7](1972), Barnett and Kleitman[8](1973), Barnett[6](1974), Turnquist and Blume[9](1980), Abkowitz and Engelstein[10](1984), Eberlein[11](1995), and O'Dell[12](1997).

All of these papers formulate the holding problem to minimize passenger waiting time, with the threshold headway and holding time as the decision variables.

Osuna and Newell (1972) focused on an idealized public transportation system, consisting of a single service point only. They formulate the dispatching of trains as a dynamic programming problem. In the simplified system, the travel times of successive trips are independent and identically distributed, and passenger arrival rate is constant. Osuna and Newell analyzed two scenarios, with one and two vehicles respectively. The objective of their programming problem was to minimize the total waiting time of all passengers.

The programming formulation showed that even idealized problems are difficult to analyze. Even though those two scenarios gave some typical properties of optimal strategies, many other approaches and scenarios would have to be analyzed before this kind of problem was fully understood. However, because of the state of the art at that time, it was not possible for Osuna and Newell to analyze more scenarios. They also suggested in their paper that other problems might require more intuition and less mathematics, and they believed that more sophisticated mathematics would not obviously help solve this type of problem.

Based on the principles that were obtained by Osuna and Newell, Barnett and Kleitman used a transportation system with one vehicle and several service points. In their system, passenger arrival rate was assumed to be constant, the capacity of the vehicle was unlimited, and the travel times between stations were randomly distributed. Barnett and Kleitman tried to minimize the average waiting time for passengers of the system. They found a simple statement of the optimal policy for systems with only one terminal stop at which interval control could be employed. They also studied a system with two terminals in detail and an optimal solution very closely related to the single-terminal solution was suggested.

Barnett (1974) analyzed a transit line with two terminals and one control station. His objective was to minimize the sum of passenger waiting time downstream from the control station and the average delay for passengers on held trains. His decision variable was still the threshold headway. To replace the complex assumption of general continuous probability distribution of train arrival headway, Barnett used a simpler discrete distribution approximation. Barnett found that the simple holding strategy that he obtained in his paper could probably be implemented very easily, and thus represented a feasible basis for improvement of bus, trolley, and rail transit operation.

Turnquist and Blume (1980) adopted the idea of discrete arrival headway distribution from Barnett (1974), but used a more general probabilistic model in their analysis. The objective of their research was to analyze holding strategies by using a very simple probability model of vehicle arrival time to get some insight into the problem. They tried to find when and where in the system vehicles should be held. Turnquist and Blume also noted the negative correlation between successive headways due to the pairing problem; however, they did not obtain a mathematical solution to the correlation relationship because of the complexity of getting reliable estimates of the covariance.

Using a general model of the probability distribution of headways between successive vehicles, Turnquist and Blume studied two simple cases that provided approximate upper and lower bounds of the potential benefits of a holding strategy, given the objective of minimizing the total passenger waiting time and average delay time of passengers on held trains. These upper and lower bounds were thought to be helpful in deciding whether a holding strategy would be beneficial.

Heavily based on empirical data analysis, Abkowitz and Engelstein (1984) developed a method that was simple to use and did not require extensive data from the transportation agencies. Their research was organized into six parts: determination of mean running time, determination of running-time variation, determination of headway variation, determination of passenger waiting time, identification of optimal control strategies, and establishment of operator compatibility with the developed strategies. The first four parts served as the input to the fifth part, and the sixth part concerned translating the research results into the operator strategies. Their objective was to minimize the total passenger time along the route. Their decision variables were the location at which to hold the train and the threshold time to hold it. Abkowitz and Engelstein also tested their methods in a case study, for three routes in Los Angeles, and found that holding was an effective strategy that could reduce the total passenger waiting time by 5%. They also suggested that the optimal holding point should be just before the high demand station.

Eberlein (1995) represents the definitive work to date on the real-time control problem. She considered many control strategies, including holding, expressing and deadheading, independently and in combination, in her dissertation. She studied these control strategies in two different types of transit systems: the one of ultimate interest is called system G, which stands for a "general" transit system; the other is called system F, which is a special case of system G, where F stands for "fixed" parameters. Using data from the MBTA Green Line, Eberlein found that the result of the first model, which had various simplifying assumptions, could shed some light on the solution of the second model, which was intractable. She concluded that holding is the best individual strategy, while the most effective and least disruptive control policies were found to result from the use of coordinated combined strategies. The results of her thesis provided important information regarding the design and implementation of real-time control systems and help shed light on the potential improvements of more advanced control systems.

O'Dell (1997) studied the use of holding and short-turning strategies to minimize passenger waiting time following a disruption. She sought to develop a real-time control decision support system to help the operator. Similar to Eberlein, O'Dell also used two models, one of which is a simplified model, and the other is a more realistic, but deterministic, generalized model of a rail system. She presented linear and mixed integer programming formulations for several holding and short-turning strategies. Using the MBTA Red Line as her case study, O'Dell concluded that passenger waiting time could be significantly reduced by applying the optimal set of holding and short-turning controls which were output by her model.

1.3.3 Expressing and Short Turning

Since these two strategies have only been analyzed in the last 10 years, and almost all papers came from MIT, we discuss these two strategies together.

Macchi[13](1989) wrote the first paper dealing with the expressing strategy in real-time control. He developed an express decision-making model to evaluate the waiting time impacts of expressing trains on the MBTA Green Line. Using real and randomly generated data, he coded his model in a simulation program and used the simulation to analyze expressing on two different segments on the Green Line. He also concluded that an automatic vehicle identification (AVI) system on the Green Line would supply more information and significantly improve the quality of expressing decisions. There are several simplifying assumptions in his simulation: 1) no train capacity constraints; 2) no variation of headway downstream; 3) the expressing train does not influence the next trip. However, this paper still sheds some light on the expressing problem.

While Macchi studied the expressing strategy, Deckoff[14] (1990) examined the short turning strategy in his thesis. He tried to measure and predict the impacts on transit passengers of short-turning trains. Choosing the MBTA Green Line B and D branches as case studies, he presented a model that could help dispatchers make short-turn decision based on accurate predictions of passenger impacts. Deckoff concluded that his model with the help of the AVI system would significantly improve the quality of the service on the Green Line. In his thesis, he assumed that there was only one station where trains could be short-turned, and the headways were constant. Unlike Macchi, he included train capacity constraints in his model.

Soeldner[15](1993) extended the previous work of Macchi and Deckooff. After comparing expressing and short-turning strategies in his model, he tried to combine them under the same assumptions of previous papers. Finally, Soeldner concluded that control decisions should be made as soon as possible. He also argued that the short-turning strategy might be more appropriate to deal with disruptions than the expressing strategy, since expressing strategy are usually restricted by the preceding train which reduces its positive benefits. However, both control strategies could reduce the passenger waiting time.

Coor[16](1997) focused on the short-turning strategy, and developed a model to simulate short turning on the MBTA Blue Line. The objective was to minimize the total passenger waiting time. The inputs to the model included passenger arrival rate and passenger alighting proportions for each station on the line, average inter-station running times, and initial sequences of train headways. The output of the model was the change in total passenger waiting time for the system from short turning. Coor also introduced the train dwell time as a function of total passenger boardings and alightings. The model showed that many short-turns on the Blue Line could make the situation worse, rather than better, and short-turning should not be used to compensate for insufficient allowed round-trip running time.

1.3.4 Others

There are other papers dealing with the analysis of operation control and passenger delay, including Newell[17](1971), and Furth[18] (1985.)

Newell (1971) used a single route with single origin and destination points as his system, under the assumption that arrival rate of passengers is some continuous function of time. He used an analytic approach to get some insight into the dispatching problem, even though the result of an analytic approach might not be as accurate as that of a computer simulation approach. The objective of Newell's model is to minimize the total waiting time of all passengers, and the decision variables are the departure times for n trains. In a system with the simplifying assumption that the capacity of the vehicles is sufficiently large, Newell concluded that the optimal flow rate of vehicles and the number of passengers served per vehicle would vary with time approximately as the square root of the arrival rate of passengers. Afterward, Newell relaxed the assumption of unlimited capacity of vehicles, and concluded that the dispatch schedule should be modified so that certain vehicles were dispatched as soon as they were full to capacity.

Furth (1985) discussed the "alternative deadheading" strategy for urban bus routes that have directional imbalance in passenger flow during some periods of time. By reducing the cycle time, deadheading could even out a gap and raise the quality of service. Furth developed a formula for the number of buses needed to meet a regular alternating deadheading schedule, and presented the advantages of this special scheduling strategy. Even though this strategy applies directly to scheduling rather than control, Furth's paper still shed some light on the analysis approach to the operation control problem.

1.4 Thesis Organization

Chapter 2 will introduce the disruption control problem. Disruptions are classified into two types, major and minor, that will be defined and described. Afterwards, a strategy for dealing with major disruptions will be briefly discussed. Finally, I will focus on the minor disruption problem, and introduce and summarize Eberlein (1995) and O'Dell's (1997) research on the holding and short-turning control strategies.

Chapter 3 begins by introducing the characteristics of the dispatching problem, as a special case of the minor disruption problem. I will also formulate the dispatching problem mathematically, and get some insights into this as background for developing a heuristic model. Finally, I will discuss some feasible strategies to deal with the dispatching problem.

Chapter 4 will describe one of the most important functions in any control system, the dwell time function. I will discuss the theory of dwell time and then analyze dwell time data from the MBTA. Finally a dwell time function will be estimated.

Chapter 5 will present the heuristic real-time dispatching control model. Based on the insights from Chapter 3, I will develop an overall strategy for the heuristic model and its structure. A proposal for the heuristic model will be presented as part of the MBTA's OCS. Then, we use a simulation model to test our heuristic model, based on data from the MBTA Red Line Braintree branch. Finally, we will compare the simulation results with actual performance and evaluate the performance of the heuristic model.

Chapter 6 will summarize the findings and conclusions of this research. Directions for the future research will also be proposed.

Chapter 2

Disruption Control System

In this chapter, we will first discuss two different types of disruptions. Then we will present the strategies developed to deal with major disruptions and the specific application on the MBTA Red Line. Finally, we will introduce the strategies for minor disturbances, which are partly based on O'Dell's (1997) thesis.

2.1 Types and Frequencies of Disruptions

Based on discussions with MBTA staff, we decided to classify disturbances into two categories: major disruptions and minor disturbances. Correspondingly, we also divide the control strategies into two categories: planning and real time. In the real world, disruptions happen frequently on any rail system. Most of them are minor, and we can use real-time control strategies to deal with them; however, some of them are major, and for these we need an alternative operations plan to put into effect until the problem is resolved. Generally, the difference between planning and real-time control strategies is whether we need to change the operation plan. The planning control strategies involve changes of a persistent nature. On the other hand, real-time control strategies are designed for immediate but short-run implementation to remedy specific operational problems, without exerting any influence on the longer-term operation plan.

With regard to different kinds of disruptions, different control strategies are appropriate. For example, if there is a bomb threat report at Kendall Square, the passengers at the Kendall Square will be evacuated and station has to be closed to investigate. Therefore, we could not operate trains through Kendall Square following the schedule due to the safety concern. The only choice in this situation is to operate one loop between Alewife and Central Square and other loops between the two branches and Charles. We also have to call for buses to deal with the operation between Central Square and Charles. This event can be categorized as a major disruption. On the other hand, if a train at Kendall Square has a malfunctioning door which could not be closed for a while, it will be categorized as a minor disturbance. We do not have to change the operation plan in this situation since the disturbance only lasts for a very short time period and the delay can be dealt with through control strategies.

During the two-year period ending 31 August 1996, there were 323 incidents or disruptions (approximately three disruptions per week) on the MBTA Red Line that resulted in recorded passenger waiting time delays larger than 15 minutes. We chose a sample of 57 incidents and reviewed the dispatcher's log to determine their causes, and to study the operations control strategies that were used to resolve them. We were able to determine 47 of these incidents with a high degree of certainty.

We separate the 57 incidents into 2 sets:

(1) Delays longer than 20 minutes: There were 33 incidents of this type with disabled train being the major cause (13 incidents), along with bomb threats (4), fires on trains (3) and other causes.

(2) Delay between 15 and 20 minutes: There were 24 incidents of this type. "Disabled Train" was still the major problem, with a total of 5. There were 7 incidents for which no definite reason was recorded.

Based on further discussion with MBTA staff, we think delays longer than 20 or 30 minutes should be treated as major disruptions, with others classified as minor disturbances. From these 33 major disruptions, we found only 2 incidents in which the MBTA used substitute buses (both of them were to deal with bomb threats), and we did not find any case where the dispatchers used single-track operation. However, we believe that single-track operation should be considered when we have a major disruption, because it may avoid calling for buses. The MBTA tends to prefer other control strategies rather than calling for buses except when busing is the only option that can supply the needed capacity, because calling for buses is usually hard to organize and expensive. In the following section, we will discuss major disruptions, and assess the role that single-track operation can play in dealing with them. We will also identify specific situations in which busing is necessary.

2.2 Major Disruption

When an incident lasts more than 20 or 30 minutes, we classify it as a major disruption. These disruptions are usually caused by serious problems such as a fire, police action or bomb threat or severe technical problems.

2.2.1 Objective and Methodology

Generally, in a major disruption, we will lose part of the track and/or a station. Because we cannot afford to wait until normal operation can be restored, as when a minor disruption occurs, we need to reschedule or redesign the operating plan to run a single-track operation where possible. Obviously, this will decrease the capacity of the entire system. Thus, the objective of the rescheduling in major disruption is to maximize the reduced capacity of the rail system at the most constrained point and to see whether this can carry the volume of passengers traveling. If not, then busing is needed. In some situations, such as station fire or bomb threat, we will lose both tracks. Busing around the location is then the only choice we have and so it is beyond our analysis scope. In this thesis, we will focus on the single track operation plan redesign to try to avoid busing whenever possible.

Due to the different track and station configurations, different strategies may be appropriate to achieve our objective depending on the location of the blockage or disruption. Therefore, first of all, we summarize the states of the track (and stations) and the strategies which may be appropriate for each state, when the disruption happens at a specific station or between specific stations. Using the MBTA Red Line as our case study, we find that there are 8 possible states when the disruption happens at a station; and 4 possible states when the disruption happens between stations (see the following diagrams.) Various strategies are possible including a single loop, two loops plus a shuttle, two overlapped loops, etc. Second, we try to identify all the feasible strategies when the disruption happens at a particular station or track section. Thirdly, we compare these strategies and find the best solution for each scenario based on the resulting headways and capacities. Finally, we compare the capacity of the best strategy with the passenger demand, and find out whether we need to call for buses.

2.2.2 Alternative States and Strategies

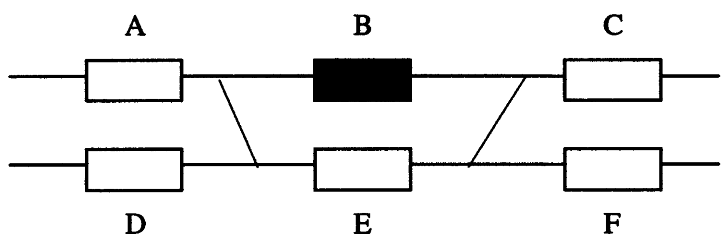
Based on the configurations of the track and crossovers, we can identify 8 states when the blockage happens at station and 4 states when the blockage occurs between stations. In the following figures, we use lines to represent tracks and crossovers, and use squares to represent stations. A solid square means that there is a disruption at that station, while a solid circle represents a blockage between two stations. The stations and crossovers are labeled by letters. Dwell times are assumed to be constant cross the stations.

For each type of blockage location, we redesign the operation plan and compute the headways for the new operation plan. To illustrate this process, we describe two states and their corresponding strategies in this section, one for a disruption at a station and the other for a disruption between stations. The other states and strategies are included in Appendix A.

I. State SA (disruption at a station)

In State SA (Figure 2-1), there is a disruption at platform B and the track configuration includes crossovers as shown between A and B and between B and C. The crossovers are indicated by location and direction. On the Red Line, Central Square Southbound is an example of a station of this type. There are two possible strategies which would be used to deal with this disruption. Strategy SA1 is based on two loops which overlap at E, while SA2 uses a single loop with single track operation through station E. Figure 2-2 can then be used to compute the minimum headways achievable for strategy SA1 under three scenarios.

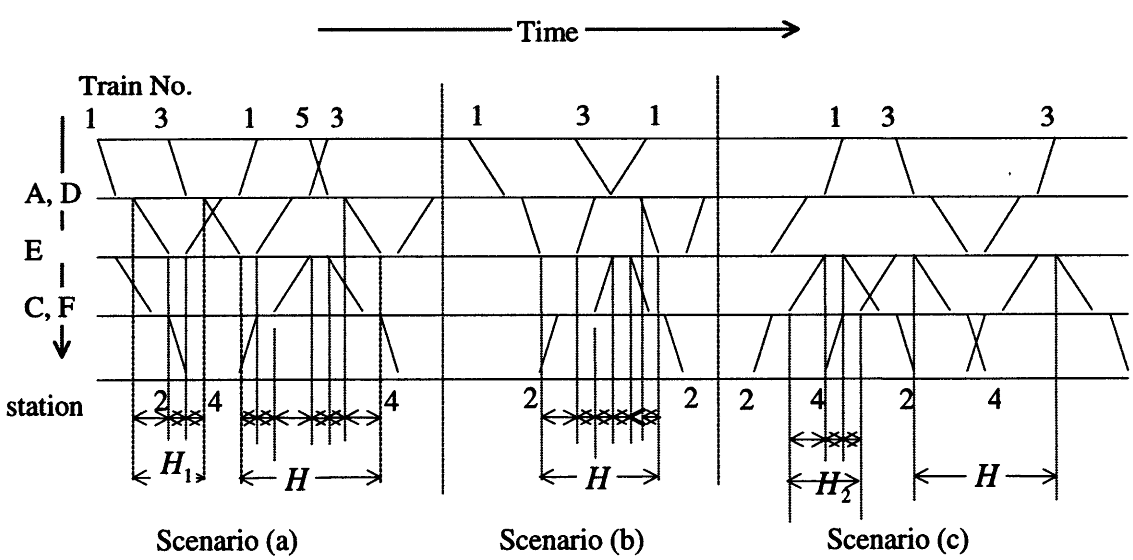
Figure 2- 1 State SA



Strategy SA1: Loop1: D-E-A
Loop2: C-E-F

Strategy SA2: Loop: D—E—F—...—C—E—A

Figure 2- 2 Strategy SA1



In this figure, the horizontal axis represents time, while the vertical axis is the station. The trains are distinguished by number, and their movement is represented by the angled lines linking stations. Three scenarios can be defined for strategy SA1, the two overlapped loops, based on where the maximum passenger flow occurs. For each scenario, the headways are computed using the following variables:

t_{dwell} : train's standard dwell time at station;

t'_{dwell} : train's extended dwell time at turning station;

t_{switch} : the time to change the switch at the crossover track;

$t_{x,y}$: Minimum running time from station x to station y;

t_x^{signal} : The time for a train clear station x.

Scenario (a)

In this scenario, the left loop AED has the heaviest passenger flow, and so we operate more than one train (up to n trains) in the heavy loop, between consecutive trains running in the light loop. In the figure, we show 2 successive trains operating in loop AED, between successive trains in loop CEF (that is n=2). Then H_1, the minimum headway between trains on loop AED, and H, the headway between trains on loop AED when a train a train on loop CEF intervenes, are as follows:

$$H_1 = t_{D,E} + t'_{dwell} + t_{E,A}^{signal} + t_{switch}$$

$$H = 2 \cdot t'_{dwell} + 2 \cdot t_E^{signal} + t_{C,E} + t_{D,E}$$

$$H_{AED} = H_1 \text{ or } H \quad \text{(alternating for n=2)}$$

$$H_{CEF} = (n-1)H_1 + H$$

$$H_{CEF} = H_1 + H \qquad \text{when n=2}$$

Scenario (b)

In this scenario, the heaviest passenger flow is through station B/E. We must guarantee that the train movements through station E are balanced with trains alternating in each loop.

$$H = 2 \cdot t'_{dwell} + 2 \cdot t_E^{signal} + t_{C,E} + t_{D,E}$$

Scenario (c)

In this scenario, loop CEF has the heaviest passenger flow. Similar to scenario (a,) but H_1 is replaced by H_2:

$$H_2 = t_{C,E} + t'_{dwell} + t_{E-F}^{signal} + t_{switch}$$

$$H = 2 \cdot t'_{dwell} + 2 \cdot t_E^{signal} + t_{C,E} + t_{D,E}$$

$$H_{CEF} = H_2 \text{ or } H \qquad \text{(alternating for n=2)}$$

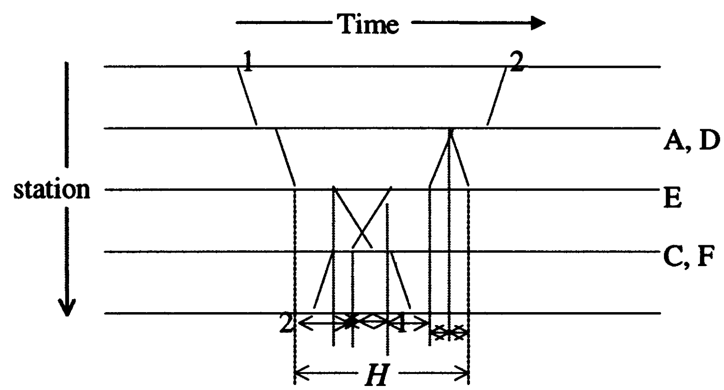
$$H_{AED} = (n-1)H_2 + H$$

$$H_{AED} = H_2 + H \qquad \text{when n=2}$$

In State SA, we can also choose strategy SA2, a single track loop. No matter where the heaviest passenger flow is, strategy SA2 has just a single headway (see Figure 2-3) as follows:

$$H = 2 \cdot t_{dwell}^{E-NB} + t_{E,F}^{signal} + 2 \cdot t_{switch} + t_{C,E} + t_{E,A}^{signal} + t_{D,E}$$

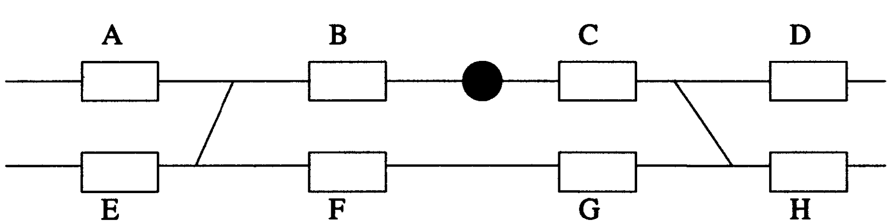
Figure 2- 3 Strategy SA2



II. State TA (disruption between stations)

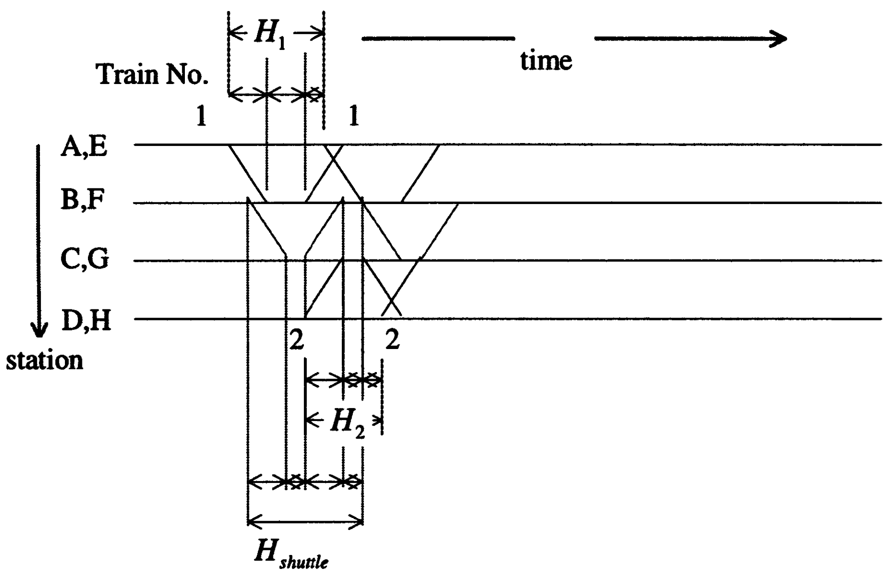
In State TA, there is a disruption between stations C and B, and the track from C to B has to be closed. Track and crossovers are as shown in figure 2-4. An example of this would be a disruption occurring between Downtown Crossing and Park Street Northbound. In this case, there is only one strategy possible TA1, consisting of two loops and a shuttle. We run two loops on each side of the blockage, and run a single train as a shuttle between these two loops on the available track. Since the shuttle will use a different track and platform from the loop, the operations of shuttle and loop will run independently and not interfere with each other. Figure 2-5 shows the train movements for this strategy.

Figure 2- 4 State TA



Strategy TA1: Loop1: E—B—A Loop2: D—C—H Shuttle: F—G

Figure 2- 5 Strategy TA1



$$H_1 = t_{A,B} + t'_{dwell} + t^{signal}_{B,A} + t_{switch}$$

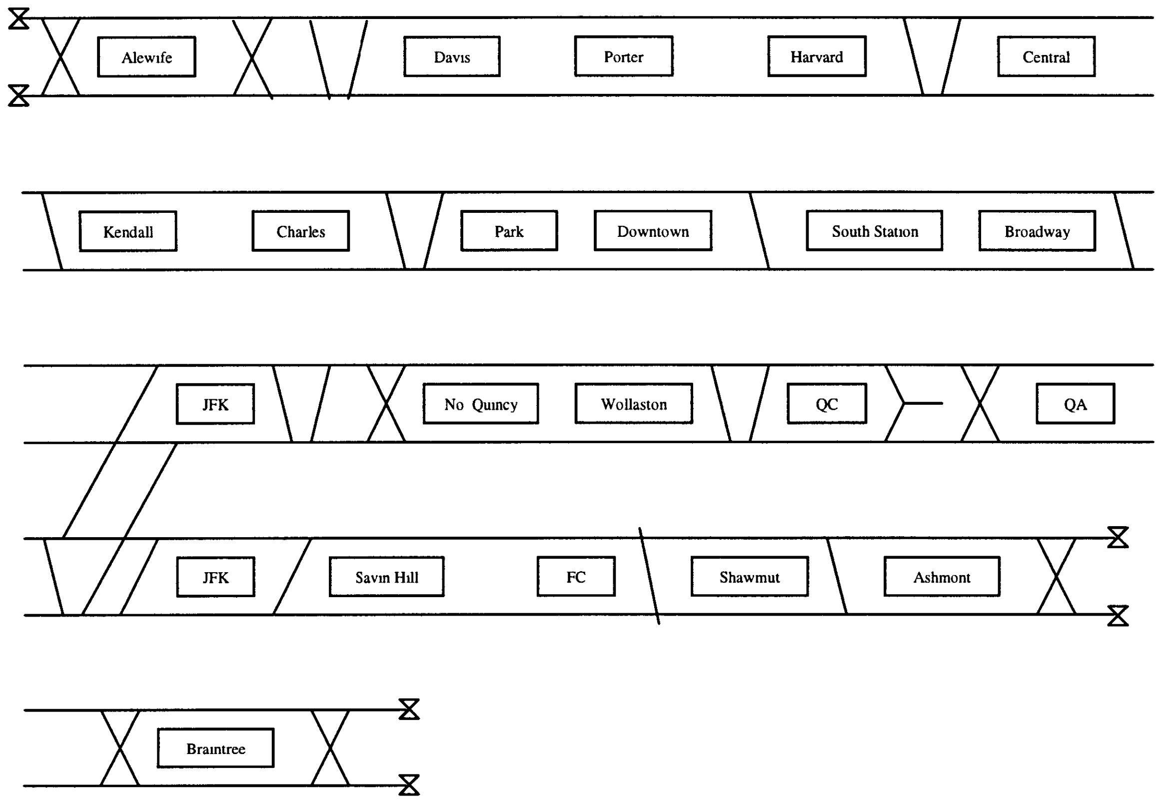
$$H_2 = t_{D,C} + t'_{dwell} + t^{signal}_{C,D} + t_{switch}$$

$$H_{shuttle} = t_{F,G} + 2 \cdot t'_{dwell} + t_{G,F}$$

2.2.3 Red Line Analysis

Based on the states defined in section 2.2.2 and Appendix A, we can classify all possible major disruptions on the MBTA Red Line (Figure 2-6) as shown in Appendix A. After defining all states and strategies, we compute the minimum headways and thus the maximum capacities for each scenario. Comparing these results with the actual passenger flow on the Red Line, we can find the scenarios under which the MBTA must provide substitute bus service because the reduced rail capacity is inadequate.

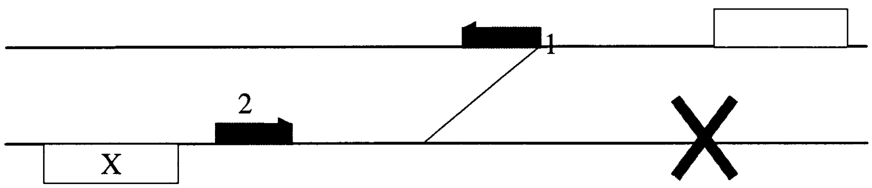
Figure 2- 6 MBTA Crossover Configurations



In this analysis, we make the following assumptions:

1. The train will start at the nearest feasible point to the blockage with an initial speed of 0.

Figure 2- 7 Train Operation During the Disruption



For example, in Figure 2-7, train 2 will stop at the nearest section (as opposed to stopping at platform X) permitted by the control lines until train 1 has cleared the crossover. In this case, we have to operate this train on the reverse track. The MBTA uses special control lines, called reverse control lines, to determine the maximum permitted speed in reverse operations. Based on the safety concern, this speed is usually lower than the normal situation. However, the reverse control lines are unavailable or malfunctioning on some sections of the Red Line (Table 2-1). Therefore, we have to operate under station-block rules to ensure safe operation. In other words, we have to make sure there will be only one train on the reverse tracks between two stations at any time, and the maximum speed for these station-blocked trains is 25mph.

Table 2- 1 Reverse Control Lines

From	*To*	*Control System*	*Current Situation*
Alewife	Harvard	Rev. control lines	
Harvard	Park Street	Rev. control lines	They are non-operational
Park Street	JFK	No Rev. control lines	
JFK	Ashmont	No Rev. control lines	
JFK	North Quincy	Rev. control lines	Just work in the northbound direction.
North Quincy	Braintree	Rev. control lines	

2. The time to throw a power switch is 10 seconds, and the time to throw a hand switch is 300 seconds. Table 2-2 lists the types of switches on the MBTA Red Line.

Table 2- 2 Switch Type

Switch Location	**Type**
Alewife	Power
Davis	Power
Harvard	Power
Kendall	Hand
Park Street	Power
Downtown/South Station	Hand
Broadway	Hand
Andrews	Hand
JFK junction (all switches)	Power
Fields Corner	Hand
Shawmut	Hand
Ashmont	Power
Braintree branch (all switches)	Power

3. The maximum speed of trains on a crossover is 10 mph. However, the speed can be 25 mph if the crossovers are entering a terminal.

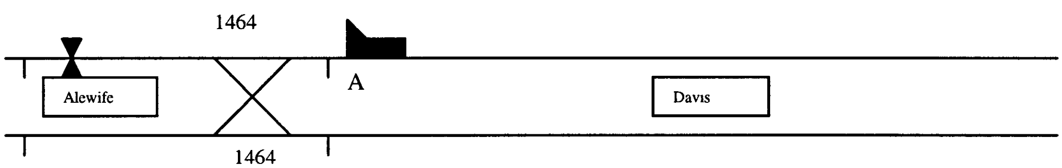
4. Since disruptions are most critical in peak periods, I assume that there are 6 cars in each train, and each car can carry 160 passengers without overcrowding. The length of the train (6 cars) is 420 feet.

5. The acceleration rate for the train is 2.75 mphs, and deceleration rate is 3 mphs.

6. For non-terminal stations, the normal dwell time is 30 seconds; at route termini and the extended dwell at disrupted station, the dwell time is 120 seconds.

Using the above data, we can compute the headways for each operation plan. An example is presented as follows:

Figure 2- 8 Example of Headway and Capacity Computation



In the above figure, there is a disruption occurring at Alewife northbound platform, which makes northbound platform unavailable. We can run one single loop to deal with this problem. The train has to stop before the crossover section until clearance is received for the crossover section. After the preceding train clears the section and the interlocking is reset, the train will receive a signal to approach Alewife. The length of train's move is 1464 feet. The minimum headway for the Davis → Alewife loop could be calculated as follows:

1. Acceleration time from point A = 25/2.75 = 9 seconds. The train will start from Point A and accelerate to 25 mph.

2. Travel time at 25mph = $(1464 - ½ * 9.09^2 * 2.75 * 1.467 - ½ * 8.33^2 * 3 * 1.467)/(25*1.467)$ = 31 seconds. Train can pass this crossover section at 25 mph.

3. Deceleration time from 25mph = 25/3 = 8 seconds.

4. Dwell time at Alewife = 120 seconds.

5. Acceleration time from Alewife = 25/2.75 = 9 seconds.

6. Travel time past the crossover section = $(1464 - ½ * 9.09^2 * 2.75 * 1.467)/(25*1.467)$ = 35 seconds.

7. Time to reset the crossover = 10 seconds, because the crossover is a power switch.

Total headway = 9 + 31 + 8 + 120 + 9 + 35 + 10 = 222 seconds ≈ 4 minutes.

The capacity of the loop = 60 / 4 * (6 * 160) = 14400 passengers/hour. There are 6 cars in each train, which can accommodate 960 passengers/train.

The results of the entire analysis for all possible disruption locations are summarized in Tables 2-3 through 2-6. For each location blockage, we obtain the optimal strategy and the corresponding capacity. To determine the feasibility of the optimal strategy, we list the peak demand for each portion of the operation plan, as well as the off-peak demand. The peak demand is estimated based on the AM and PM line volume in CTPS data, while the off-peak demand is based on the CTPS data between 12:00-13:00 as shown in Table 2-7. The number is marked in boldface when the optimal strategy can not supply enough capacity and thus busing is needed.

Table 2- 3 Major Station Disruption on the MBTA Red Line Northbound

Location of Blockage	Northbound				
	Optimal Strategy	Feasible Headway	Feasible Capacity	Off-peak Demand	Peak Demand
Alewife	Davis → Alewife loop	4	14400	1923	9390
Davis	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 3200	1923 662	9390 3562
Porter	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 **3200**	1923 662	9390 **3562**
Harvard	Harvard→Porter loop Porter→Ale. shuttle	11.5 12.5	**5000** 4600	1923 245	**9390** 2215
Central	Park → Kendall loop Kendall → Ale. loop	11 12.5	**5200** **4600**	1923 1491	**9390** **6886**
Kendall	Park → Charles loop Harvard→Central loop Cent.→Charles shuttle	7 11 12.5	**8200** 5200 **4600**	1923 1491 1710	**9390** 4964 **7760**
Charles	Downtown→Park loop Harvard→Central loop Park→Central shuttle	12 11 16	**4800** 5200 **3600**	1923 1491 1869	**9390** 4964 **7760**
Park Street	South→Down. loop Down.→Charles loop	7.5 9.5	**7600** **6000**	1666 1923	**9390** **7760**
Downtown Crossing	Andrew→Broad. loop Charles→Park loop Broad.→Park shuttle	9.5 6.5 13	**6000** 8800 **4400**	1444 1869 1923	**9390** 7760 **8657**
South Station	Andrew→Broad. loop Charles→Park loop Broad.→Park shuttle	9.5 6.5 13	**6000** 8800 **4400**	1444 1869 1923	**9390** 7760 **8657**
Broadway	JFK→Andrew loop Charles → Park loop Park→Andrew shuttle	10.5 5 19.5	**5400** 11500 **2900**	1355 1869 1923	**9298** 7760 **9390**
Andrew	Busing Andrew→ JFK Broad→Andrew loop JFK→Ashmont loop JFK→Braintree loop	 13.5 9 6	 **4200** 6400 9600	1355 1923 573 573	9298 **9269** 3108 6442
JFK/Umass (Ash.)	Alewife→Ash. loop	9	6400	573	3108
Savin Hill	Shawmut→FC loop FC→JFK shuttle	10 13	5700 4400	522 573	2913 3108
Fields Corner	Ashmont→JFK shuttle	21.5	**2600**	573	**3108**
Shawmut	Single loop	16.5	3400	573	3108
Ashmont	No effect			573	3108
JFK/UMass (Bra.)	North Qui.→JFK loop	6	9600	573	6442
North Quincy	QC.→Wollaston loop Woll.→JFK shuttle	7.5 22	7600 **2600**	450 573	5438 **6442**
Wollaston	NQ.→ Braintree loop NQ.→JFK shuttle	16.5 12	**3400** **4800**	573 573	**6442** **6442**
Quincy Center	Wol.→ Braintree shuttle Woll.→JFK shuttle	22 24	**2600** **2400**	450 573	**5438** **6442**
Quincy Adams	Single loop	12	**4800**	573	**6442**
Braintree	Single loop	4	14400	573	6442

Table 2- 4 Major Station Disruption on the MBTA Red Line Southbound

Location of Blockage	Southbound				
	Optimal Strategy	Feasible Headway	Feasible Capacity	Off-peak Demand	Peak Demand
Alewife	Davis → Alewife loop	4	14400	2091	9269
Davis	Central→Harvard loop	6	9600	2091	9269
	Harvard→Ale. shuttle	18	**3200**	730	**5968**
Porter	Central→Harvard loop	6	9600	2091	9269
	Harvard→Ale. shuttle	18	**3200**	730	**5968**
Harvard	Harvard→Por. loop	11.5	**5000**	2091	**9269**
	Porter→Ale. shuttle	12.5	4600	534	4367
Central	Park → Kendall loop	11	**5200**	2091	**9269**
	Kendall → Ale. loop	12.5	**4600**	1985	**8572**
Kendall	Park → Charles loop	7	**8200**	2091	**9269**
	Harvard→Central loop	11	**5200**	1679	**7613**
	Cent.→Charles shuttle	12.5	**4600**	2047	**8572**
Charles	Park→Charles loop	11	**5200**	2091	**9269**
	Central→Charles loop	12	**4800**	2047	**8572**
Park Street	South→Charles loop	16	**3600**	2091	**9269**
	Central→Charles loop	14.5	**3900**	2047	**8572**
Downtown Crossing	Andrew→Broad. loop	9.5	**6000**	1512	**9100**
	Charles→Park loop	5	11500	2091	8572
	Broadway→Park shuttle	15	**3800**	1821	**9269**
South Station	Charles→Down. loop	7	**8200**	2091	**8572**
	Broadway→Down. loop	17.5	**3200**	1821	**9269**
Broadway	JFK→Broadway loop	14.5	**3900**	1444	**9100**
	Down.→Broadway loop	15	**3800**	2091	**9269**
Andrew	Down. → Broad. loop	13	**4400**	2091	**9269**
	Broad.→Andrew loop	13	**4400**	1512	**8642**
	JFK→Braintree loop	10	5700	605	4979
JFK/Umass (Ash.)	JFK→Ashmont shuttle	21	2700	531	2591
Savin Hill	JFK→Ashmont shuttle	21	2700	531	2591
Fields Corner	JFK→Ashmont shuttle	21	2700	531	2591
Shawmut	JFK→Ashmont shuttle	21	2700	531	2591
Ashmont	JFK→Ashmont shuttle	21	2700	531	2591
JFK/UMass (Bra.)	North Qui.→JFK loop	6.5	8800	605	4979
North Quincy	QC.→Wollaston loop	7.5	7600	501	4136
	Woll.→JFK shuttle	22	**2600**	605	**4979**
Wollaston	NQ.→ Braintree loop	16.5	**3400**	501	**4136**
	NQ.→JFK shuttle	12	**4800**	605	**4979**
Quincy Center	Wol.→ Braintree shuttle	22	**2600**	501	**4136**
	Woll.→JFK shuttle	23	**2500**	605	**4979**
Quincy Adams	Single loop	11	5200	605	4979
Braintree	Single loop	4	14400	605	4979

Table 2- 5 Major Inter-Station Disruption on the MBTA Red Line Northbound

Location of Blockage	Northbound				
	Optimal Strategy	Feasible Headway	Feasible Capacity	Off-peak Demand	Peak Demand
Alewife-Davis	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 **3200**	1923 662	9390 **3562**
Davis-Porter	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 **3200**	1923 662	9390 **3562**
Porter-Harvard	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 **3200**	1923 662	9390 **3562**
Harvard-Central	Kendall→Central loop Central→Ale. shuttle	15 23.5	**3800** **2400**	1923 662	**9390** **4964**
Central-Kendall	Park → Charles loop Harvard→Central loop Cent.→Charles shuttle	7 11 12.5	**8200** 5200 **4600**	1923 1491 1710	**9390** 4964 **7760**
Kendall-Charles	Park → Charles loop Harvard→Central loop Cent.→Charles shuttle	7 11 12.5	**8200** 5200 **4600**	1923 1491 1710	**9390** 4964 **7760**
Charles-Park Street	Downtown→Park loop Charles→Park loop	12 18.5	**4800** **3100**	1923 1869	**9390** **7760**
Park-Downtown	South→Down. loop Down.→Charles loop	9.5 7	**6000** 8200	1666 1923	**9390** 7760
Downtown-South Station	Andrew→Broad. loop Charles→Down. loop Broad.→Down. shuttle	9.5 12 10.5	**6000** **4800** **5400**	1444 1923 1666	**9390** **7760** **8657**
South Station -Broadway	Andrew→Broad. loop Charles→Down. loop Broad.→Down. shuttle	9.5 12 10.5	**6000** **4800** **5400**	1444 1923 1666	**9390** **7760** **8657**
Broadway-Andrew	JFK→Andrew loop Charles → Park loop Park→Andrew shuttle	10.5 5 19.5	**5400** 11500 **2900**	1355 1869 1923	**9298** 7760 **9390**
Andrew - JFK/UMass	Busing Andrew→ JFK Broad→Andrew loop JFK→Ashmont loop JFK→Braintree loop	 13.5 9 6	 **4200** 6400 9600	1355 1923 573 573	9298 **9269** 3108 6442
JKF-Savin Hill	Shawmut→Savin loop Savin→JFK shuttle	16 7.5	3600 7600	522 573	2913 3108
Savin Hill - Fields Corner	Shawmut→FC loop FC→JFK shuttle	10 13	5700 4400	522 573	2913 3108
Fields Corner - Shawmut	Ashmont→JFK shuttle	21.5	**2600**	573	**3108**
Shawmut-Ashmont	Ashmont→JFK shuttle	21.5	**2600**	573	**3108**
JFK - North Quincy	JFK→Wollaston shuttle Wollaston→Bra. loop	24 7.5	**2400** 7600	573 450	**6442** 5438
NorthQuincy – Wollaston	JFK→NorthQui. loop Wollaston→Bra. loop North.→Wol. shuttle	5 7.5 7.5	11500 7600 7600	573 450 573	6442 5438 6442
Wollaston -Quincy Center	Wol.→ Braintree shuttle Woll.→JFK shuttle	22 24	**2600** **2400**	450 573	**5438** **6442**
Quincy Center - Quincy Adams	Wollaston.→ QC loop QC→Braintree shuttle	5.5 18	10400 **3200**	573 450	6442 **5438**
Quincy Adams– Braintree	QC → QA loop QA → Braintree	4 12	14400 4800	573 206	6442 3055

Table 2- 6 Major Inter-Station Disruption on the MBTA Red Line Southbound

Location of Blockage	Southbound				
	Optimal Strategy	Feasible Headway	Feasible Capacity	Off-peak Demand	Peak Demand
Alewife-Davis	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 **3200**	2091 730	9269 **5968**
Davis-Porter	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 **3200**	2091 730	9269 **5968**
Porter-Harvard	Central→Harvard loop Harvard→Ale. shuttle	6 18	9600 **3200**	2091 730	9269 **5968**
Harvard-Central	Kendall→Harvard loop Harvard→Ale. shuttle	21 18	**2700** **3200**	2091 730	**9269** **5968**
Central-Kendall	Park → Charles loop Harvard→Central loop Cent.→Charles shuttle	7 11 12.5	**8200** **5200** **4600**	2091 1679 2047	**9269** **7613** **8572**
Kendall-Charles	Park → Charles loop Harvard→Central loop Cent.→Charles shuttle	7 11 12.5	**8200** **5200** **4600**	2091 1679 2047	**9269** **7613** **8572**
Charles-Park Street	Down.→Charles loop Charles→Central loop	16 14.5	**3600** **3900**	2091 2047	**9269** **8572**
Park-Downtown	South→Park loop Park→Charles loop	12 4.5	**4800** 12800	1821 2091	**9269** 8572
Downtown-South Station	Andrew→Down.. loop Charles→Down. loop	17.5 12	**3200** **4800**	1821 2091	**9269** **8572**
South Station -Broadway	Andrew→Down.. loop Charles→Down. loop	17.5 12	**3200** **4800**	1821 2091	**9269** **8572**
Broadway-Andrew	JFK→Broadway loop Down.→Broadway loop	14.5 15	**3900** **3800**	1444 2091	**9100** **9269**
Andrew - JFK/UMass	Busing Andrew→ JFK Broad→Andrew loop JFK→Ashmont loop JFK→Braintree loop	 13.5 9 6	 **4200** 6400 9600	1286 2091 531 605	8300 **9269** 2591 4979
JKF-Savin Hill	JFK→Ashmont shuttle	21	2700	531	2591
Savin Hill - Fields Corner	JFK→Ashmont shuttle	21	2700	531	2591
Fields Corner - Shawmut	JFK→Ashmont shuttle	21	2700	531	2591
Shawmut-Ashmont	JFK→Ashmont shuttle	21	2700	531	2591
JFK - North Quincy	JFK→Wollaston shuttle Wollaston→Bra. loop	24 7.5	**2400** 7600	605 501	**4979** 4136
NorthQuincy – Wollaston	JFK→NorthQui. loop Wollaston→Bra. loop North.→Wol. shuttle	5 7.5 7.5	11500 7600 7600	605 501 605	4979 4136 4979
Wollaston - Quincy Center	Wol.→ Braintree shuttle Woll.→JFK shuttle	22 23	**2600** **2500**	501 605	**4136** **4979**
Quincy Center -Quincy Adams	Wollaston.→ QC loop QC→Braintree shuttle	6.5 16.5	8800 3400	605 214	4979 2636
Quincy Adams – Braintree	QC → QA loop QA → Braintree	4 11.5	14400 5000	605 117	4979 1660

Table 2- 7 Line Volume on the MBTA Red Line

	Off-peak Demand		Peak Demand	
Station	NB	SB	NB	SB
Alewife		283		1880
Davis	245	534	2215	4367
Porter	464	730	3562	5968
Harvard Square	662	1679	4964	7613
Central Square	1491	1985	6886	8572
Kendall/MIT	1710	2047	7760	7887
Charles/MGH	1869	2091	7301	7558
Park Street	1923	1622	6918	6473
Downtown Crossing	1666	1821	6294	8448
South Station	1651	1586	8657	9269
Broadway	1444	1512	9390	9100
Andrew	1355	1391	9298	8642
JFK/UMass	1239	1286	8753	8300
Savin Hill	573	531	3108	2591
Fields Corner	522	407	2913	1841
Shawmut	350	350	2164	1603
Ashmont	326		1897	
North Quincy	573	605	6442	4979
Wollaston	450	501	5438	4136
Quincy Center	383	214	4313	2636
Quincy Adams	206	117	3055	1660
Braintree	99		1264	

1. In all off-peak disruptions, we can provide the needed capacity by using a single-track operation plan. We conclude that single-track operation is indeed useful in dealing with major disruptions, especially during the off-peak.

2. There are however many cases where we may need to call for substitute buses. Because we chose the maximum line volume in the AM and PM peaks, the peak demand is usually very high and makes the single-track operation often appear impracticable. However, given our very conservative assumptions about the travel speed and hand crossover setting time, we believe that we may obtain shorter headway and thus higher capacity in reality.

3. When a disruption happens on one branch, it will not influence the operation on the other branch. Therefore, we can always operate a full loop, which connects the trunk portion and one branch on the Red Line, with a disruption occurring on the other branch. The passengers on the affected branch can get on the trains at JFK, where these two branches merge.

4. Theoretically, we know that if we can organize two loops and one shuttle to serve the system, it will give us the largest possible capacity. In this strategy, these three elements can operate largely independently. Thus, the headway will not increased due to interference. However, it is very hard to find this situation because of the critical crossover configuration needed for this strategy. When a disruption occurs on the track between North Quincy and Wollaston southbound, we can run two loops and one shuttle, which will supply enough capacity to all sections of the line.

5. It is more difficult to deal with station disruptions, since we lose the tracks at both ends of the station, as well as one platform. For example, we can deal with a disruption on the track between North Quincy and Wollaston northbound readily, even in the peak period. However, we are not able to find a feasible strategy to deal with a disruption, at either Wollaston or North Quincy during the peak period.

6. We do have some very difficult situations, particularly when disruptions happen on the trunk portion of the Red Line. The most difficult parts of the Red Line to deal with major disruptions are around Harvard, Broadway and Andrew, because of the lack of crossovers around these stations. When a disruption happens at Andrew, busing is inevitable. Generally we can operate buses between Andrew and JFK to serve the system.

7. When we must use a hand-powered crossover, it causes long headways and low capacity, since we assume that the time to throw a hand-power crossover is 5 minutes.

8. Single-track operation also has disadvantages:

- It increases the number of transfer passengers. Some passengers may have to transfer twice to reach their destination.

- It increases the number of passengers on the platform, especially when just one platform is used. Moreover, when we run shuttles, the entire train will be emptied at each terminal station and thus the platforms may become very crowded.

- It increases the passenger confusion, since we may change the direction of the train, and/or the function of the platform.

From the above analyses, we found that the single-track operation can be helpful in terms of supplying enough capacity, saving operation cost, and reducing passenger waiting time in the off-peak. Given the track and crossover configuration, we want to operate trains wherever we can supply enough capacity and thus avoid busing. We will prefer those strategies that do not involve hand-powered switches, long reverse running sections and low speed crossovers. Tables 2-3 through 2-6 might be considered as a menu to help dispatcher choose the optimal strategy in the case of major disruptions, although extensive further discussions with the MBTA and subsequent refinement of the strategies would be required first.

2.3 Minor Disruption

Incidents lasting less than 20 minutes, are classified as minor disturbances, and are typically caused by a disabled train, door jam or malfunctioning signal. Minor disturbances occur more frequently than major disruptions. In the peak periods when the Red Line is operating close to capacity, a relative minor disturbance can lead to serious degradation in system performance if appropriate control actions are not taken immediately. The dependence of a train's dwell time on its preceding headway causes the long headways ahead of the blockage to lengthen further, while the short headways following the blockage are further shortened. Therefore, even after the disturbance has been cleared, the problem is amplified in an uncontrolled, or poorly controlled, system.

2.3.1 Objective and Methodology

When there is a minor disturbance, we usually will not lose any part of track or station for an extended period as in major disruption. Thus we do not need to change the operation plan, even though the schedule has to be adjusted in real-time to minimize the impacts of the disruption. However, passengers will inevitably suffer from longer waiting times and more crowded trains, thus reducing the quality of service. Therefore, the primary objective in minor disturbance control is to maximize the transit service quality given the disturbances.

Because service quality can be measured in several different ways, and some aspects even conflict with each other, it is very difficult to choose a single objective that will satisfy all passengers and all transit agencies. Generally, we can choose among the following options:

1. Passenger waiting time

   Obviously, the passenger waiting time is one of the most important measures of transit service quality. Passengers are very sensitive to waiting time especially in bad weather or with unsafe waiting areas. Passenger waiting time might also be the easiest measure of service quality to estimate. Therefore, most papers on public transportation service quality control choose minimizing passenger waiting time as the objective.

2. On time performance

On time performance is another important measure, which is often chosen by transit agencies to evaluate their performance. It also can be easily calculated. However, passengers of high-frequency transit system might not be sensitive to it, since they are less likely to pay attention to the schedule because of the high frequency nature of the service.

3. The number of crowded trains

   Based on discussions with MBTA staff, there is a strong relationship between the number of crowded trains and the number of complaints the transit agency receives, which also reflects the perceived service quality.

4. The number of affected passengers

   The number of affected passengers is defined as the total number of passengers who are delayed by the disturbance. Intuitively, we want to minimize the number of affected passengers, and the extent of impact. However, sometimes we have to sacrifice some passengers' benefits to get the best overall solution.

There are still other measures, such as the number of passengers who are left at the station, or the number of complains that transit agencies receive. To decide on the objective, we have to study the passengers' behavior carefully, and choose one appropriate measure or a combination of several measures.

Eberlein (1995) and O'Dell (1997) chose minimizing the passenger waiting time as the objective of their models, which represent the basis of our current research. First they analyzed the holding and short-turning strategies using a simple, idealized system model for which they obtained closed-form results, in order to gain a better understanding of the problem. Many of the simplifying assumptions were then relaxed to develop mathematical programming formulations for a more realistic, generalized model. Finally, O'Dell tested her model, which was developed based on Eberlein's work, on several problem instances using data from the MBTA Red Line. Their results of the holding, expressing, and deadheading control strategies provide the basis for the following discussion.

2.3.2 Systems Description

In this section, we will describe the general system, which is used in O'Dell's thesis to analyze control strategies for minor disturbances. This general model has the following features:

1. The transit system can be a two-branch system, rather than a single-loop system, to accommodate the configuration of the MBTA Red Line, which has two branches.

2. Dwell time is a linear function of the total alighting and boarding passengers.

3. Passenger arrival rate, and alighting fraction of the load of train are station-specific parameters.

4. Trains have capacity constraints so that passenger may be left at some stations in the case of very crowded trains.

5. A train will not depart from a station until it can travel to the next station at the free-running speed.

2.3.3 Control Strategies

Generally, there are four main control strategies that the transit agencies can choose. They are holding, expressing, deadheading and short turning. Among these strategies, both deadheading and expressing are restricted by the location of the preceding train and the control lines. The benefits in terms of passenger waiting time savings by employing either of these two strategies will be limited by this restriction. Therefore, O'Dell focused on the other two strategies, holding and short turning as described below.

I. Holding

Holding is the easiest strategy to employ among all those available to transit agencies. Therefore, it is also the strategy most frequently implemented in real time operations control of transit systems. The core idea behind holding is that a train may be held at one or more control stations for a time, even though it is ready to depart from those stations. The decision variables are the location and duration of the holding strategy being employed. The objective of holding is to even out a long and short headway sequence between these trains.

There are several advantages of holding, compared to other control strategies.

1. It is easy for transit controllers to execute. Dispatchers can use phone, radio, or signal to inform the train operator of the holding action.

2. Since holding does not change the operations plan substantially, it will minimize the confusion to both operators and passengers.

3. Unlike short turning, expressing, and deadheading, holding does not result in any station being skipped. This can reduce the frustration of passengers who might be passed by a train even after a relatively long wait.

In the general transit system, we can not find closed form results due to the complexity of the problem. We can however use mathematical programming to get the optimal result. In O'Dell (1997) model, she used the departing time as decision variable, and chose minimizing passenger waiting time as the objective. Therefore, her objective function resulted in a large quadratic mixed integer formulation. Intending to use her model in the real-time, O'Dell developed a piecewise linear approximation for this function. She used the Cplex software package in two case studies on the MBTA Red Line, and supplied valuable insights into the holding strategy.

Obviously, consideration of a larger number of trains to control will increase the size of the problem but will in general give a better result. However, O'Dell only considered a limited set of trains and stations, which were in front of and behind the blockage at the time of the disturbance, as the "impact set". O'Dell's finding on the appropriate impact set size included:

1. Larger numbers of trains and stations will increase the amount of computation, and jeopardize the ability to produce real-time decisions.

2. There is no need to consider a very large number of trains and stations. Because we only consider minor disturbances, the duration of the blockage is relatively short and the number of impacted trains and stations are also small.

3. Based on the existing signal and communication systems on rail systems such as the MBTA Red Line, it is not recommended from a technical standpoint to attempt to control a large number of trains.

O'Dell defined three alternative holding strategies.

1. "Hold All"

In this strategy, we can hold any train at any station in the impact set. Obviously, this will be the most effective strategy in a narrow sense, because it is the least constrained. However, there are some technical and feasibility concerns in executing this strategy in the real world. For example, how easy will it be to hold a train or control a train's departure time at several intermediate stations.

2. “Hold at First”

This is the simplest holding strategy in which we can hold a train only at the next station in the “impact set” after the blockage occurs.

3. “Hold Once”

In this strategy, any train in the impact set can be held no more than once at an optimally chosen station.

Results of the mathematical programming solution show that the holding strategy can be very effective and reduce passenger waiting time by 15-40% compared with “do nothing” case. Moreover, “Hold at First” and “Hold Once” are virtually as effective as “Hold All” for most of the cases tested. Therefore, O’Dell recommended using the “Hold at First” strategy, since it is the easiest one to implement and it does not seriously compromise overall effectiveness.

II. Short turning

Short turning is another useful control strategy that is often employed by transit agencies. However, it is restricted by the availability, configuration and ease of use of crossovers and so can not be employed everywhere. In the normal operating situation, the average running speeds of trains based on the control lines are in the range of 20-40 mph. When we plan to short turn a train, we have to use reverse track operation, where the maximum running speed is 25mph. Generally, it may take about 6 minutes to short turn a train from the current platform to the opposite platform. Therefore, the short turning strategy is only appropriate when we have longer disturbances or unusual circumstances.

Suppose that there is a blockage that causes a delay as in Figures 2-9 and 2-10.

Figure 2- 9 Short-turning Behind the Blockage

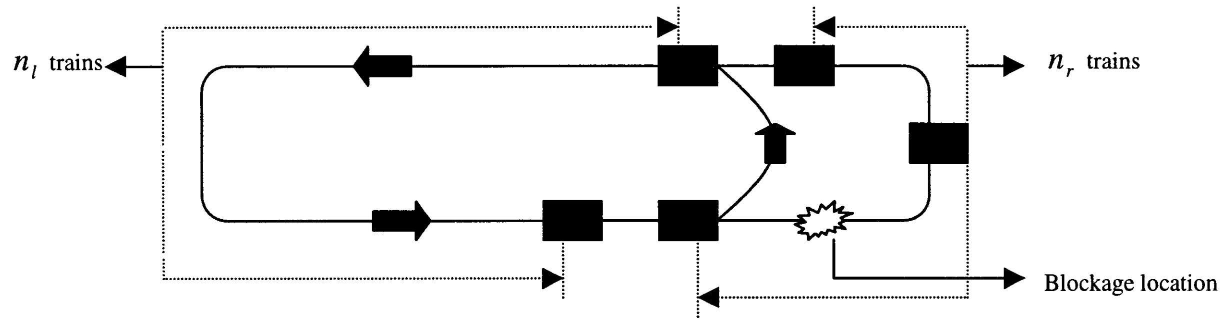
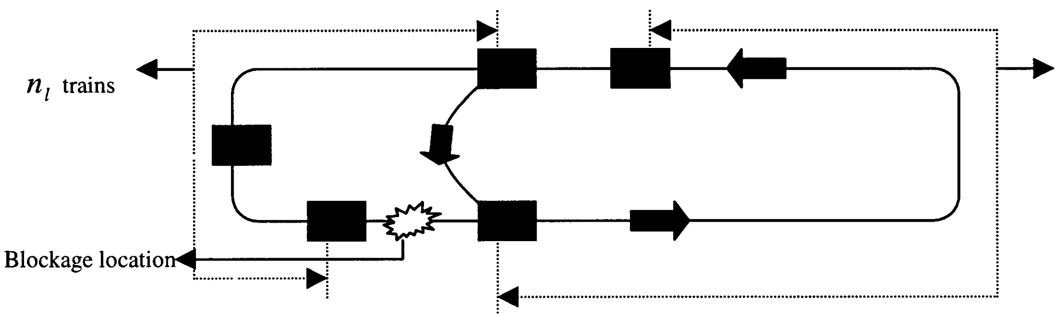


Figure 2- 10 Short-turning in Front of the Blockage



We can short turn trains behind, or in front of, the blockage as shown in the above figures, and the decision is based on the tradeoff between the waiting time savings in the fully served areas versus waiting time increases in skipped areas. Usually, we will short turn a train so that this train can serve the heavy passenger flow direction. We may also short turn a train when the skipped section is a small portion of the entire trip.

In figure 2-9, we run n_l trains in the left short-turning loop, while holding n_r trains at the right part of track during the blockage. In this strategy, we reduce the waiting time of passengers in the left loop and increase the waiting time of passengers in the right part. In the generalized system, it is impossible to find closed form results for the optimal control strategy. However, as with the holding strategy, O'Dell developed a series of mathematical programming formulations to solve the combined short turning and holding problems, given that a specific train and short-turn location had been identified.

We have to hold some other trains, while we short turn a train in the designated short-turning loop. Therefore, there is no pure short-turning strategy. O'Dell's short-turning mathematical programming model was based on combining short turning with the "Hold All" strategy. O'Dell also assumed that the train order after short turning would be predetermined, and discussed how it could be extended to the undetermined order problem.

O'Dell's results show that the short turning strategy, which is combined with the holding strategy can bring up to 50% passenger waiting time savings. It is much more beneficial in the situation when the duration of the blockage is 20 minutes (or greater.) This result reaffirmed the conclusion from the simplified system.

2.3.4 Conclusion

From the above analysis, we find that we can use mathematical programming to obtain optimal solution of the control strategy. However, it has three disadvantages which suggest investigation of other approaches as well.

1. Mathematical programming methods are very complicated and restrictive, especially for transit agency managers. Transit agencies usually do not have the human and software resources to feel comfortable employing mathematical programming. Furthermore such simple single objectives as minimizing passenger waiting times may not be acceptable to many transit managers. Moreover, the objective function of mathematical programming is not flexible. It is like a black box, and dispatchers only can accept or reject the solution without any real ability to change it. The model may need to be redeveloped when a new objective is chosen. Sometimes, it will not even be possible to use mathematical programming to solve larger and more complicated problems.

2. The computation time required can be unpredictable. In the case of O'Dell's model, it took from 20 to 2458 seconds to compute a solution. This is not reliably fast enough to be part of a real-time decision support system.

For both these reasons, we believed that a heuristic model, which is based on the rules derived from the mathematical programming models, might be more acceptable to transit agencies as the basis for effective and implementable control decision support tools.

We will discuss the dispatching problem, one specific part of a general heuristic disruption control system, in the remainder of this thesis.

Chapter 3

Dispatching Control Problem

In this chapter, I first introduce the characteristics of the dispatching problem, as a special case of the minor disruption problem. Then I formulate the dispatching problem mathematically, and get some insight into the dispatching problem as a basis for the heuristic dispatching control model. Finally, I will discuss some feasible strategies to deal with the dispatching problem on the MBTA Red Line.

3.1 Dispatching Problem

When delays occur at, or around, the terminals, the normal dispatching operation at the terminus will need to be modified. One typical problem is that a train might not arrive at the terminal early enough to be dispatched on schedule. This problem, often referred as the dispatching problem, is critical to the quality of the transit service since it causes uneven headways and increases passenger waiting times, especially when the dispatching direction is the heavy passenger flow direction. Because dispatching problems are often caused by minor disruptions, the dispatching problem can be treated as a special case of the minor disturbance problem. However, the dispatching problem also has the following special characteristics:

1. The dispatch headway is a critical determinant of service quality along the route, as shown by Eberlein (1995). It is particularly important to get the headway right in the heavy passenger flow direction.

2. Because there is a scheduled departure time at the terminal for every train, we can identify the problem most easily by comparing the estimated arrival or departure time with the scheduled time. For the intermediate stations, we have to monitor every train's movement and find

whether there is a delay or disturbance, which can be complicated, since there is no scheduled time at these stations.

3. Usually, there are more comprehensive facilities available at the terminal, which make passengers more comfortable during their wait. These facilities, such as long benches, vending machines, etc. will reduce the sensitivity of passengers to waiting.

4. There is usually an inspector at the terminal. They can help dispatchers or operators to execute the control actions more effectively. Staying on the platform, they can also guide the passengers and reduce confusion.

5. When delays occur around a terminal, it makes it easier for transit agencies to employ control strategies. Every train has a recovery time at each terminal as well as the dwell time. This recovery time can be adjusted to enable trains to follow the schedule. When a train arrives at a terminal earlier than scheduled, we just extend the normal recovery time of this train; when a train arrives at a terminal later than scheduled, we reduce the normal recovery time. Obviously, using recovery time to adjust the operation will be constrained by the minimum recovery time, and is only useful when the delay is short.

When the dispatcher decides that it is appropriate to dispatch a train from the terminal, he will send a signal to that terminal; this is called the ring-off. The train is supposed to leave the terminal as soon as possible after the ring off time. Therefore, the ring off time can be used to control the operation. When the dispatcher wants to hold a train, he can just change the ring off time for the train to be held. The held trains will simply have a longer recovery time, and passengers and operators are less likely to notice the change of the scheduled dispatching time and this should minimize the possible frustration.

When the dispatcher chooses an expressing or deadheading strategy, he only need inform the passengers at the terminal. He does not need to inform the passengers in the train and this can avoid confusion in the train, which can result when expressing starts at an intermediate station.

Based on the nature of the dispatching problem, we believe that the solution to it will also shed light on the solution to the general minor disruption problem. In the following sections, we will analyze the dispatching problem and derive optimal headways based on mathematical analysis.

3.2 Mathematical Analysis

3.2.1 Notation

We will introduce some notation that will be used in the rest of this chapter.

$H_{i,k}$: departure headway of train i from station k;

$D_{i,k}$: dwell time of train i at station k;

$T_{i,k}$: travel time of train i from station k-1 to station k;

P_{off}: number of alighting passengers;

P_{on}: number of boarding passengers;

c_k, l_k: coefficients in the running time function;

$T_{open+close}, \alpha, \beta$: coefficients in the dwell time function;

t_i: departure time of train i;

$a_k(t)$: passenger arrival rate at station k at time t (a_k is used when the passenger arrival rate is constant);

$F(t_i)$: equivalent cumulative passenger arrival when train i departs from the terminal at t_i;

$f(t_i)$: equivalent passenger arrival rate at time t_i, $f(t_i) \equiv dF(t_i)/dt$

The following points should be noted:

The dwell time function will be discussed in chapter 4; however the general function is assumed to be of the following form:

$$D_{i,k} = T_{open+close} + \alpha \cdot P_{off} + \beta \cdot P_{on}$$

The running time between consecutive stations is affected by the location of the preceding train, which can be measured by the headway. When the headway is small, the train has a lower permitted speed than in the normal situation based on the minimum safe stopping distance. Therefore, the running time is usually larger in short headway situations. While the headway is large, the train can run at a higher maximum permitted speed. Therefore, the running time will be shorter. We can approximate the running time as follows:

$$RT(i,k) = c_k + \frac{l_k}{H_{i,k-1}}$$

Where c_k is the minimum running time.

Passengers may board at any station along the route. It is difficult to estimate the equivilent cumulative passenger arrivals or equivalent passenger arrival rate at the terminal because of the variable travel time between consecutive stations. Therefore, we have to simplify the variable travel time, and assume that it is independent of the alighting and boarding passengers. Since the travel time between stations is assumed constant, we can use the following equation to get the equivalent passenger arrival rate at the terminal.

Suppose that tt_k is the travel time from the terminal to station k, we can convert any trip at station k at time t to the equivalent trip from the terminal at time $t - tt_k$. Therefore, we can compute the equivalent passenger arrival rate at the terminal using the following equation:

$$f(t) = \sum_k a_k (t + tt_k)$$

3.2.2 Headway

We will assume that the dwell time functions for every station are identical. Therefore, the headway can be represented by the following equation:

$$H_{i,k} = H_{i,k-1} + (D_{i,k} - D_{i-1,k}) + (RT(i,k) - RT(i-1,k))$$

When we consider the dispatching problem, we only deal with the stations that are close to the terminal. It is quite likely that there are not a significant number of alighting passengers at these stations. In other words, P_{off} can be approximated as 0.

$$D_{i,k} = \alpha + \beta \cdot P_{off} + \gamma \cdot P_{on} \quad \Leftrightarrow \quad D_{i,k} = \alpha + \gamma \cdot P_{on} = \alpha + \gamma \cdot a_k \cdot H_{i,k}$$

Because the dwell time functions are identical, the coefficients are the same. Thus:

$$D_{i,k} - D_{i-1,k} = \gamma \cdot a_k \cdot (H_{i,k} - H_{i-1,k})$$

We can also get the difference in running times of trains i and i-1:

$$RT(i,k) - RT(i-1,k) = k_k \left(\frac{1}{H_{i,k-1}} - \frac{1}{H_{i-1,k-1}}\right) = k_k \cdot \frac{H_{i-1,k-1} - H_{i,k-1}}{H_{i,k-1} H_{i-1,k-1}}$$

From the above equations, we can get the following result:

$$H_{i,k} - H_{i,k-1} = \gamma \cdot a_k \cdot (H_{i,k} - H_{i-1,k}) + (RT(i,k) - RT(i-1,k))$$

$$\Leftrightarrow H_{i,k}(1 - \gamma \cdot a_k) = H_{i,k-1} - \gamma \cdot a_k \cdot H_{i-1,k} + (RT(i,k) - RT(i-1,k))$$

$$\Leftrightarrow H_{i,k} = (H_{i,k-1} - \gamma \cdot a_k \cdot H_{i-1,k} + (RT(i,k) - RT(i-1,k)))/(1 - \gamma \cdot a_k)$$

$$H_{i,k} - H_{i,k-1} = (\gamma \cdot a_k \cdot (H_{i,k-1} - H_{i-1,k}) + (RT(i,k) - RT(i-1,k))/(1 - \gamma \cdot a_k)$$

If we want to find the situation where the headway of the current train increases at the following station, that is $H_{i,k} > H_{i,k-1}$:

$$H_{i,k} - H_{i,k-1} \geq 0 \Leftrightarrow \gamma \cdot a_k \cdot (H_{i,k-1} - H_{i-1,k}) + (RT(i,k) - RT(i-1,k) \geq 0$$

$$\Leftrightarrow \gamma \cdot a_k \cdot (H_{i,k-1} - H_{i-1,k}) \geq RT(i,k) - RT(i-1,k)$$

In the above inequality, the left hand side is the increase of the train's dwell time at the previous station compared with the preceding train's dwell time at the current station, while the right hand side is the saving in the train's running time. The inequality shows that the headway will increase, if the saving of running time can not compensate for the increase in dwell time. From the above relationship, we find that if we want to ensure that train i's headway is not increasing, we have to constrain its headway at station k-1. When the headway is large enough to avoid any speed restriction, the running time saving will be 0, and train i's headway at station k-1 cannot be larger than the preceding train's headway at the following station if train i's headway is to be non-increasing.

3.2.3 Optimal Headway

When the equivalent cumulative passenger arrival rate is constant, we have the following relationship:

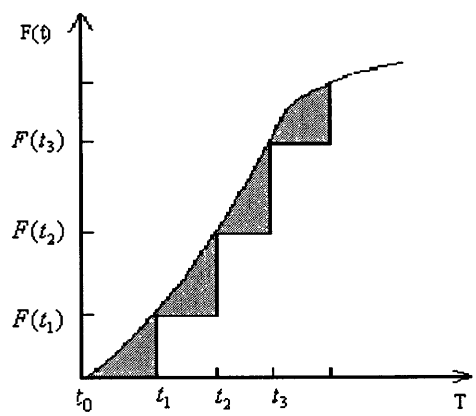
1. Even headways produce the minimum passenger waiting time. This can be easily proven. Because the passenger waiting time is a quadratic function of headway, it is the optimal

solution to divide a fixed time period evenly, that is to have each train depart from the terminal at even headways, in terms of minimizing the passenger waiting time.

2. Unless there is a severe delay, trains departing from the terminal evenly will minimize the probability of passengers experiencing overcrowded trains.

However, there may be some variations in cumulative passenger arrival rate during the peak periods, and the analysis of the variable arrival rate situation, which is based on the Newell's (1971) research, is shown below.

Figure 3- 1 Cumulative Passenger Load



In Figure 3-1, the shaded area between F(t) and the step function with heights $F(t_i)$ for $t_i \leq t \leq t_{i+1}$ is the total passenger waiting time. According to the earlier discussion of rail transit systems, we know that the passengers can be assumed to arrive randomly in any short time interval. Therefore, we have the following relationship for total passenger waiting time:

Total Passenger Waiting Time = Total Shaded Area = $TW = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} [F(t) - F(t_i)]dt$.

Where we initialize $F(t_0) = 0$ and $t_n = T$.

To obtain the optimal headways that minimize total passenger waiting time, we have the following relationships:

$$\frac{\partial TW}{\partial t_i} = F(t_i) - F(t_{i-1}) - f(t_i)(t_{i+1} - t_i) = 0 \Leftrightarrow F(t_i) - F(t_{i-1}) = f(t_i)(t_{i+1} - t_i)$$

$$\Leftrightarrow t_{i+1} - t_i = [F(t_i) - F(t_{i-1})]/f(t_i)$$

Using any t_0 and trial value of t_1, we can sequentially get the optimal times for t_2, t_3, ..., t_n. According to the assumption $t_n = T$, we can finally determine the optimal time for t_1.

If n is sufficiently large, we can use the Taylor series expansion to estimate the value of $F(t_{i-1})$ from $F(t_i)$:

$$F(t_{i-1}) \approx F(t_i) + f(t_i)(t_{i-1} - t_i) + f'(t_i)(t_{i-1} - t_i)^2/2 + 0(t_{i-1} - t_i)$$

1. If the passenger arrival rate $f(t)$ is constant in a time period, $f'(t) = 0$, then $F(t_{i-1}) \approx F(t_i) + f(t_i)(t_{i-1} - t_i)$. We substitute this into the above equation:

   $$t_{i+1} - t_i \approx t_i - t_{i-1} \Leftrightarrow H_{i+1,term.} \approx H_{i,term.}$$

   This result also proves that even headways will bring about minimum passenger waiting time when the equivalent cumulative passenger arrival rate is constant.

2. If the passenger arrival rate $f(t)$ is not constant in a time period, $f'(t) \neq 0$, then $F(t_{i-1}) \approx F(t_i) + f(t_i)(t_{i-1} - t_i) + f'(t_i)(t_{i-1} - t_i)^2/2$. We substitute this into the above equation,

$$t_{i+1}-t_i \approx [1-\frac{(t_i-t_{i-1})f'(t_i)}{2f(t_i)}](t_i-t_{i-1}) = \prod_{j=m}^{i}[1-\frac{(t_j-t_{j-1})f'(t_j)}{2f(t_j)}](t_m-t_{m-1})$$

$$= e^{-\sum_{j=m}^{i}\frac{(t_j-t_{j-1})f'(t_j)}{2f(t_j)}}(t_m-t_{m-1})$$

If $t_j - t_{j-1}$ is sufficiently small, we can replace it by $dt = t_j - t_{j-1}$.

$$t_{i+1}-t_i \approx e^{(-1/2\int_{t_{m-1}}^{t_i} f'(t)/f(t)dt)}(t_m-t_{m-1}) = f^{1/2}(t_{m-1})/f^{1/2}(t_i)(t_m-t_{m-1})$$

$$\Leftrightarrow (t_{i+1}-t_i)f^{1/2}(t_i) = (t_m-t_{m-1})f^{1/2}(t_{m-1}) = \text{constant}$$

The above formula shows that $t_{i+1}-t_i$ is proportional to $f^{1/2}(t_i)$. We conclude that the optimal departure time of trains is approximately proportional to the square root of the arrival rate of passengers. From the above equation, we also find that the variation of the headway is much smaller than the variation of the cumulative passenger arrival rate. Therefore, running even headways might still be a good decision even when the passenger arrival rate is not constant.

While the above analysis is based on the objective of minimizing the passenger waiting times, we can choose different objectives as discussed in the chapter 2. For example, if we want to minimize the number of overcrowded trains, we will have the following relationship:

Probability of overcrowded trains = $\sum_i (P(F(t_i)-F(t_{i-1})) > TrainCapacity)$

$$= \sum_i (P(f(t_i)(t_i-t_{i-1})) > TrainCapacity)$$

$$= \sum (P((t_i-t_{i-1}) > TrainCapacity/f(t_i)))$$

If $f(t_i)$ is constant in a time period, we can always choose $t_i - t_{i-1} < TrainCapacity / f(t_i)$ to minimize the possibility of overcrowded trains, and the right hand side of the above inequality is constant. Therefore, we still can use even headways to minimize the probability of overcrowded trains. If $f(t_i)$ is not constant, we also have to keep $t_i - t_{i-1} < TrainCapacity / f(t_i)$ to avoid the overcrowding. The optimal headway is a reciprocal function of the equivalent passenger arrival rate in this case. However we can still use the maximum $f(t_i)$ to get the threshold value of headway. As long as our headway is smaller than this threshold value, we can run trains at even headways without causing any overcrowded trains.

To summarize our analysis, we have shown that operating even headways is close to the optimal solution, even when the cumulative passenger arrival rate varies cross the time. In our heuristic model, we will use this idea to deal with the dispatching problem.

3.3 Control Strategies for MBTA Braintree Branch Dispatching Problem

Even though the dispatching problems at any terminus have a similar structure, the solutions will differ due to the differences in track and station configuration, signal systems, and passenger flow. In this section, we will discuss the possible strategies based on the characteristics of the Braintree branch of the MBTA Red Line.

1. Track Configuration

   The Braintree branch has 6 stations, beginning with JFK/UMass and running to the Braintree terminus. Besides the Braintree station, there are other three stations, Quincy Adams, Quincy Center and Wollaston, where the trains can be short-turned. Among them, the track configuration at Quincy Center is most conducive for short turning, and almost all short turns are executed at Quincy Center. Therefore, the short turning strategy is constrained to Quincy Center.

2. Passenger Flow

To determine the passenger flow on the Braintree branch, a set of passenger data were collected by the MBTA at the 5 branch stations, excluding JFK/UMass. Each station has 3-hour AM peak period data collected on one weekday. The following figures (Figures 3-2 through 3-6) show the passenger flows on the Braintree branch.

Figure 3- 2 Braintree Passenger Flow

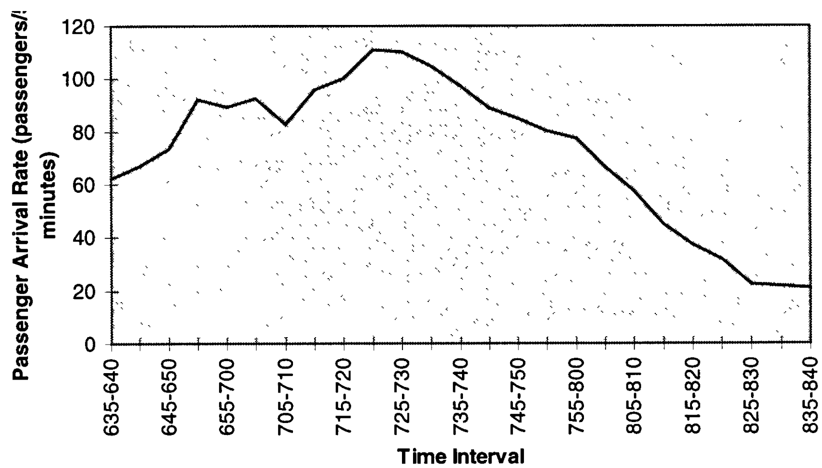


Figure 3- 3 Quincy Adams Passenger Flow

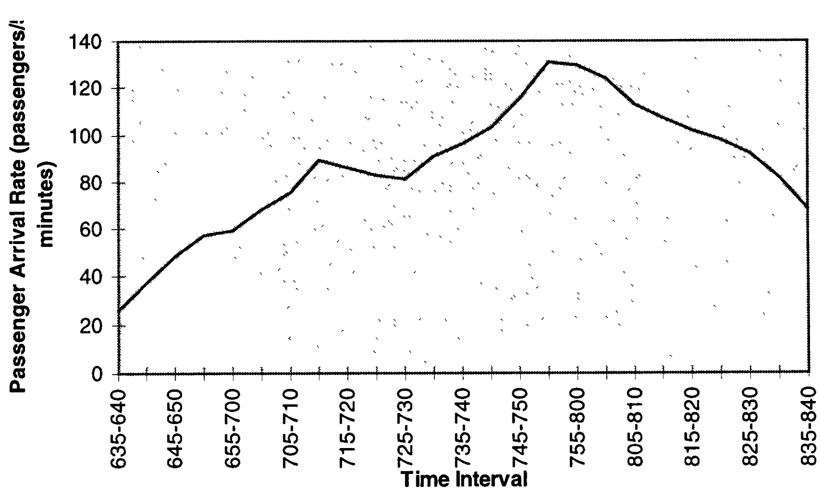


Figure 3- 4 Quincy Center Passenger Flow

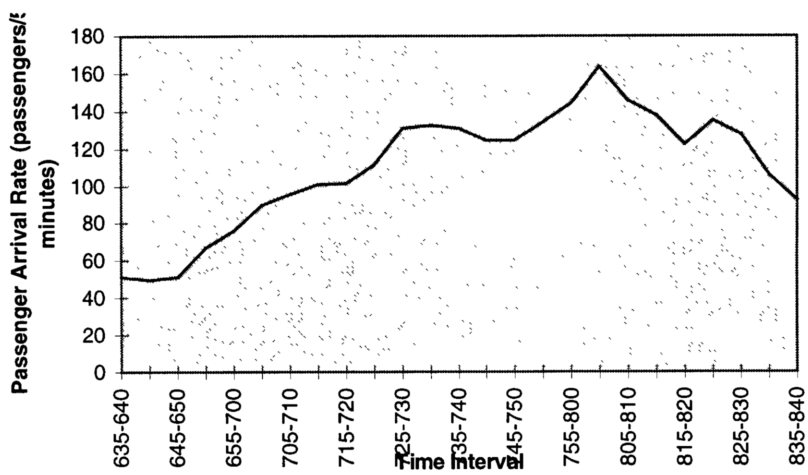


Figure 3- 5 Wollaston Passenger Flow

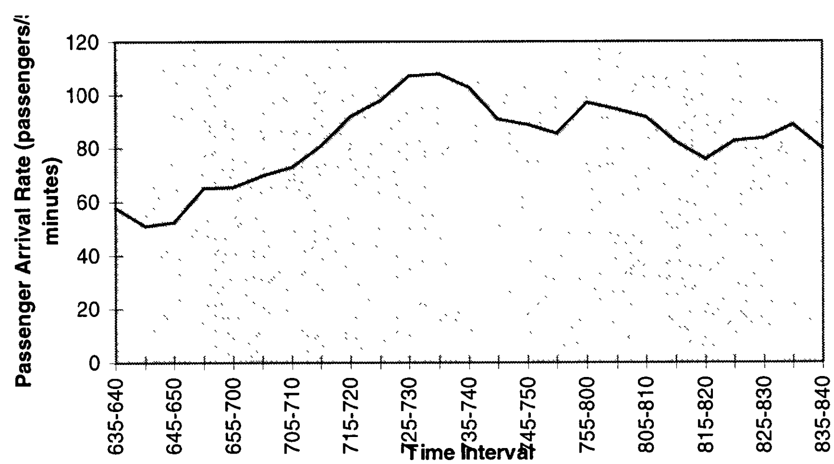
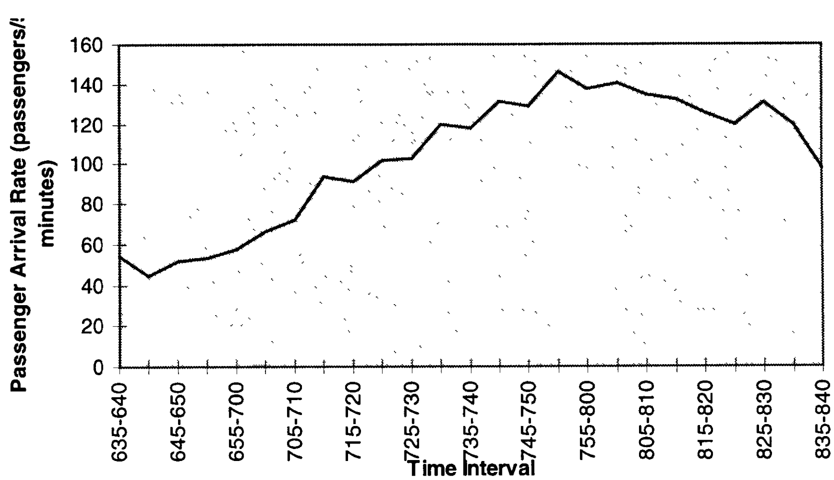


Figure 3- 6 North Quincy Passenger Flow

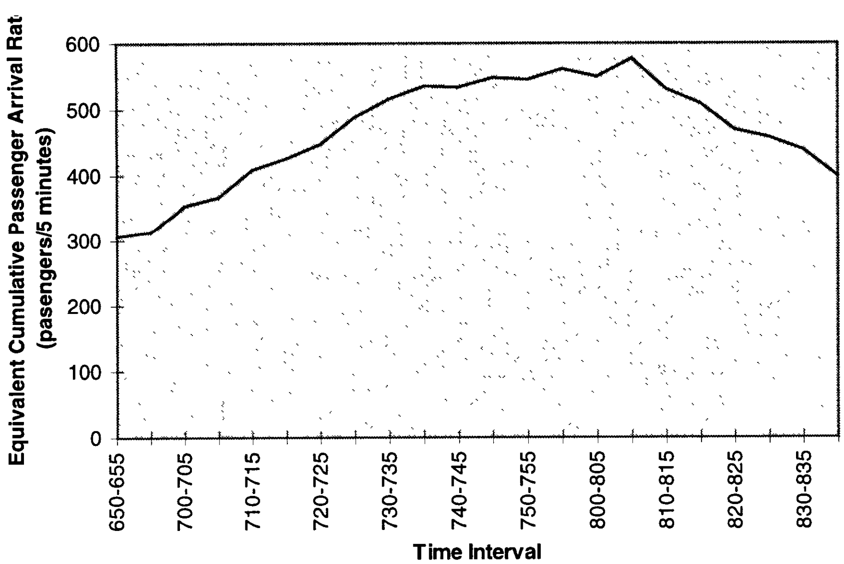


We assume that the travel time from Braintree to each station northbound is constant as shown in Table 3-1, and calculate an equivalent passenger arrival rate at Braintree (see Figure 3-7.) We find that the equivalent passenger flow at Braintree is quite flat during the AM peak half hour. Therefore, the equivalent average arrival rate can be considered to be constant during this period and only drops by about 10 percent in the next heaviest half hour period. Given the square root relationship between the optimal headway and passenger arrival rate, this implies that a constant headway during the peak hour is close to optimal.

Table 3- 1 Travel Time From Braintree to Other Branch Stations

Station	Travel Time (seconds)
Quincy Adams	235
Quincy Center	438
Wollaston	606
North Quincy	801

Figure 3- 7 Equivalent Passenger Arrival Rate at Braintree



Since the highest equivalent passenger arrival rate at Braintree is less than 600 during a 5 minute interval, the average equivalent arrival rate can be treated as 120 passengers/minute. During the peak period, the six-car Red Line train can readily accommodate 960 passengers. Therefore, we can constrain the headway to be less than 960/120=8 minutes during the peak period, which will avoid overcrowded trains and should tend to minimize passenger waiting time.

Chapter 4

Dwell Time Function

When a train arrives at a station, the doors must be opened for a sufficient time to allow all passengers who want to alight and board to do so. This time is called the dwell time.

Due to the high speed of rail transit systems, the running time between stations is relatively short; however, the dwell time can be relatively long particularly if there is high passenger flow. Even though the dwell time still represents less than half of total travel time, its high variation can be a critical element in increasing headway variation. Therefore, the dwell time is very important for real-time control, simulation and line capacity, and is quite worthy of detailed analysis as presented in this chapter.

Focusing on real-time dispatching control on the MBTA Red Line, we believe that the dwell time is one of the key functions in our entire model, and it will also play an important role in the disruption control system. Our objective is to develop a good dwell time function, based on data obtained from the MBTA. We also expect that the same methodology employed here can be used in other situations.

Despite the fact that dwell time is simply the time to let passengers alight and board, it is very difficult to develop a reasonable dwell time function. In fact, the dwell time depends on many complex factors, such as passenger flow, congestion in the train, passenger behavior, doorman behavior, and car type, etc. Among these factors, human behavior alone will result in significant variation in the dwell time.

In this chapter, we first discuss the theory of dwell time, then the data collection and estimation are presented in the next section. Finally, we develop the dwell time functions in the last section.

4.1 Theory

Generally, the dwell time can be separated into three elements:

1) Constant time: this time includes the time to open and close the doors. For a given type of car and train length, this time should be similar across trains and stations (except for some special stations, such as Park Street on the MBTA Red Line, which will be discussed below).

2) The alighting time: this time is for passengers to get off the train.

3) The boarding time: this time is for passengers to board the train.

Obviously, these three elements may not be entirely independent, indeed, they may overlap to some extent. For example, some passengers will be aggressive, and try to board when other passengers are getting off; or because of the distributions of passengers across cars in a train and along the platform, boarding may be occurring at one door while alighting is still occurring through other doors.

As stated previously, the dwell time depends on many factors, which are summarized below:

1. Passenger behavior: passenger behavior is one of the most important factors that affect the dwell time. Some passengers are fast walkers, while others are slow. Some passengers are aggressive, while others are not. This can make a big difference in the dwell time. We also notice that even the same person might have different behavior at different times. For example, a passenger in the peak period may walk faster than at other times. While passenger behavior can affect dwell time, it can not be included in the dwell time model, except in an aggregate sense, and will always result in uncontrolled variance in the model forecast.

2. Doorman behavior: During the peak periods, the MBTA operates six-car trains, whose doors are controlled by a doorman. Therefore, the doorman behavior is also a very important factor

in determining the dwell time. If the doorman is very generous, he will give passengers more time to board or alight. That is, even if no passengers are trying to get on or off, the doorman will still leave the doors open for a while. On the other hand, if the doorman is tough, he or she may close the doors in some passengers' faces.

3. Platform number and configuration: Except at Park Street (see following discussion), the MBTA uses only one platform at every station for each direction of operation. However, the position of the stairways or other aspects of the platform configurations, which may cause different effects on the dwell time, vary across stations. In some stations, passengers will jam around the entrance and make it difficult for other passengers to get off the train.

4. Type of cars: Currently the MBTA uses 1500, 1700 and 1800 series cars on the Red Line. The newer Bombardier cars (also referred as to 1800 series cars) have four doors on each side, while the Silverbird cars (1500 and 1700 series) only have three doors per side. For the same passenger loads, we would expect that the boarding and alighting time for each passenger would be lower for a Bombardier train, reducing the dwell time, because there are more doors to accommodate passenger movement. Therefore, we may expect that the marginal dwell time for a Bombardier train will be about 30% lower than for a Silverbird train, all other factors being equal.

5. Crowding factor includes four types of potential conflicts.

    1) The conflict between the alighting passengers and the passengers staying in the car.

    2) The conflict between the alighting passengers and the passengers on the platform (include the boarding passengers and those passengers who had just got off)

    3) The conflict between the boarding passengers and those passengers already on board.

4) The conflict between boarding passengers and the passengers on the platform.

6. Marginal time for each alighting passenger and boarding passenger. Marginal times for alighting and boarding passengers may be different at different time and under different conditions. In some situations, the marginal time for an alighting passenger may be larger than that of a boarding passenger; in some other conditions, the reverse may be true.

   There are a couple of reasons for this: First, the marginal time is may be influenced by the load in the car or the number of passengers on the platform. When there is a heavy arriving load and the number of passengers on the platform is also very large, the marginal alighting passenger time will usually be large due to passenger conflicts. Second, some boarding passengers try to get on when there are still passenger getting off. This phenomenon not only causes conflicts, but also increases the marginal boarding time.

7. Other control operations: These operations include holding, short-turning or expressing which result in a larger than normal dwell times.

From the above analysis, we will consider including the following explanatory variables in our dwell time function:

1. Number of alighting passengers;

2. Number of boarding passengers;

3. Load in the train on arrival;

4. Type of car;

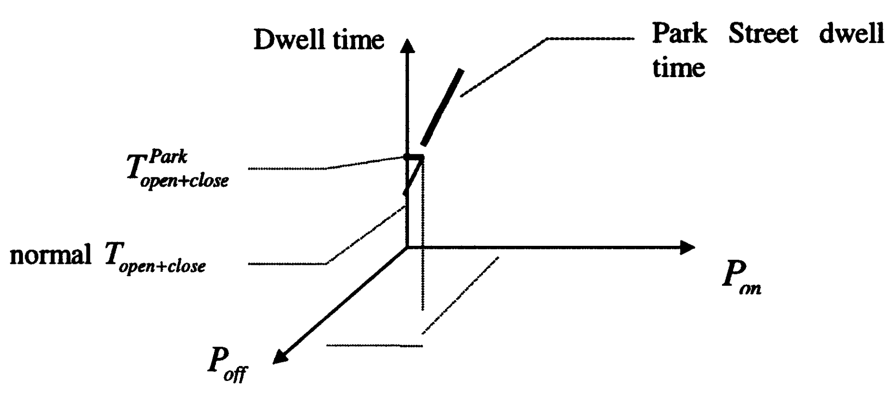
5. Particular station;

6. Any control operation.

Among these explanatory variables, the first and second ones are expected to be the most important, and also the easiest to obtain. The third variable, which is the load in the train, is also important, however, due to the limited time, it is difficult to collect this information accurately. Therefore, we might use a categorical variable here or just approximate the number.

At some special stations, such as Park Street on the Red Line, we have a more complicated situation. At Park Street, doors are opened on the both sides of the trains, and two platforms are used simultaneously. In this case, the constant part in the dwell time will be larger since it takes longer to open and close both sets of doors. However, the time at Park Street will still consist of the same three parts discussed above, even though it may be dominated by either platform. Because the door man has to open and close doors on both sides at Park Street, $T^{Park}_{open+close}$ is larger than $T_{open+close}$ at other stations. During this long open-and-close time, many passengers can get on or off without increasing the dwell time, as shown in Figure 4-1.

Figure 4- 1 Park Street Dwell Time



In the Park Street case,

the P_{on} will be: 1) number of boarding - A, if number of boarding > A

2) 0 if number of boarding $\leq$ A

and P_{off} will be: 1) number of alighting - B, if number of alighting > B

2) 0 if number of alighting $\leq$ B

Where A and B are the number of passengers who can get on and get off during the long open+close time. If the number of boarding passengers is smaller than A, and the number of alighting passengers is smaller than B, then passenger will finish their alighting and boarding during the long door open and close time. We will expect small marginal passenger boarding and alighting times, and a large constant term, even though we could estimate A and B for Park Street in our analysis due to the limited data and technical difficult.

4.2 Data Collection and Estimation

The ability to estimate credible dwell time function depends heavily on having a good data set available capturing all the important variables. However, as will be described in this section, obtaining an adequate data set is a significant challenge.

4.2.1 Data Available

1. CTPS data

We have two sets of passenger flow data that were collected by Central Transportation Planning Staff (CTPS). Both sets provide counts of passengers entering and departing from station platforms in 15 minute time intervals. A full data set for the entire Red Line was collected in 1989, even though the data in Park Street and Downtown/South Station were not collected on the same day. The other data set, collected in 1997, only includes the passenger flow data from Alewife to Charles. The format of the CTPS data is shown in Table 4-1:

Table 4- 1 CTPS Passenger Flow Data

	Station 1				Station 2			
	Arriving Volume	Alighting Volume	Boarding Volume	Leaving Volume	Arriving Volume	Alighting Volume	Boarding Volume	Leaving Volume
…	…	…	…	…	…	…	…	…
8:00-8:15	760	105	155	810	810	60	50	800
…	…	…	…	…	…	…	…	…

2. Checker Data

Checker data was collected by the MBTA staff, to get a clearer understanding of the relationship between passenger load and dwell time. This data was collected at Park Street, Downtown and South Station in April 1997 during weekday PM peak period, organized by the index of incoming trains. Table 4-2 shows the format of the checker data. There was one instance of holding control in the data, resulting in a 200 second headway; this observation was deleted from the data set.

Table 4- 2 Checker Data

No.	Branch	Arr. Time	Dept. Time	# of doors	Load in most crowded car	Boarding
1	Braintree	6:57:20	6:58:05	3	100	23
2	Ashmont	7:03:03	7:03:48	4	80	12
…	…	…	…	…	…	

4.2.2 Data Needed

To estimate the dwell time function, we need the following data:

1. number of boarding passengers for train;

2. number of alighting passengers for train;

3. arriving (and departing) load of train;

Neither CTPS data set could provide the above data directly. The data in 1989 is too dated, while the data in 1997 does not include passenger flow information at Park Street, Downtown or South Station. Therefore, we had to transform the CTPS data and checker data to estimate the data needed for the dwell time function estimation, as described below:

1. Update the old CTPS data at Park Street, Downtown and South Station using the new CTPS data

Since we have CTPS data in 1997 from Alewife to Charles, we can update the passenger flow information at Park Street, Downtown and South Station by comparing the CTPS data from 1997 with 1989 for the station with both data set available.

We compared the CTPS data in 1997 and 1989, at the aggregate 1-hour time interval level with the following results:

1) The 1997 PM peak southbound data (i.e. the light flow direction) was not significantly different from the old data. From Harvard to Charles, the line volumes were nearly identical. Even though there were large percentage increases in volume at Alewife, Davis and Porter, the actual increments of passenger number at these stations were not large (See Figures 4-2 ~ 4.5)

Figure 4- 2 Ratio of 1997 CTPS Data to 1989 CTPS Data (southbound line volume)

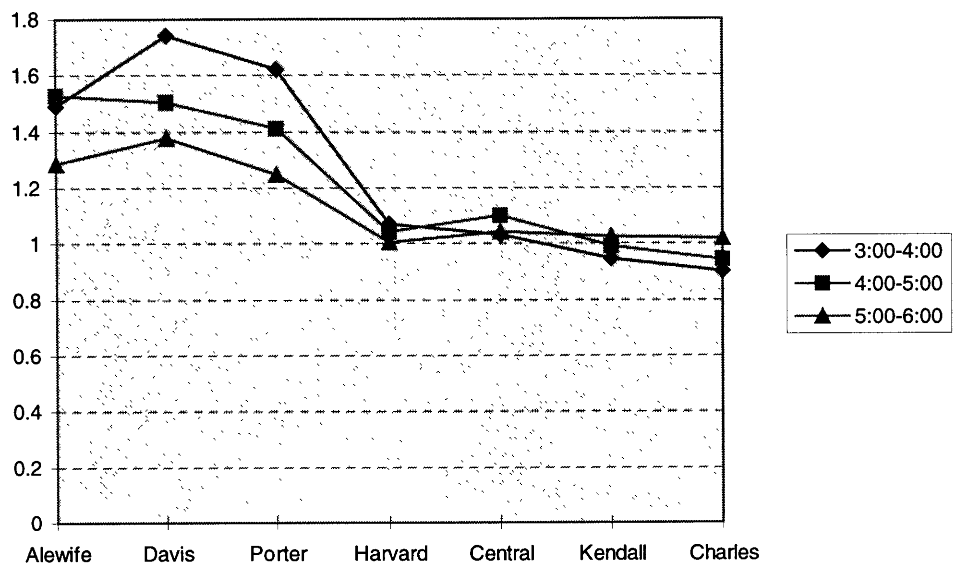


Figure 4- 3 Line Volume From 3:00-4:00 P.M. (southbound)

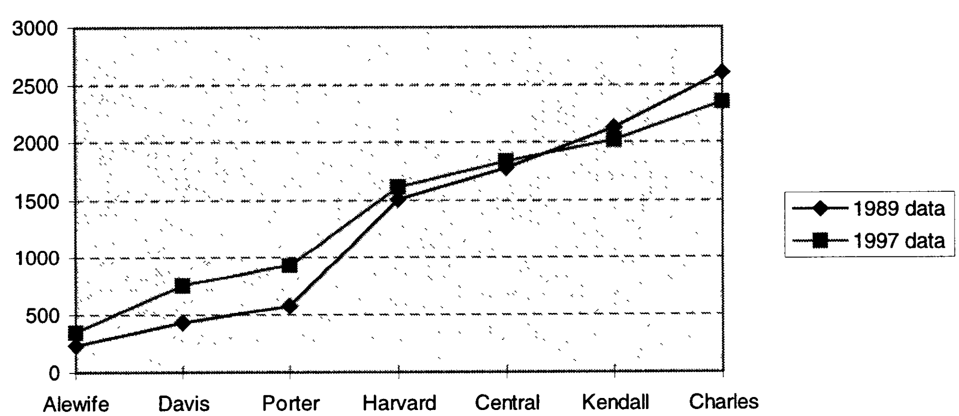


Figure 4- 4 Line Volume From 4:00-5:00 P.M. (southbound)

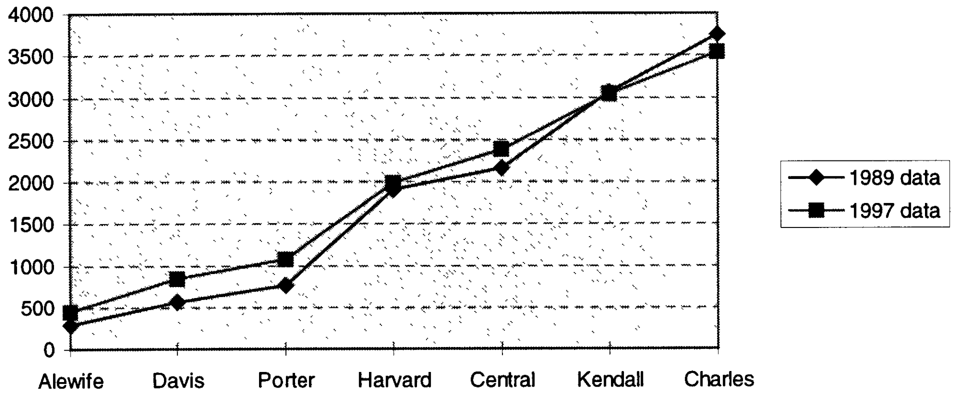
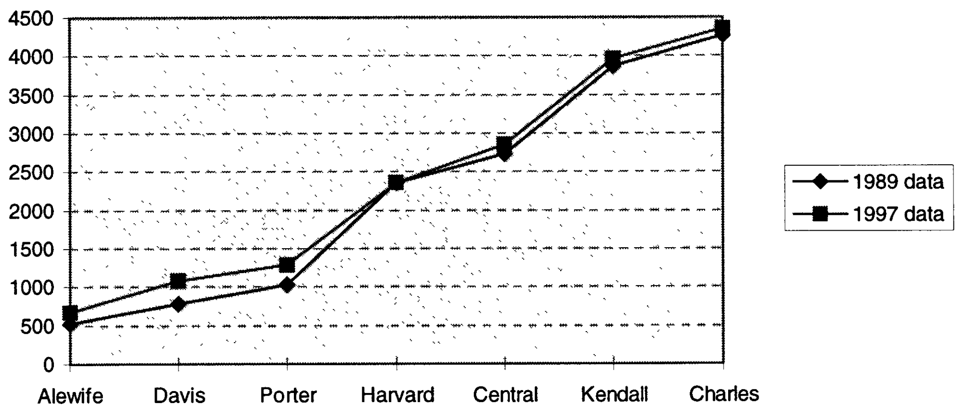


Figure 4- 5 Line Volume From 5:00-6:00 P.M. (southbound)



2) The situation northbound was more complicated. From 2:00-5:00 PM, the new volumes were generally smaller than the old volumes. However, the new volumes from 5:00-6:00 were larger than the old volumes (See Figures 4-6 ~ 4-10) :

- From 2:00-3:00, there was a 500 passenger decrease in line volume at Park Street.
- From 4:00-5:00, there was a 1000 passenger decrease in line volume at Park Street, Charles, and Kendall.

- During 5:00-6:00, there was a 1000 passenger increase in line volume at every station.

- Otherwise, there was little difference in line volume.

Figure 4- 6 Ratio of 1997 CTPS Data to 1989 CTPS Data (northbound line volume)

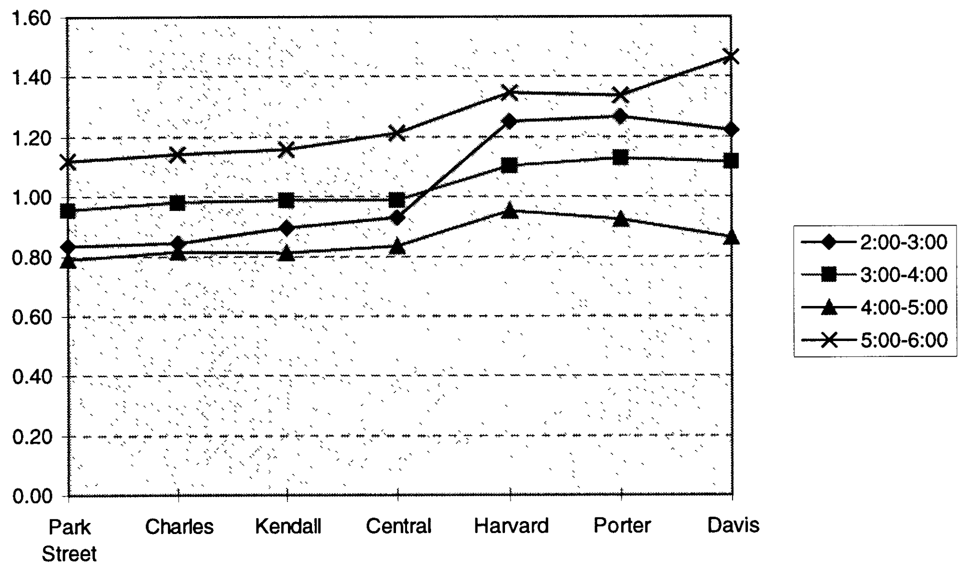


Figure 4- 7 Line Volume From 2:00-3:00 P.M. (northbound)

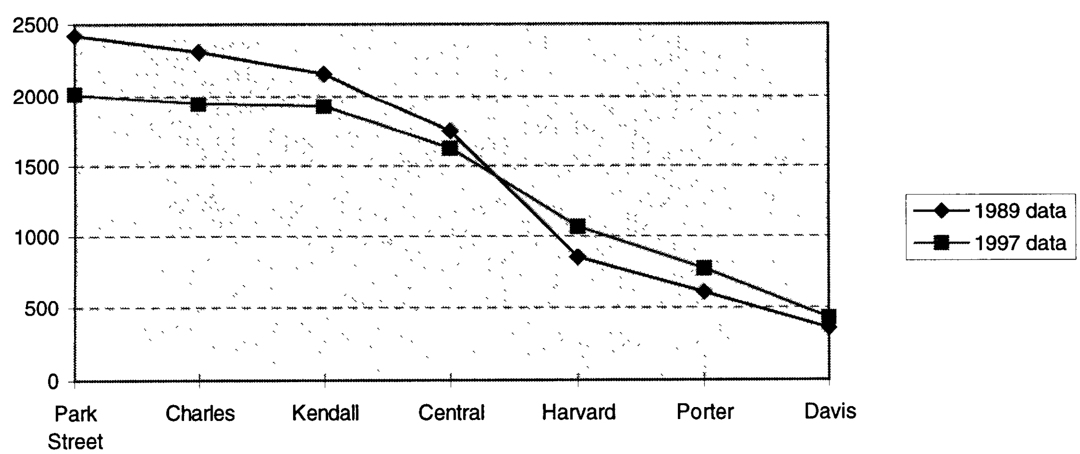


Figure 4- 8 Line Volume From 3:00-4:00 P.M. (northbound)

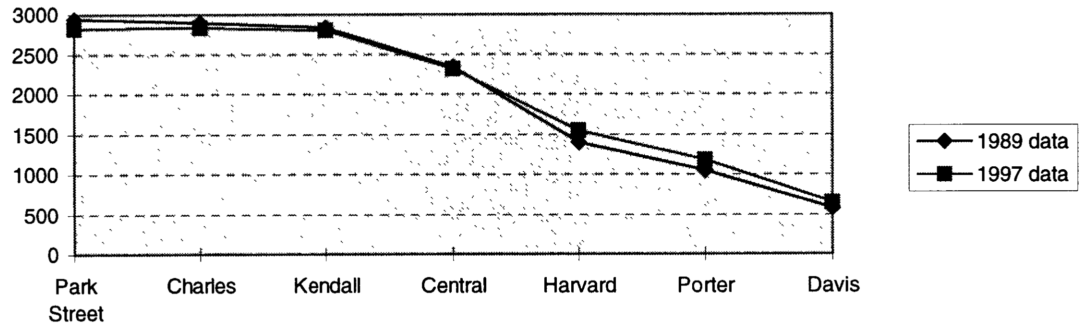


Figure 4- 9 Line Volume From 4:00-5:00 P.M. (northbound)

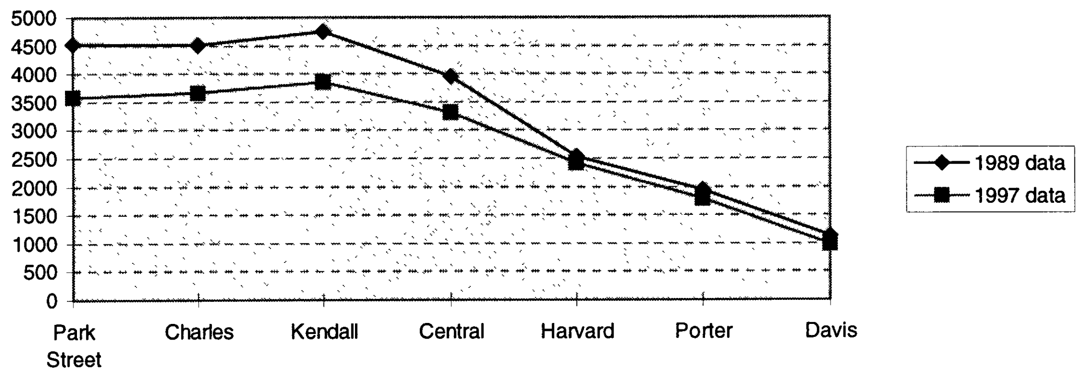
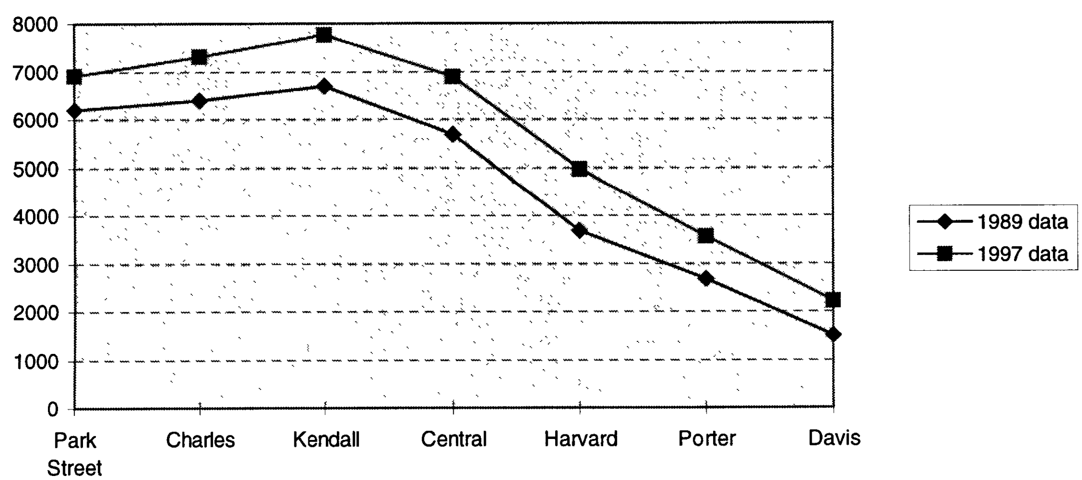


Figure 4- 10 Line Volume From 5:00-6:00 P.M. (northbound)



From the above results, it was very hard to get a consistent trend or conclusion about the relationship between the 1997 and 1989 CTPS data, except for the northbound 5:00-6:00 data. Therefore, it is not possible to estimate the current CTPS data at Park Street, Downtown and South Station based on the current information. We will try another approach to estimate the new CTPS data at Park Street

2. Update the 1989 CTPS data from the checker data

We also got new checker data about the boarding passenger number and dwell time at Park Street, Downtown and South Station, which we tried to use to estimate passenger flows.

First we organized the checker line volume data by 15-minute time intervals, then computed the ratios of new line volume to the corresponding 1989 CTPS data. We took this ratio as the ratio of the change of number of passengers compared with the 1989 CTPS data. We found that new line volume at South Station increased, while the line volume at Park Street decreased. The line volume at Downtown Crossing fluctuated around the old CTPS boarding data (see Figures 4-11 and 4-12.)

Figure 4- 11 Ratio of Boarding Number to 1989 CTPS Boarding Data

Figure 4- 12 1989 CTPS Boarding Southbound

Time Interval	Park Street	Downtown	South Station
3:45 - 4:00 p.m.	625	265	176
4:00 - 4:15 p.m.	750	645	199
4:15 - 4:30 p.m.	769	376	159
4:30 - 4:45 p.m.	1,127	705	358
4:45 - 5:00 p.m.	1,080	788	330
5:00 - 5:15 p.m.	984	1,027	609
5:15 - 5:30 p.m.	1,152	836	511
5:30 - 5:45 p.m.	879	631	252

3. Number of Boarding Passengers

Using the 1989 CTPS data, we first divided the boarding passenger number into time intervals to get the average passenger arrival rate at each station in each time interval. Then we got the headway from the checker data. Multiplying the passenger arrival rate by headway, we estimated the number of boarding passengers for each train. Finally, we used the ratio that we got from the data transformation (see above) to adjust the number of boarding passengers and estimate the current number of boarding passengers.

When we computed the number of boarding passengers in the southbound direction, we separated the arriving passenger number into three parts, which are Ashmont Branch, Braintree Branch and shared trunk parts. Obviously the applicable headways will be different.

4. Alighting Fraction

The alighting fraction is a station-specific variable which we assume has not changed since 1989. Therefore, we can divide the alighting passenger number by the arriving passenger load from the 1989 CTPS data to obtain the alighting fraction at that station.

5. Arriving Load

The checker data only provided the leaving load in the most crowded car of each train. We have to transform this load into the current arriving load, which is the load in the most crowded car when the train arrives. We could use the following relationships:

leaving load - boarding passengers + alighting passengers = arriving load;

arriving load * alighting fraction = alighting passengers

Therefore, arriving load = (leaving load - boarding passengers) / (1 - alighting faction)

Using the parameters computed previously, we estimated the arriving load.

6. Number of Alighting Passengers

The number of alighting passengers is simply a fraction of the arriving load as follows:

alighting load = arriving load * alighting faction.

7. Crowded Car Ratio

During the peak period, the MBTA Red Line operates with 6-car trains. The load of train should be less than 6 times the load of the most crowded car. To get the ratio of the load on the train to the load in the most crowded car, we divided the total leaving load from the revised 1989 CTPS data by the leaving load on the most crowded car from the checker data.

8. Train Leaving Load

We simply multiplied the leaving load in the most crowded car by the crowded car ratio calculated above.

9. Number of Boarding and Alighting Passengers in the Most Crowded Car

The number of boarding and alighting passengers in the most crowded car is obtained by dividing the total numbers of boarding and alighting passengers by the crowded car ratio.

4.3 Model Estimation

After estimating all these data, we then estimated a series of dwell time models. As discussed previously, there are various factors which affect the dwell time, not all of which could be included in the model specification. Therefore, we tried different specifications and combinations of variables as follows:

4.3.1 Model 1

Because the dwell time generally consists of three parts, our first model used the following specification, where α and β are the marginal alighting and boarding time respectively

$$T_{dwell} = T_{open+close} + \alpha \cdot P_{off} + \beta \cdot P_{on}$$

This is the most straightforward model. Because the dwell time generally is controlled by the number of boarding and alighting passengers, we defined the dwell time as a function of P_{on}, the number of passengers boarding the train, and P_{off} the number of passengers alighting from the train. Since we know that Park Street is fundamentally different from other stations, we estimated a separate model from Park Street data only. The regression results are shown in Table 4-3.

Table 4- 3 Model 1 Estimation Results (by station)

	Park Street	t Stat.	Downtown/ South Station	t Stat.
$T_{open+close}$	50.22	6.53	18.94	10.34
α	0.13	0.65	0.25	17.25
β	0.08	0.31	0.10	7.96
R^2	0.02		0.75	
# Observation	74		146	

The results show that dwell time at Park Street has a large and significant constant term with neither variable being significant. This is consistent with our prior expectations. At Downtown Crossing and South Station, all variables in model 1 are statistically significant. The marginal boarding time is much lower than the marginal alighting time.

4.3.2 Model 2

From the previous discussion of dwell time theory, we also expect that the marginal alighting and boarding times may vary with different types of car. Therefore, in Model 2 we separate the data further based on the car type. The results are shown in Table 4-4.

Table 4- 4 Model 2 Estimation Results (by station and car type)

	3-door Cars				4-door Cars			
	Park Street		Downtown & South Station		Park Street		Downtown & South Station	
		t Stat.		t Stat.		t Stat.		t Stat.
$T_{open+close}$	52.6	5.6	21.2	9.6	64.3	4.68	17.3	5.9
α	-0.01	-0.2	0.23	10.4	0.10	1.02	0.26	12.8
β	0.03	0.4	0.09	6.6	-0.12	-1.10	0.11	5.1
R^2	0.004		0.72		0.05		0.77	
# observation	45		71		29		75	

From the regression results, we find that there is some evidence to support our expectation that different types of car have different dwell time functions. The results also reinforce the Model 1 finding that there is no strong linear relationship between alighting and boarding passengers with the dwell time at Park Street.

As discussed above, Park Street is a special station. Passengers have more chance to get on or off during the constant door opening and closing time. Therefore, we attempted to capture this characteristic. However we were not able to find a good statistical model to fit our prior knowledge because of the limited data set. We also attempted to include several expressions reflecting conflicts in passenger movements. After developing several models, we found that we could not avoid strong multi-collinearity among the variables. Since we have to estimate many variables, from limited data it is almost inevitable for us to face the multi-collinearity problem when we attempt to model conflicts.

4.4 Conclusion

From the results of these dwell time models, we found that the first and second model could explain the dwell time very well. The best model is the first one, because of its high R^2 and t-statistic. Even though we expect that there be implications from passenger conflicts on the dwell time, we could not find a model that shows these conflicts to be significant.

There are two likely reasons for not being able to estimate such a model.

1. Poor data. The sample size is not large enough, and there were some serious concerns about the accuracy of the passenger volume estimates, since we have to update the 1989 CTPS data to reflect current conditions.

2. We did not find a satisfactory expression to reflect crowding. We believe there is some influence of crowding, however, it is not easy to capture it.

Even though Model 1 and 2 are not perfect dwell time models, they indeed provide same important insight. Generally, the marginal time for an alighting passenger is larger than the marginal time for a boarding passenger. The main reason might be that there are many passengers alighting and boarding at these three stations, and the platform configuration causes congestion around the doors of cars, thus alighting passenger can not quickly get off the train and thus have a longer marginal time.

We found that the marginal alighting and boarding time for 3-door car train was slight smaller than that for 4-door car trains, which is contrary to our expectations. One possible explanation is that the door in the new Bombardier car is recycled if an object interferes with door closing, hence the marginal time is partly controlled by the passengers. Therefore, the marginal time for the Bombardier car may be larger than that of the Silverbird car. However, this phenomenon may not be as clear at other stations, due to the lower passenger flows. The constant term for the Bombardier car is smaller reflecting faster opening and closing processes on this newer car.

At Park Street, the dwell time is not a simple function of boarding or alighting, but has a high constant term associated with doors on both sides of the train having to be opened and closed independently.

To find a better dwell time model, we may need to redesign the data collection procedure, and resolve the data accuracy concerns. For example, we can follow the same train, and try to eliminate the variance introduced by the doorman behavior. We can also try to obtain accurate alighting and boarding passengers counts, rather than estimating them. At the same time, we should try other expressions for crowding. Also, we should record data on special days, such as July 4, or for other major events when serious crowding can occur. Using these data, we may be able to identify the effect of serious crowding. In the period of observation, serious crowding did not occur on the Red Line, even during the peak period. However, we believe that during incidents, severe crowding can develop, which can have a strong influence on dwell time. So far, we do not have the data to capture these effects.

Chapter 5

Heuristic Real-time Dispatching Control Model

From the analysis in Chapters 2 and 3, we understand that an optimal real-time dispatching control strategy can conceptually be found using mathematical programming tools, such as Cplex. However, due to the magnitude of the problem, the number of variables, the reliability of implementation, and limited computational resources available at most transit agencies, the mathematical programming methods are usually not feasible for real-time control. Moreover, there are other measures of service quality and objectives of real time control, which could be very difficult to model using mathematical methods. Therefore, heuristic methods are often developed instead, which can be more efficient than mathematical programming, especially for very large and complex problems. Essentially, this approach is based on a series of rules derived from the mathematical analysis. Then these rules are used to make dispatching decisions to yield near-optimal solutions to the real-time dispatching control problem.

In the context of this chapter, we will first present the structure and rules of our heuristic model; then implement them on the Braintree Branch of the MBTA Red Line as a case study. Finally we will develop and use a simulation model to evaluate our implementation on the Red Line.

5.1 Dispatching Structure and Strategy

The structure of the heuristic model is presented first. Then based on the discussion in Chapter 3, the rules for the dispatching control strategy are developed.

5.1.1 Definition of the Dispatching Problem

As we discussed previously, the dispatching problem occurs when a train can not follow its schedule, as a result of earlier delays or other disruptions. If we can identify the delay or disruption early, we will have more time to choose the most appropriate control strategies to resolve the problem. Therefore, identifying disruptions which will result in a dispatching problem as early as possible is vital for the heuristic real-time dispatching control system.

There are various ways to identify dispatching problems. For example, we can install radios on the trains and at the stations to report any disruption or disturbance along the line. We can also use an automatic vehicle indication (AVI) system to track the movement of trains. In our heuristic model, track indication, which is available in most rail transit systems, is used to monitor the movement of train along the entire line. While all track circuits could be used, we will use only track circuits at each station. When a train arrives at, or departs from, a monitored station, we can estimate the running time from current station to the terminal for that train and compare the resulting expected terminal arrival time with the schedule. If we find that it is impossible for that train to arrive at the terminal early enough to depart on its next trip on time, the dispatcher will be notified about the problem. If the train has only a small delay, the recovery time is reduced so that the train can depart on schedule. This track indication method will not require extra investment in hardware or monitoring facilities, but it can supply relatively reliable monitoring of the operation of trains to detect serious disruptions that will affect the schedule unless rectified. However, because the length of the one circuit ranges from several hundred feet to more than a thousand feet, the track indication is just an approximate monitoring method.

5.1.2 The Strategy Choice Set for Dispatching Control System (BBDOCS)

Four strategies are often chosen by a transit agency when dispatching problems arise: holding, expressing, deadheading and short turning. To decide whether and how each of those strategies should be included, we must understand their advantages and disadvantages.

1. Holding.

   Holding is the most straightforward control strategy that can be used. In the dispatching problem, we can use the ring off time, which does not rely on advanced communication technology, to implement the holding action. Obviously, holding will be the most appropriate strategy in the choice set, as long as it can satisfy the objectives and constraints. However this may not be the case, especially when there is a long delay.

2. Expressing

   Expressing deals with the dispatching problem having the train with a long preceding headway skip some low volume stations to reduce the gap in the heavy passenger flow direction. The dispatcher must inform the operator about the origin and destination point of the expressing section; and passengers are informed so they can alight, if necessary, at the start of the express segment. The disadvantage of expressing is that the travel time saved by expressing might be small, especially when we consider the extra dwell time needed at the origin station of the expressing segment to allow everyone to decide whether they can use the train being expressed.

3. Deadheading

   Like expressing, deadheading is also used to reduce the gap in the heavy passenger flow direction. The distinct characteristic of deadheading is that an empty train is dispatched from the terminal to an intermediate station at, or before, the heaviest passenger flow. The disadvantage of this strategy is also the limited saving of travel time and hence the minimal impact on a long headway.

4. Short turning

   Short turning is the most complicated control strategy to implement. To short turn a train, the dispatcher has to block the operations in both directions simultaneously, inform the operator and passengers in the train, inform the passengers waiting on the platform of the short-turning station. The most difficult part is to short turn a train in the middle of a big gap, without causing continuing uneven headways in that direction. It is also constrained by the location and configuration of the crossover tracks. Therefore, it is principally employed as a last resort when other control strategies can not supply enough capacity or shorten a very long headway acceptably.

5.1.3 Choose Feasible Choice Subset

Based on the time of the day and expected passenger flow, the track configurations and the nature of the disturbance or delay, we have to select a feasible subset from the entire choice set, presented in section 5.1.2, to form the basis of the dispatching control strategy.

Since short turning is constrained by the location of crossovers, if there is no crossover available for all "controlable" trains, we can consider only holding, expressing or deadheading. The nature of the disturbance also constrains our choice of strategy. For example, if a disabled train, which must be repaired at the terminal, is the cause of the delay, this train cannot be short-turned even if short-turning this train is "optimal" in terms of minimizing passenger waiting time or travel time. In this case, short turning must be excluded from the feasible choice set.

5.1.4 Dispatching Control Strategy

The dispatching control strategy is the heart of the heuristic model. Following the discussion in Chapter 3, we developed the following set of rules for our heuristic real-time dispatching strategies:

1. Estimate the train travel time from the current station to the terminal. The travel time is estimated using a function of the headway, system effect and the characteristics of the train being monitored (see the Su Shen's report in Appendix B for detailed information)

    a) Headway can have either a positive or negative effect on the travel time. At one extreme, very short headways can slow a train in order to maintain a safe separation from the preceding train. At the other extreme with a very long headway, more passengers will be on the train and waiting at stations, which will result in large dwell times and hence travel times. At some intermediate headway the travel time will be minimized.

    b) The "system effect" reflects the current characteristics of the rail system, which may be influenced by the weather, signal and track abnormalities. For example, in snow and ice or other bad weather, the maximum permitted speed is often reduced on exposed sections of the track, and the travel time will be longer than normal.

    c) The train may be malfunctioning and not operating normally, which may result in a longer than normal travel time.

2. If the terminal arrival time estimated above is later than the scheduled terminal arrival time of that train, we will calculate the difference between scheduled departure time and estimated arrival time. When the difference is larger than the minimum recovery time at the terminal, we do not have to use any control strategy, we can simply adjust the recovery time so that the train can depart on schedule. When the difference is smaller than the minimum recovery time,

meaning that this train can not depart on schedule, even with a minimum recovery time, we must choose a control strategy from the choice set. In this case, the train will dwell at the terminal for only the minimum recovery time, and we estimate its earliest possible departure time. To avoid a long gap we will even out the headways by holding a number of trains that can be controlled by the dispatcher. This may be sufficient to resolve the dispatching problem, as discussed in the next step. Obviously, the more trains we can hold, the more possibility that we can deal with a long delay. The number of trains that can be held depends on the communication and control methods that are employed in the transit agencies, the number of monitored stations and the schedule. Generally, we can hold any train that is in the range of monitoring.

Expressing and deadheading are usually not chosen by our dispatching control model. The time saving of expressing or deadheading is the sum of dwell time and deceleration and acceleration time. However, comparing with the long headway at the start of the express segment, the travel time saving is very small, and sometimes can be totally offset by the extra dwell time at the terminal or start of express segment. Moreover, expressing and deadheading can cause confusion and frustration, especially for those passengers who are bypassed by an express train after a long wait.

3. As discussed previously, there will be more passengers in the train and waiting on the platform when the headway is longer. A long headway may cause an over-crowded train and force some passengers to wait for the next train, which is most frustrating. Therefore, we constrain the holding strategy so as not to create a headway so great as to probably result in an over-crowded train. Obviously, this threshold headway, which is determined by the passenger flow, depends on the time of day.

4. If we have a situation in which the holding headway is longer than this threshold headway, we must use short turning to fill the gap caused by the delay.

5. During the peak period, the total numbers of trains and trips are constant. Therefore, we will also shorten the headways of trains that are behind the disturbance to make sure that all trains are used productively during the peak period.

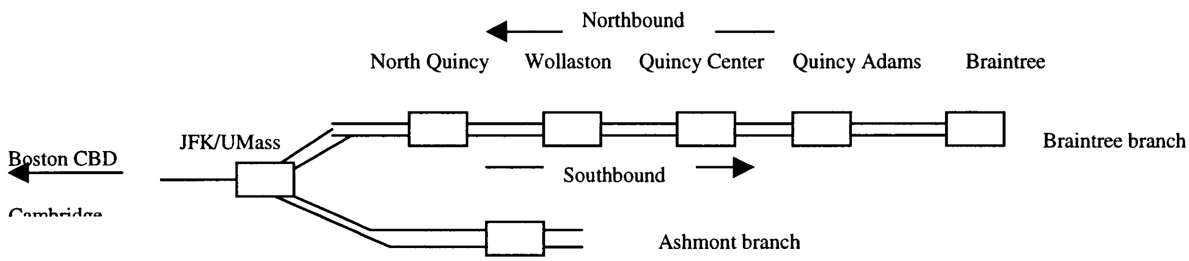
5.2 Braintree Branch Implementation

Using the above heuristic dispatching control strategies, we develop an implementation for the Braintree branch of the MBTA Red Line. In this section, I will first describe the Braintree Branch briefly and the data we need to implement the heuristic dispatching control strategy. Then I will illustrate the control system (BBDOCS) by using a series of examples. Finally we will fully develop the BBDOCS.

5.2.1 Braintree Branch Description

The Braintree branch (Figure 5-1), which is one of two branches on the MBTA Red Line, has the heaviest passenger flow. The operation on it will influence the entire service quality on the Red Line, especially in the AM peak period when the heavy passenger flow direction is northbound into the Boston Central Business District. A disturbance southbound may lead to a large gap at Braintree, making it difficult to depart on schedule from Braintree northbound. Our heuristic dispatching control model is designed to identify these disturbances earlier on the southbound track, and then supply the near-optimal solution to the dispatcher to minimize the impacts on northbound service.

Figure 5- 1 Braintree Branch



We need the following data to design the control strategy.

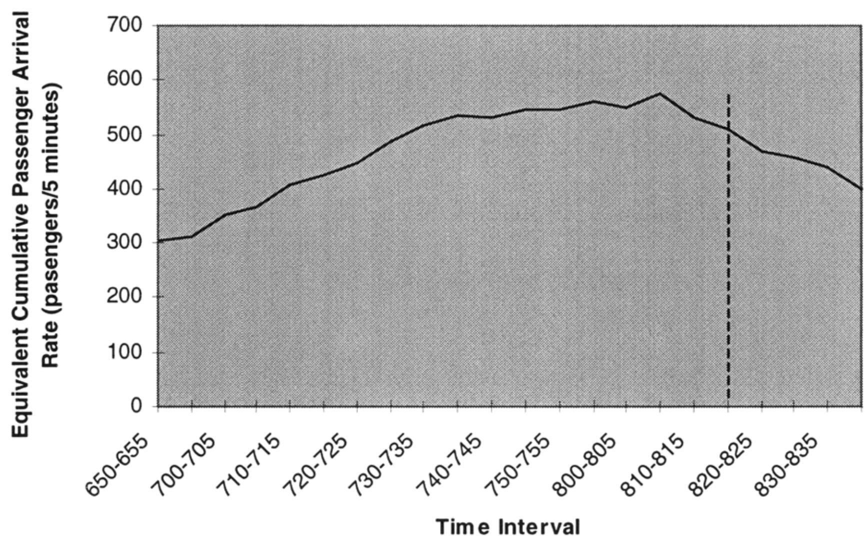
A. Passenger flow: We have to use the passenger flow data to estimate the threshold headway H_{crit} that is the maximum headway without over-crowding. We used data collected by the Central Transportation Planning Staff (CTPS) and also used supplementary passenger flow survey data on the Braintree branch.

The CTPS data (Appendix C) provides composite one-day counts of train-by-train passenger boardings and alightings on all Red Line Braintree Branch trains scheduled to leave or arrive at Braintree from about 6:30 AM to 9:30 PM on a weekday. The data from Braintree to North Quincy were collected on a single day, May 2, 1997. Using these data, we compute the cumulative passenger volume at North Quincy northbound every minute from 7:00 to 9:00 AM, to obtain H_{crit}, the maximum headway which allows normal operation, i.e. without excessive station dwell times or passengers being left at stations.

The results (Appendix C) showed that the 2-hour time period could be divided into 2 parts based on the passenger flow: with the heaviest volume occurring before 8:20 AM. The peak heavy passenger flow on the Braintree branch is the first part, and is close to the "peak of the peak" definition by the MBTA, which is 7:15-8:30.

In the first part, the average cumulative passenger flow rate is 110~120 passenger/minute (see Figure 5-2). This implies that when we have an eight-minute dispatching headway from Braintree, we can get a crowded train at North Quincy northbound with an assumed crowding threshold of 960 passengers. After 8:20 AM, the average cumulative passenger number per minute drops below 90, implying an H_{crit} of 10 minutes or more.

Figure 5- 2 Cumulative Passenger Arrival Rate on the Braintree Branch



B. Track Indication: From a control point of view, the entire track system is composed of circuits. Whenever a train moves into, or exits, a circuit, a signal is sent to the operation control system (OCS) showing this change in circuit occupancy. Combining this track indication information and the train ID, we can find the time of any train at specific locations and thus can estimate other important information, such as travel time and station dwell time.

C. Schedule: The schedule is the base for our dispatching decision. Because the schedule is created based on the passenger flow, operating on schedule generally is close to the optimal solution to the dispatching problem. We will try whenever possible to follow the schedule. The Braintree branch scheduled headway during the AM peak period is 6 minutes.

D. Travel time estimation southbound: Obviously, the travel time on any line segment will depend on the location of the preceding train and other factors, which can be modeled by a travel time function as discussed in section 5.1. In the examples shown later in this section, for simplicity we just use the historical mean of 20 minutes for train travel time from departing JFK to arriving Braintree.

   A similar estimation function is used to forecast the travel time from departing JFK to arriving Quincy Center southbound. We use the average travel time, 11 minutes, in the following tables. Travel time between departing Braintree and arriving Quincy Center northbound is assumed to be the average travel time of 6 minutes.

E. The time to short-turn a train at Quincy Center: We assume 6 minutes, which is the running time from Quincy Center southbound departure to Quincy Center northbound arrival, as the short-turning time.

F. Recovery time at Braintree: We use 8 minutes, which is the average recovery time in OCS data as the scheduled recovery time. From the analysis of OCS data and direct observation, we found that 2 minutes was the minimum recovery time.

G. Dwell time for short-turned train. When a train is short-turned at Quincy Center, it will need more time to make sure passengers are informed. From the OCS data, the dwell time is assumed to be 1.5 minute.

5.2.2 Dispatching Control Strategy

We should monitor the southbound trains at a particular station (for example JFK/UMass), and identify which trains will have problems northbound without some control intervention. As trains arrive at the following stations, these initial decisions will be revised as necessary based on the new estimated running time to Braintree. When trains are to be dispatched from Braintree, we will check the situation again before the ring off, and decide whether to revise the decision.

Using examples, we will define three scenarios reflecting a minor, medium and major gap, and choose the appropriate strategies. Table 5-1 presents the normal situation, which can be considered as our base case.

Table 5- 1 Normal Operation

	Obs. JFK	Obs. Headway	Est. QC(S)	Normal Schedule	Sch. Headway	EPDT (Computed at Bra.)	Actu. Departure Time at Bra.	DDT	Desired Headway	Est. QC(N)
1	6:45		6:56	7:13		7:07	7:13	7:13		7:19
2	6:51	6	7:02	7:19	6	7:13	7:19	7:19	6	7:25
3	6:57	6	7:08	7:25	6	7:19	7:25	7:25	6	7:31
4	7:03	6	7:14	7:31	6	7:25	7:31	7:31	6	7:37
5	7:09	6	7:20	7:37	6	7:31	7:37	7:37	6	7:43
6	7:15	6	7:26	7:43	6	7:37	7:43	7:43	6	7:49
7	7:21	6	7:32	7:49	6	7:43	7:49	7:49	6	7:55
8	7:27	6	7:38	7:55	6	7:49	7:55	7:55	6	8:01
9	7:33	6	7:44	8:01	6	7:55	8:01	8:01	6	8:07
10	7:39	6	7:50	8:07	6	8:01	8:07	8:07	6	8:13
11	7:45	6	7:56	8:13	6	8:07	8:13	8:13	6	8:19
12	7:50	5	8:01	8:18	5	8:12	8:18	8:18	5	8:24

Where:

Obs. JFK: Observed departure time at JFK (the monitor station).

Obs. Headway: Observed headway at JFK.

Est. QC: Estimated arrival time at Quincy Center southbound, where we can short turn trains.

Normal Schedule: Scheduled dispatching time from Braintree.

Normal Headway: Scheduled dispatching headway at Braintree.

EPDT: The earliest possible dispatching time from Braintree is computed at JFK (or any other monitor station).

DDT: Based on the decision made by the dispatcher, the scheduled dispatching time will be revised to the desired dispatching time (DDT) from Braintree.

Desired Headway: Based on the DDT, we can get the desired dispatching headway at Braintree. If the operator executes the decisions very well, the DDT will be the same as the real dispatching time and the desired headway will be as same as the real dispatching headway at Braintree.

Est. QC(N): Estimated arrival time at Quincy Center northbound.

In Table 5-1 operations are normal and no control intervention is required to maintain scheduled service northbound. However if one train has a delay, we have to choose a control strategy to deal with the potential dispatching problem according to the severity of the delay

A. Minor gap

In this scenario, the gap is small, and we can simply follow the schedule by adjusting the recovery time at Braintree. The maximum headway to be classified under this scenario is given by the following inequality:

$$DT_{Bra}^{i-1} + H > DT_{Mon}^{i} + T_{M-B}^{i} + RT \text{ ,}$$

DT_{Bra}^{i-1} : Estimated dispatching time from Braintree for train i-1.

DT_{Mon}^{i} : Departure time for current train i at monitor station.

T_{M-B}^{i} : Estimated running time from monitor station to Braintree for train i.

RT : Minimum recovery time.

H : Scheduled headway.

And $EPDT = DT_{Mon}^{i} + T_{M-B}^{i} + RT$.

For example (Table 5-2), we assume that the 4th train arrives at JFK at 7:06, which is 3 minutes behind schedule. The EPDT for this train is 7:28, which is earlier than 7:31, which is the scheduled dispatching time. Therefore, we can simply adjust the scheduled recovery time for this train from 8 minutes (including 2 minutes minimum recovery time) to 5 minutes, and don't need to change the DDT. Thus, the schedule is not changed, and no control intervention is required.

Table 5- 2 Minor Gap

	Obs. JFK	Obs. Headway	Est. QC(S)	Normal Schedule	Sch. Headway	EPDT (Computed at Bra.)	Real Dispatching Time at JFK	DDT	Desired Headway	Est. QC(N)
1	6:45		6:56	7:13		7:07	7:13	7:13		7:19
2	6:51	6	7:02	7:19	6	7:13	7:19	7:19	6	7:25
3	6:57	6	7:08	7:25	6	7:19	7:25	7:25	6	7:31
4	7:06	9	7:17	7:31	6	7:28	7:31	7:31	6	7:37
5	7:09	3	7:20	7:37	6	7:31	7:37	7:37	6	7:43
6	7:15	6	7:26	7:43	6	7:37	7:43	7:43	6	7:49
7	7:21	6	7:32	7:49	6	7:43	7:49	7:49	6	7:55
8	7:27	6	7:38	7:55	6	7:49	7:55	7:55	6	8:01
9	7:33	6	7:44	8:01	6	7:55	8:01	8:01	6	8:07
10	7:39	6	7:50	8:07	6	8:01	8:07	8:07	6	8:13
11	7:45	6	7:56	8:13	6	8:07	8:13	8:13	6	8:19
12	7:50	5	8:01	8:18	5	8:12	8:18	8:18	5	8:24

B. Medium gap

In this scenario, since the headway is longer, we will use holding as well as adjusting recovery times to even the headways across several trains. The basic idea of this method is to hold a number of trains at Braintree (or stations northbound), to create more even headways between these trains.

If we have a long headway kH (k>1, means this headway is longer than the scheduled headway), and we have N trains between Braintree (departure) and the monitor station (not included the train at the monitor station), we should hold the first train $\frac{(k-1)H}{N+1}$, hold the second train $\frac{2(k-1)H}{N+1}$, ..., hold the Nth train $\frac{N(k-1)H}{N+1}$. In this case, the headways between these trains are $H+\frac{(k-1)H}{N+1}$. If $H+\frac{(k-1)H}{N+1}$ is smaller than the threshold headway H_{crit}, our strategy should not result in any overcrowded trains on the Braintree branch. If $H+\frac{(k-1)H}{N+1}$ is larger than H_{crit} we may consider increasing the number of held trains.

For example (Table 5-3), suppose the current time is 7:12:30, and the 4th train will arrive at JFK in half a minute (so far, we don't know this). At 7:12:30, BBDOCS will check whether to ring-off the 1st train, since the scheduled dispatching time for this train is 7:13. BBDOCS will find that there is already a large gap at JFK even if the 4th train is about to arrive there. Therefore, BBDOCS will hold the 1st train 1 minute.

Half a minute later, the 4th train arrives at JFK, 10 minutes behind schedule. We compute the EPDT for train 4 as 7:35, which is 4 minutes behind schedule. At this time, there are 3 trains between JFK and Braintree departure. According to the formula presented earlier, we find that the achievable headway is 7 minutes which is below H_{crit} of 8 minutes. Therefore, we decide to hold the 1st train by 1 minute, the 2nd train by 2 minutes, and the 3rd by 3 minutes. We change the Desired Dispatching Times (DDT) accordingly.

The DDT then replaces the original schedule for the train dispatch system at Braintree.

To avoid bunching problems in the peak period, we also need to revise the DDT of the following trains. From 7:35 to 8:18, which is the end of our peak window, we have 8 trains to deal with over these 43 minutes. Therefore, the average headway of these trains is 5.4 minutes. We also need to revise the DDT for these trains to make sure that all scheduled trains do, in fact, provide service in the peak period.

Since all headways are smaller than H_{crit}, we do not need to hold any trains at stations northbound. Otherwise, we may want to hold one or two trains at stations northbound, to avoid short turning any train.

All EPDT and DDT values will be updated when each train arrives at each subsequent station, as well as when trains depart from Braintree.

Table 5- 3 Medium Gap

	Obs. JFK	Obs. Headway	Est. QC(S)	Normal Schedule	Sch. Headway	EPDT (Computed at Bra.)	Real Dispatching Time at JFK	DDT	Desired Headway	Est. QC(N)
1	6:45		6:56	7:13		7:07	7:14	7:14 (holding)		7:20
2	6:51	6	7:02	7:19	6	7:13	7:21	7:21 (holding)	7	7:27
3	6:57	6	7:08	7:25	6	7:19	7:28	7:28 (holding)	7	7:34
4	7:13	16	7:24	7:31	6	7:35	7:35	7:35 (holding)	7	7:41
5	*7:16*	*3*	*7:27*	7:37	6	*7:38*	7:40	7:40	5	7:46
6	*7:19*	*3*	*7:30*	7:43	6	*7:41*	7:46	7:46	6	7:52
7	*7:22*	*3*	*7:33*	7:49	6	*7:44*	7:51	7:51	5	7:57
8	7:27	5	7:38	7:55	6	7:49	7:57	7:57	6	8:03
9	7:33	6	7:44	8:01	6	7:55	8:02	8:02	5	8:08
10	7:39	6	7:50	8:07	6	8:01	8:08	8:08	6	8:14
11	7:45	6	7:56	8:13	6	8:07	8:13	8:13	5	8:19
12	7:50	5	8:01	8:18	5	8:12	8:18	8:18	5	8:24

Avoid bunching

C. Major gap

As the delay increases, the holding strategy will eventually lead to crowded trains and some passengers being left at stations. In this situation, we should use a short turning strategy. Due to the technical complexity and loss of capacity, short turning will not be used unless holding results in such long headways that it results in crowded trains.

For example (Table 5-4), suppose the current time is still 7:12:30, and the 4th train will arrive at JFK in 6.5 minute (so far, we don't know this). At 7:12:30, BBDOCS will check whether to ring-off the 1st train, since the scheduled dispatching time for this train is 7:13. BBDOCS find that there is a large gap at JFK even if the 4th train arrives immediately. Therefore, BBDOCS will hold the 1st train for one minute.

One minute later, at 7:13:30, BBDOCS will check the situation again and ring-off the 1st train, which leaves Braintree at 7:14.

At 7:19, the 4th train arrives at JFK, and EPDT of this train is 7:41, which is 10 minutes behind schedule. There are 2 trains between JFK and Braintree departure, so we have to change the DDT of these two trains. Using the formula presented earlier, we find that holding these two trains will result in 9.3 (6+10/3=9.3) minute headways. Since 9.3 is larger than H_{crit}, which is 8 minutes, we know that we can only hold the 2nd train 1 more minute, and hold the 3rd train 2 more minutes, with both headways being H_{crit}. There is still a 3 minutes difference between DDT and EPDT of the 4th train. Therefore, we need to short-turn one train or hold more trains at stations northbound. Using the formula, we find that we need to hold at least 2 more trains northbound to get all headways no higher than H_{crit}.

Let us now consider the short turning alternatives. When we consider short turning a train, we need to look back at the trains on the trunk portion of the Red line, to decide which train is the best short-turn candidate. Usually we will have two candidates, the train at the monitor station and the following train.

In our example, if we short-turn the 4th train, we find the Est. QC(N) is 7:36, which is close to the arrival time of a train departing from Braintree at 7:30. Therefore, we decide to short-turn the 4th train, and hold the 3rd train at Braintree until 7:38. In this case, the northbound sequence will have switched the order of trains 3 and 4.

Table 5- 4 Major Gap (Scenario 1)

	Obs. JFK	Obs. Headway	Est. QC(S)	Normal Schedule	Sch. Headway	EPDT (Computed at Bra.)	Real Dispatching Time at JFK	DDT	Desired Headway	Est. QC(N)
1	**6:45**		**6:56**	**7:13**		**7:07**	**7:14**	**7:14 (holding)**		**7:20**
2	**6:51**	**6**	**7:02**	**7:19**	**6**	**7:13**	**7:22**	**7:22 (holding)**	**8**	**7:28**
3	**6:57**	**6**	**7:08**	**7:25**	**6**	**7:19**	~~7:38~~	**7:38 (holding)**	**8**	**7:44**
4	7:19	22	7:30	**7:31**	**6**	7:41	**Short-turning**	**7:30 (s-t)**	**8**	**7:36**
5										
6										
7										
8										
9										
10										
11										
12										

Sometimes, we can not short-turn the train which is the cause of the major gap, due to some technical problems (such as it being disabled) or other reasons, and we have to adjust our control strategy. In our example, suppose that we are informed that we can not short-turn the 4th train. At 7:26, we observe that the 5th train arrives at JFK (Table 5-5), with a 7 minute headway. If we short-turn the 5th train, we find the Est. QC(N) is 7:43, which is equivalent to the time a train departing from Braintree at 7:38. Therefore, we decide to short-turn the 5th train and hold it by 1 minute at Quincy Center northbound, hold the 3rd train until 7:30, and not short turn the 4th train but hold it at Braintree until 7:46. In this case, the 4th and 5th would be in the reverse sequence northbound.

When we short turn a train, this train has arrived at a headway H_{crit}, because our strategy prefers holding to short turning. If the headway of the short-turning train is the same as the following train northbound, the short-turning train inevitably will have less passengers since it skipped several stations, while the following train will have much heavier passenger volume since it has to accommodate several stations that have not been served for almost 16 minutes. This will make the following train over-crowded, while the short-turning train will still have unused capacity. Therefore, we have to calculate the headway for the short-turning train and its following train, and try to even the load between these two trains.

If we want to short turn a train at Quincy Center, and the short-turning train's headway is assumed to be X minutes, we have the following relationship:

$$X(a_{QuincyCenter} + a_{Wollaston} + a_{NorthQuincy} + a_{JFK}) =$$
$$(2 * H_{crit} - X)(a_{QuincyCenter} + a_{Wollaston} + a_{NorthQuincy} + a_{JFK}) + 2 * H_{crit} * (a_{Brain.tree} + a_{QuincyAdams})$$

Using the OCS data, we find that X = 12.5 minutes, while H_{crit} = 8 minutes. Therefore, we should run the short turning train at a headway of 12.5 minutes, and its following train at a headway of 3.5 minutes, because the passenger flows at Braintree and Quincy Adams are heavy.

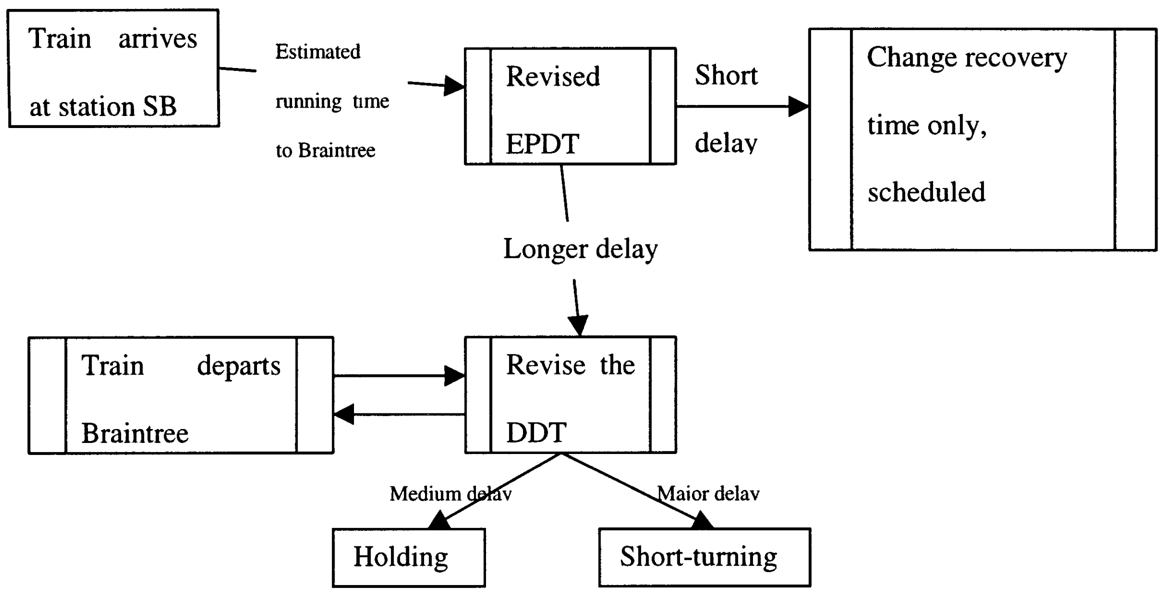
Table 5- 5 Major Gap (Scenario 2)

	Obs. JFK	Obs. Headway	Est. QC(S)	Normal Schedule	Sch. Headway	EPDT (Computed at Bra.)	Real Dispatching Time at JFK	DDT	Desired Headway	Est. QC(N)
1	6:45		6:56	7:13		7:07	7:14	7:14 (holding)		7:20
2	6:51	6	7:02	7:19	6	7:13	7:22	7:22 (holding)	8	7:28
3	6:57	6	7:08	7:25	6	7:19	7:30	7:30 (holding)	8	7:36
4	7:19	22	7:30	7:31	6	7:41	7:46	7:46 (holding)	8	7:52
5	7:26	7	7:37	7:37	6	7:48	Short-turning	7:38 (s-t)	8	7:43
6	*7:29*	*3*	*7:40*	7:43	6	*7:51*	7:51	7:51	5	7:57
7	*7:34*	*5*	*7:45*	7:49	6	*7:56* Avoid bunching	7:56	7:56	5	8:02
8	*7:38*	*4*	*7:49*	7:55	6	*8:00*	8:01	8:01	5	8:07
9	*7:42*	*4*	*7:53*	8:01	6	*8:04*	8:06	8:06	5	8:12
10	*7:47*	*5*	*7:58*	8:07	6	*8:09*	8:10	8:10	4	8:16
11	*7:49*	*2*	*8:00*	8:13	6	*8:11*	8:14	8:14	4	8:20
12	*7:52*	*3*	*8:03*	8:18	5	*8:14*	8:18	8:18	4	8:24

5.2.3 Detailed Design

In BBDOCS, we have following flow chart (Figure 5-3).

Figure 5- 3 Dispatching Control System Flow Chart



Based on the control strategies discussed above, we use C++ to implement our heuristic dispatching control model on the Unix platform. The data structure and pseudo-code are presented below.

A. Data structure

We have two data structures in the model for stations and trains.

A1. Station

$\tilde{TT}(k)$: the smoothed travel time from the preceding station to current station.

TT(k~QA): Estimated time function from current station k to Quincy Adams. (At every station, there is a function, which has several parameters, to estimate the travel time from station k to Quincy Adams.)

TT(k-QC): Estimated time function from current station k to Quincy Center southbound.

Minimum Recovery Time: If the station is a terminal or station where short turning can be employed, this number will be larger than 0. This is set to 2 minutes at Braintree based on the OCS data.

A2. Train

Current station (we get this information from the current circuit occupancy)

Next event type (this specifies the current state of the train)

EPDT (earliest possible departure time. See B1)

DDT (desired dispatching time at Braintree. It should be initialized as the scheduled time, and revised according to the subsequent events and control actions. See B2)

EQCS (estimated Quincy Center SB time. See B4)

EQCN (estimated Quincy Center NB time. See B5)

Dheadway (desired headway. See B3)

$AT(k)$: arrival time at station k

$DT(k)$: departure time at station k

$DW(k)$: dwell time at station k. $DW(k)= DT(k) - AT(k)$

$H(k)$: arriving headway of at station k. $H(k)=AT(k)-train[I-1].AT(k)$

EDT(QA): Estimated departure time at Quincy Adams southbound

$TT(k)$: travel time from preceding station k-1 to station k, when k is not Quincy Adams.

$TT(k)=AT(k)-AT(k-1)$

travel time from QC to QA is defined as

$TT(QA)=DT(QA)-AT(QC)$

Pullout Indicator: the indicator for the pullout train, which is a train pulling out from Cabot into the service based on the schedule.

B. Formulas

All formulas that will be used to decide the control strategy are included in this section.

B1 EPDT is one of the most important variables in the model. We use a function developed by Su Shen, which is defined in the Appendix B, which also includes the methods to calculate the EQCS and EQCN.

B2 train[i].DDT = train[i-1].DDT + train[i].dheadway. In other words, the current train's DDT is the preceding train's DDT plus the new desired headway, which is computed in B3.

B3 When a train departs from any station southbound, dheadway = (train.EPDT - last departure time from Braintree) / (the number of trains from current station to Braintree, including the current train and the pull-out service train)

When a train is to depart from Braintree, dheadway = (EPDT of any train at JFK/UMass - last departure time from Braintree) / (the number of southbound trains on Braintree Branch, including that train at JFK/UMass).

B4 Estimated Arrival Time at Quincy Center southbound = current time + estimation time (current station k → Quincy Center southbound) TT(k-QC). (Obviously, the current stations only include JFK/UMass, North Quincy and Wollaston southbound.)

B5 Estimated Arrival Time at Quincy Center northbound = current time + estimated time (current station k → Quincy Center northbound) TT(k-QC). (Obviously, the current stations includes Braintree and Quincy Adams northbound.)

B6 We have two options for headways: we can use either the arriving headway or departing headway. In our analysis and design, we use the departing headway.

C. Constant Values

We have various constants in the model, which are very important in terms of making correct decisions. Some of them such as H_{crit}, are based on the current data, might need to be revised in the future as riderships levels and patterns change.

C1 H_{crit} = 8 minutes

C2 Minimum recovery time = 2 minutes

C3 Estimated short-turning time from Quincy Center southbound to Quincy Center northbound = 6 minutes.

C4 Small headway thresholds (in estimated time function, see Appendix 5.1)

For all in-service trains at all stations, the small headway threshold is 180 seconds.

For pull out train, the threshold varies by station as shown in Table 5-6:

Table 5- 6 Small Headway Thresholds (pull-out train)

Type \ Station	JFK	NQ	WOL	QC
To Braintree	270	245	220	200
To Quincy Center	240	215	200	

C5 Historical mean travel time between consecutive stations (JFK-Quincy Adams) southbound

Table 5- 7 Travel Time

Station	JFK-NQ	NQ-Wol	Wol-QC	QC-QA
Value	361	122	162	200

C6 Proportions of running time

Table 5- 8 Proportion of Running Time

	P(JFK~NQ/JFK~QA)	P(NQ~Wol/NQ~QA)	P(Wol~QC/Wol~QA)
Regular Train	0.427	0.252	0.448
Pull-out Train	0.444	0.213	0.441

D. Initialization

When the OCS receives a track indication with a train ID, it considers this as a train movement, and adds this train to the control list. At the beginning of the peak period, the heuristic sets the travel time between consecutive stations to the historic means. We also set the headway of the first train at every station to the scheduled headway, 6 minutes.

E. Logic of Model

We can classify the events in the model into two categories, which are defined below:

1) OCS receives a track occupancy indication with a train ID at a platform track.

2) OCS receives a track vacancy indication with a train ID at a platform track, or half minute before the ring off time at Braintree

The reason that we split the train operation events into two categories is that we want to generalize the movement of trains, and also find the source of disturbances. Using the times of events 1 and events 2, we can also estimate the train travel times and dwell times. When an event 2 happens, BBDOCS will check the situation based on the following flow-chart (see Figure 5-4) and make a decision accordingly. Each step in the flow chart will be explained later in this section.

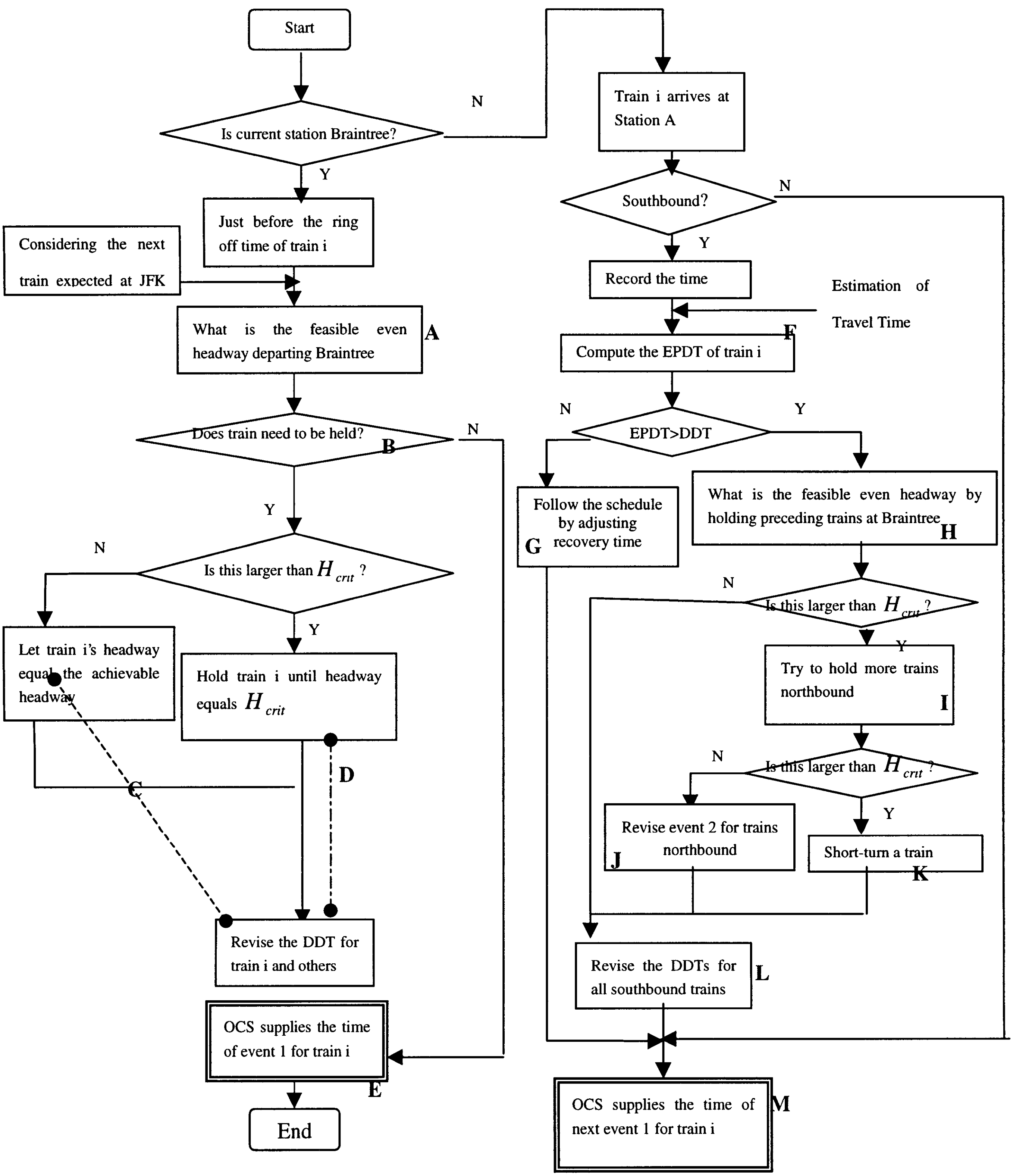
Generally there are several things to be done in this step:

1. Check the situation and find whether there is a disturbance, and whether holding or short turning should be employed to handle it.

2. Specific suggestions are made by BBDOCS, such as the ring off time, or which train should be crossed.

3. The location and during of holding actions.

Figure 5- 4 Event 2 Flow Chart



Step A

The feasible even headway departing Braintree is:

(Train[expected at JFK].EPDT - Station[Braintree].last departure time) / (the number of trains between JFK SB and Braintree)

If there is not a train at JFK SB, we assume that one will arrive at JFK instantly. Therefore, we will increase the number of trains between JFK SB and Braintree by one.

Moreover, since the number of trains between JFK SB and Braintree is very important in BBDOCS and will be used many times, we can alias it as StrainCANBEhold.

Step B

If the headway we just computed is larger than train i's dheadway (Train[i].next time - Station[Braintree].last departure time), then the result is "Yes", otherwise is "No".

Step C

C1 Revise train i's dheadway to the headway obtained in Step A, and set Train[i].DDT = Station[Braintree].last departure time + Train[i].dheadway;

C2 Train[i+1].dheadway = Train[i].dheadway, Train[i+1].DDT = Train[i].DDT + Train[i].dheadway;

C3 Repeat Step C2 until train j which is expected at JFK

C4 Revise the next event time of any train at Braintree (usually trains i and i+1) to their new DDT times.

Step D

D1 Revise train i's dheadway to H_{crit} (8 minutes), Train[i].DDT = Station[Braintree].last departure time + 8 minutes;

D2 Train[i+1].dheadway = 8 minutes, Train[i+1].DDT = Train[i].DDT + 8 minutes;

D3 Repeat Step D2 until train j which is expected at JFK

D4 Revise the next event time of any train at Braintree (usually trains i and i+1) to their new DDT time.

Step E

E1 Change the next event type to 1.

E2 Revise the dheadway and DDT, smooth the following trains headways if necessary.

E3 Revise the last depart time at Braintree.

Step F

Based on the estimation function, we compute EPDT = Train[i].next time + estimation(current station, preceding headway, history average, system condition)

Step G

In this step, we do not change anything, but remind the dispatcher that this train may not have the usual recovery time. If we find that the estimated recovery time is close to the minimum recovery time, we should take some actions, such as announcements at Braintree, to reduce the recovery time as much as possible.

Step H

H1 Compute the number of southbound trains that are ahead of train i.

H2 The feasible even headway = (Train[i].EPDT - Station[Braintree].last departure time) / number from step H1.

Step I

This Step is similar to Step H.

I1 Compute the number of trains that are ahead of train i on the Braintree branch, not only southbound, but also northbound.

I2 The feasible even headway = (Train[i].EPDT - Train[the preceding train of current first train northbound].DDT) / number from step I1.

Part J

J1 Hold first train northbound by (new headway from Step I - Train[current first train northbound].dheadway).

J2 Hold second train northbound by (new headway from Step I - Train[current first train northbound].dheadway) + (new headway from Step I - Train[current second train northbound].dheadway).

J3 Repeat Step J2 Until train at Braintree

Step K

K1 We need to compute train i's EQCN and compare it with other train's EQCN.

When the train is beyond Quincy Center southbound, we will use the estimation function to estimate EQCN; otherwise EQCN equals the train's DDT + estimation running time between Braintree and Quincy Center northbound. We can choose the most appropriate EQCN to match the current train's EQCN, then short turn the current train into the order northbound of that train who has most appropriate EQCN.

K2 If train i's EQCN is less than train j's EQCN, it means train i can replace train j's position northbound.

K3 Set Train[j].DDT = Train[j+1].DDT, and keep revising until train i.

K4 In event 1, BBDOCS will check whether a command will be given to short turn a train at Quincy Center southbound.

K5 If train's current station is Braintree, Train[].next time = train[].DDT

Step L

L1 If we only hold trains at Braintree:

Train[current first train SB].DDT = Station[Braintree].last departure time + headway from Step H.

Train[current first train SB].dheadway = headway from Step H.

Train[current second train SB].DDT = Train[current first train SB].DDT + headway from Step H.

Train[current second train SB].dheadway = headway that we computed in Part H.

Keep revising until train i.

L2 If we hold southbound and northbound trains:

Train[i-1].DDT = Train[i].EPDT - headway from Step I, Train[i-1].dheadway = headway from Step I.

Keep revising it until finishing all trains southbound.

L3 If train's current station is Braintree, Train[].next time = train[].DDT

Step M

This Step is similar to Step E.

M1 Change the next event type to 1.

M2 Revise the last depart time at the Station.

We also check the system and make decisions when event 1 happens, however, the flow chart of event 1 will be similar to that of event 2. Therefore, we will not present Event 1's flow chart.

5.3 Simulation of Control Strategies

5.3.1 General Approach

Since we want to ensure that our heuristic control strategy will work effectively before implementation, we developed a simulation model to test it and to obtain some insight into the performance characteristics of interest.

In general, there are two basic approaches to simulation, one is fixed-time simulation, and the other is event simulation. Fixed-time simulation has a very direct and clear logic based on a selected time interval for updating. At the end of each interval, we update the system states, execute any events that have happened during the last time interval and continue. However, when the average interval between two events is large relative to the fixed interval, this method wastes resources on unnecessary computing.

Event simulation, which I use in the TD (terminal dispatching) simulation model, is based first on defining the events of interest that happen in our model, then computing the next time of each event. Next, we find the next chronological event, and this event is the next one executed. We move the clock ahead to that time and continue. In this method, each event is instantaneous and changes the state of the system, and also triggers other events. Using this approach, we can decrease the computation and increase the efficiency. However, this method is also somewhat more difficult and time-consuming to design and program on the computer.

5.3.2 Purpose of TD Simulation

The TD model is designed to simulate the operation of trains on the Braintree Branch of the Red Line in the AM period, including control strategies, such as holding and short turning. However, this model is also designed to simulate the general terminal dispatching problem on any public transport line.

In the TD simulation (the Braintree branch case in particular), we assume a series of trains entering the starting point of the simulation system, JFK southbound. The entire system state and control decisions are updated after each train leaves every station until reaching JFK northbound, which is the end of the simulated system. Once the train departs from JFK northbound, we do not continue to monitor it further.

To quantify the BBDOCS model, there are 6 stations, and 11 platforms on the Braintree branch. We need to monitor 13 trains' operations from 6:45 to 8:45 AM. We calibrate our simulation model based on actual data (from JFK/UMass southbound to Braintree) collected by the OCS system. Our simulation model will choose the appropriate strategies and try to smooth the operation following the control strategy described in section 5.2. We also run the model in evaluation mode using the actual data on the entire Braintree Branch, not just the southbound part. Since minimizing passenger waiting time is one of our primary objectives, we will compute the total passenger waiting time in both models, and compare them. The number of overcrowded trains is another important measure of service quality, which we also want to minimize. We will also compute this measure as well as the passenger waiting time.

Clearly we could not use the actual northbound data in the simulated mode, since we may use different dispatching times at Braintree.

5.3.3 Events and Logic

The definition and logic are both very important in the design of an event simulation: the definition of the events should be clear and concise, while the logic between the events should be rational and consistent.

Definition of Events

In our model, there are two types of events, which are defined below.

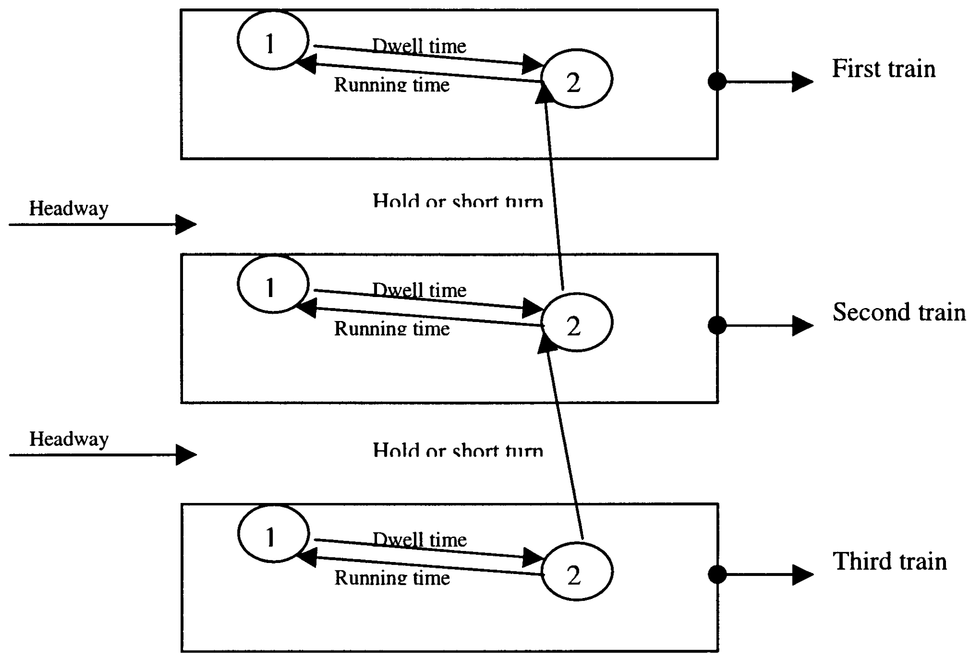
Event 1: Train arrives at a station. This event will trigger event 2 for this train.

Event 2: Train departs from a station. This event will trigger this train's event 1 at the next station

Logic of Simulation

When event 1 occurs, that is a train arrives at a station, we estimate the dwell time and calculate the time of event 2 for this train. When event 2 occurs, we will estimate the running time to the next station and thus get the time of next event 1 for this train. Moreover, we will decide whether we should employ any control strategy, based on the flow chart (see Figure 5-5).

Figure 5- 5 Logic of Simulation



5.3.4 Data

We used data from 11 weekday AM peak periods for the evaluation and simulation. Since the only information we really need to know is the station arrival and departure times, we can focus on the track circuits at every station, specifically the circuits immediately before the platform, the platform itself and that immediately after the platform. We also need to collect some data on the circuit close to Braintree to help us understand train operation through the terminal itself.

Since we did not have train ID information in the OCS data, we had to process this data very carefully, especially at Braintree. The arrival order and departure order of trains may change at

Braintree, therefore we had to link the inbound and outbound trains manually. Also, trains may enter service or leave service, which needed to be marked in our database.

The data processing is easier at other stations than at Braintree. Generally, we can just use the time when the track before the station platform became vacant as the arrival time, and the time when the track circuit after the station platform was occupied as the departure time. Obviously, there will be exceptions, such as short turning, track indication errors, etc, which need to be recognized and dealt with appropriately.

5.3.5 Evaluation & Simulation

Using the actual OCS data, we can estimate the total passenger waiting time on the Braintree branch. We also use the OCS southbound data as the basis for the simulation. Our dispatching control strategy will choose the appropriate departure time so as to minimize the total passenger waiting time. We can implement holding and short turning in our model. For each day, we ran two evaluations and two simulations, and compared the results.

Evaluation 1

In this experiment, we used actual train arrival and departure times on the entire Braintree Branch. The results represent what happened that day and provide a base case for comparison purposes.

Evaluation 2

In this experiment, we used actual train arrival and departure times for the Braintree Branch southbound and for Braintree itself; the northbound arrival and departure times are estimated using the travel time estimation function in the simulation model. Since we estimate the running time northbound in our simulation experiment, this will be a fairer basis for comparison since it uses the same estimation function in our evaluation. Thus, it will avoid any bias there might be in running time between the simulation and evaluation modes.

Simulation 1

In this experiment, we use the exact OCS data southbound. However, our dispatching control system can change the DDT and thus the northbound running times and dwell times must be estimated. We will use the mean values of travel times from each station southbound to Braintree to compute the EPDT (earliest possible departure time), one of the most important variables in our control model. Obviously, this is not a very accurate estimate, due to a variation in actual travel times. So, it will give us a lower bound on total passenger waiting time saving.

Simulation 2

As with simulation1, this model uses the exact OCS data southbound and estimates the northbound running times and dwell times. However, the EPDT will be based on the actual data, rather than the estimated values. This makes the unrealistic assumption that we know the EPDT exactly for every train, but it will give us an upper bound on total passenger waiting time savings if we have very accurate train travel time forecasts.

5.3.6 Results

The results are shown in Table 5-9, which shows the actual and simulated Braintree departure times for each train on each day. The first two columns show the actual arrival time and departure time for each train at Braintree; the third column (RO) shows the ring off time and the final two columns show the simulated departure times under the two simulated travel time scenarios. Boldface entries in the two simulation columns mean that our dispatching control model suggests a revised departure time.

Feb. 9 AAT	ADT	RO	SIM1	SIM2
6.57.18	7.06 29	7 06.02	7.06 00	7:06:00
7:01:05	7:13.41	7.13.01	7.12:00	7 12 00
7.10.36	7.20.14	7:19.03	7:18.00	7:18.00
7.15:57	7.25.29	7:25 02	7:24.00	7.24:00
7.22 33	7:31:06	7:30:04	7:30:00	7:30:00
7:27:53	7:37:32	7:37.03	7:36:00	7:36:00
7.33:21	7.43:36	7.43.02	7.42:00	7:42:00
7:40:52	7.49.33	7:48.08	7.48.00	7:48.00
7:46.01	7.54:05	7.51:49	7:54 00	7.54:00
7.51.45	8.02 21	8:01:01	8.00:00	8.00:00
7.56.22	8.07.37	8.07.02	8.06.00	8.06.00
8:04:42	8 12.39	8.11:41	8.12.00	8:12 00
8:09:53	8.17.37	8.17 07	8.18.00	8:18:00

Feb. 10 AAT	ADT	RO	SIM1	SIM2
7:02.20	7.06.23	7 06.01	7.06:00	7 06:00
7:03 43	7.13 37	7:13.03	7.12 00	7.12:00
7:10 29	7:19 43	7 19.03	7:18:00	7:18:00
7:16.44	7.26 26	7.25.02	7:24.00	7:24:00
7:23:26	7:31:33	7.31.02	7:30.00	7 30:00
7:28.48	7:37.43	7:37.02	7:36.00	7 36.00
7:33.45	7.44:16	7.43:02	7:42.00	7 42:00
7:43.12	7:49:30	7:49:03	7:48 00	7:48:00
7:46.35	7.55:25	7:55.02	7:54:00	7.54:00
7:51:46	8.01 42	8.01.02	8 00.00	8 00:00
7:57.34	8.07.30	8.07 02	8.06:00	8.06:00
8:03.56	8 13.29	8:13.01	8 12:00	8 12.00
8:10.02	8:19:03	8.18:02	8:18 00	8:18.00

Feb. 11 AAT	ADT	RO	SIM1	SIM2
6 55.38	7 08:32	7 06:53	7.06:00	7.06:00
7.07.24	7 13 30	7:13:01	**7:12:06**	**7:12:06**
7.10.59	7.19.42	7.19.01	**7:18:12**	**7:18:12**
7:17.02	7 25.30	7:25:02	**7:24:18**	**7:24:18**
7.22:01	7 31.31	7.31:02	**7:30:24**	**7:30:24**
7.27.45	7:38:13	7:37.03	**7:36:30**	**7:36:30**
7.33 57	7 43.40	7:43:02	**7:42:36**	**7:42:36**
7:40.17	7.49.56	7.49.03	**7:48:42**	**7:48:42**
7.46 30	7.55:39	7:55:02	**7:54:48**	**7:54:48**
7:52:40	8 00 03	7:59:00	**8:00:54**	**8:00:54**
7:57:54	8 07.33	8.07.02	8 07.00	8.07:00

Feb. 12 AAT	ADT	RO	SIM1	SIM2
6.58.44	7.06.35	7 06 02	7.06.00	7 06:00
7.06.48	7.13.41	7 13.02	7.12 00	7.12.00
7:13:14	7.19.38	7.19:03	7 18 00	7.18 00
7:18.26	7:25:36	7 25:02	**7:25:00**	7:24:00
7:23:56	7.31:42	7.31 02	**7:32:00**	**7:30:13**
7:28.00	7:37:46	7:37.02	**7:39:00**	**7:36:26**
7:33.47	7:44:05	7.43:02	**7:46:00**	**7:42:44**
7.46:53	7:50.14	7:49.02	**7:51:20**	**7:48:55**
7:47:26	7.55:27	7:55.02	**7:56:40**	**7:54:42**
7:52:38	8.01.53	8:01.03	**8:02:00**	**8:00:32**
7.57.38	8.07:33	8.07:02	**8:07:20**	**8:06:21**
8:04.08	8.14:20	8.13 02	**8:12:40**	**8:12:11**
8:10.01	8.19.58	8.18.02	8.18 00	8.18.00

Feb. 13 AAT	ADT	RO	SIM1	SIM2
6.57.38	7 06.25	7 06.03	7 06 00	7.06.00
7 02.29	7.13.30	7:13.02	7:12 00	7.12.00
7:09.01	7.19.39	7:19 03	7:18 00	7:18:00
7:21:49	7:25:34	7:25:16	7:24.00	7:24.00
7:19.25	7:31:30	7:31:03	7.30.00	7:30.00
7:27.44	7 37.58	7:37.02	7 36:00	7:36:00
7.33.44	7.43:50	7:43:02	7:42.00	7:42:00
7:41.27	7:49.27	7:49:03	7.48:00	7.48.00
7.46:16	7 55.41	7:55.03	7:54.00	7:54.00
7.51:34	8.01.32	8.01.02	8.00.00	8.00.00
7.57 58	8 07:38	8 07:05	8.06:00	8.06:00
8:03:49	8:13.46	8:13.02	8:12.00	8.12:00
8.09 43	8.19.35	8:18.02	8:18.00	8 18.00

Feb. 17 AAT	ADT	RO	SIM1	SIM2
6.56 47	7 07 26	7 06 02	7.06:00	7.06 00
7:06 48	7.13 38	7 13.01	7.12.00	7.12.00
7.10.32	7 19.51	7.19:01	7 18.00	7:18:00
7:17:52	7.25:42	7:25 03	7.24:00	7:24:00
7:22:15	7.31:11	7.29.08	7.30:00	7:30.00
7.27.18	7:37 52	7:37:03	7:36.00	7:36:00
7:32:58	7.43.55	7:43:02	7.42.00	7.42:00
7:41.32	7:49 52	7:49 02	7.48.00	7.48:00
7:46:00	7:56 17	7:55.02	7 54.00	7.54 00
7.54.00	8 01.59	8.01:02	8 00 00	8 00:00
7:58.02	8 07.42	8.07 01	8:06 00	8 06.00
8:04.11	8.14:18	8.13.02	8 12.00	8.12:00
8:09:27	8.19.06	8.18:01	8 18.00	8.18.00

Feb. 18 AAT	ADT	RO	SIM1	SIM2
6.56 36	7 06.30	7.06:03	7.06.00	7.06.00
7 03.20	7 15.21	7.13:02	7:12:00	7.12:00
7.14.06	7 18 01	7.17.23	7.18.00	7.18:00
7.20.15	7.24:30	7:20:51	7:24.00	7.24.00
7:26 21	7.31.01	7:29.35	7:30.00	7:30:00
7:27.55	7:37 43	7:37:02	7.36.00	7 36 00
7.33 34	7:43 17	7.42.03	7:42.00	7.42:00
7 41.11	7 50.22	7:49:02	7:48:00	7 48:00
7 47:41	7.56.58	7:56:32	7:54.00	7:54:00
7.52 25	8.02.38	8:00:30	8 00:00	8.00:00
7.59.24	8:05.53	8.03.34	8:06:00	8 06:00
8.04.36	8.10.34	8.10.05	8.12.00	8.12.00
8 09.28	8.16.21	8.15:26	8:18:00	8.18.00

Feb. 19 AAT	ADT	RO	SIM1	SIM2
6:59 34	7.06:39	7.06 02	7:06 00	7.06 00
7:04.24	7.13.42	7.13 04	7.12.00	7.12 00
7.09:46	7.18.29	7.18:31	7.18.00	7 18 00
7:17:30	7.28 03	7.25.01	7 24.00	7 24.00
7:24:41	7.31.41	7.31:02	7:30.00	7 30 00
7.29.57	7:37.56	7.37.03	7.36 00	7.36:00
7:33 23	7.43:44	7 41 28	7:42 00	7 42 00
7.40.47	7:49:12	7 45.50	7:48.00	7:48 00
7:46:02	7.54.10	7 50.54	7:54.00	7 54:00
7.59.10	8 02 01	7.59:26	8:00:00	8 00.00
7 51.04	8:05.07	8 04.05	8:06.00	8 06 00
8:03:54	8:13:42	8.13:01	8.12 00	8.12.00
8:10:11	8 19 49	8:18.03	8:18:00	8.18.00

Feb. 23 AAT	ADT	RO	SIM1	SIM2
6.46:20	7:03.26	7:01.05	7:06:00	7:06 00
6.59:10	7:12.38	7:12.00	7 12.00	7.12 00
7.09:09	7.19.57	7.12.40	7:18:00	7:18:00
7:19 50	7.25:33	7:24.48	7.24:00	7:24:00
7 24.24	7:32:00	7.30:42	7.30:00	7 30:00
7.27.25	7:34.44	7:32.28	7.36:00	7.36.00
7.33:38	7:39:31	7 35:22	7.42 00	7:42.00
7:36:23	7.42.39	7.42 07	7.48.00	7:48 00
7:41.10	7:47.50	7 46:37	7:54 00	7:54 00
7:44.18	7:53 20	7 49.02	8 00 00	8.00.00
7.51.01	7 56 36	7:55 01	8 06:00	8.06:00
7:58:17	8:04 54	8.04.27	8 12.00	8:12:00
7.55:03	8 08.06	8.07.01	8 18.00	8:18:00

Feb. 24 AAT	ADT	RO	SIM1	SIM2
6:58:49	7.06 44	7 06:03	7:06:00	7.06:00
7:04.19	7.13:42	7.13.02	7.12.00	7.12:00
7:11.07	7.20 12	7:19.02	7.18.00	**7:18:07**
7.20:02	7 26 49	7:25.02	**7:25:00**	**7:24:24**
7:25 49	7 31.43	7 31 02	**7:32:00**	**7:30:41**
7:29 09	7:38:25	7.37.02	**7:39:00**	**7:37:15**
7.40:33	7.45:49	7.44.55	**7:44:24**	**7:43:34**
7.47.17	7:50.19	7 49.02	**7:50:52**	**7:49:53**
7.54 12	7.57 20	7 56:15	**7:57:20**	**7:56:19**
7:56 30	8.03 32	8.01.02	**8:02:00**	**8:01:39**
8.01 17	8 08.14	8 07.02	**8:07:20**	**8:07:06**
8:07.33	8.14:34	8.13:02	**8:12:40**	**8:12:33**
8:13.03	8:20 02	8:18 03	8 18 00	8.18.00

Feb. 25 AAT	ADT	RO	SIM1	SIM2
6.56 53	7.06 40	7.06 01	7:06:00	7.06:00
7 03.34	7 13 36	7.13.02	7 12 00	7.12:00
7:10 30	7 19 49	7:19.02	7:18.00	**7:18:31**
7.15.17	7 28 08	7.25 02	7:24:00	**7:25:02**
7.21 30	7.32 39	7.31 02	**7:31:00**	**7:31:33**
7.29 57	7:37.59	7:37.01	**7:38:00**	**7:38:04**
7.34 14	7.43 44	7:43.02	**7:45:00**	**7:44:35**
7.49.08	7:52.17	7:49 01	**7:52:00**	**7:51:06**
7:50.51	7:57:24	7.55.02	**7:57:12**	**7:56:29**
7 56.46	8.02 18	8.01 03	**8:02:24**	**8:01:52**
7.59 04	8 07.46	8.07.02	**8:07:36**	**8:07:15**
8.03.55	8 14.03	8:13:03	**8:12:48**	**8:12:38**
8.09 25	8 19.02	8 18.02	8 18.00	8:18:00

Table 5- 9 Simulation Results

Table 5- 10 Average Passenger Waiting Time Comparison

	Average Passenger Waiting Time					Average Passenger Waiting Time Saving (%)			
	Ring off (minute)	Eva1 (minute)	Eva2 (minute)	Sim1 (minute)	Sim2 (minute)	Ring off vs. Eva2	Eva2 vs. Eva1	Sim1 vs. Eva2	Sim2 vs. Eva2
Feb. 9	3.05	3.07	3.05	2.97	2.95	0.0%	0.7%	2.6%	3.3%
Feb. 10	3.05	3.09	3.06	2.98	3.01	0.3%	1.0%	2.6%	1.6%
Feb. 11	3.17	3.23	3.20	3.01	3.05	0.9%	0.9%	5.9%	4.7%
Feb. 12	3.07	3.14	3.13	3.06	3.02	1.9%	0.3%	2.2%	3.5%
Feb. 13	3.05	3.08	3.05	3.01	3.01	0.0%	1.0%	1.3%	1.3%
Feb. 17	3.06	3.06	3.05	2.98	3.00	-0.3%	0.3%	2.3%	1.6%
Feb. 18	3.24	3.34	3.20	3.01	2.97	-1.3%	4.2%	5.9%	7.2%
Feb. 19	3.36	3.50	3.47	3.01	2.98	3.2%	0.9%	13.3%	14.1%
Feb. 23	3.25	3.23	3.11	3.01	3.01	-4.5%	3.7%	3.2%	3.2%
Feb. 24	3.13	3.19	3.14	3.05	3.02	0.3%	1.6%	2.9%	3.8%
Feb. 25	3.15	3.27	3.20	3.02	3.00	1.6%	2.1%	5.6%	6.3%
Median	3.13	3.19	3.13	3.01	3.01	0.3%	1.0%	2.9%	3.5%
Mean	3.14	3.20	3.15	3.01	3.00	0.2%	1.5%	4.4%	4.6%
Range	0.31	0.44	0.42	0.09	0.10	7.7%	3.9%	11.9%	12.8%

Table 5-10 summarizes the average passenger waiting time results northbound from Braintree under five scenarios: the ring off times were followed exactly, the actual departure time and travel time northbound (Eva1), the actual departure but simulated travel time northbound (Eva2), and the control system departures with the two assumptions on southbound travel time forecast accuracy (Sim1 and Sim2). The table also compares the actual performance against the performance if the ring off times had been followed precisely, and with the simulated control system departures. The most important conclusions from these experiments are summarized below:

1. From Table 5-10, we find that the simulated control strategy could lead to lower average passenger waiting times than the actual performance. It suggests that our proposed dispatching control system could give MBTA benefits in passenger waiting time, especially when there are large delays or disturbances. We found that the average passengers waiting time saving was up to 14%. The greater the actual passenger waiting time, the greater the savings from the dispatching system. For example, the actual passenger waiting time (Eva1)

on Feb 13 was 3.08 minutes, which was very close to the idea case of 3 minutes, half of the scheduled headway. Clearly, the dispatching control could make no significant improvements in this case. On the other hand, the actual average waiting time (Eva1) on Feb. 19 was 3.50 minutes, and the dispatching control model proposed a strategy, which would reduce this to 3.01 minutes average waiting time (Sim2), for a 14% savings.

2. With respect to the second objective, which is minimizing the number of overcrowded trains, we did not find any overcrowded train in these 11 weekdays in the evaluation and simulation models, because there were no major disruptions on these days. The maximum average waiting time in these days were 3.50 minutes on Feb. 19, which did not result in any overcrowded trains. Even though there were two trains with headway larger than 8 minutes on Feb. 19, the passenger flow at those trains was not large enough to cause overcrowded trains. However, we would expect that the disruption control system would reduce the probability of overcrowded trains when the disruption becomes longer.

3. We also found that there were some minor differences between Eva1 and Eva2. We could not simply conclude that there were biases on the different travel time northbound in the evaluation, just because Eva2 produces shorter average passenger waiting time than Eva1 in some situations. The variation in the travel time forecast that we used in Eva2 might explain the difference between Eva1 and Eva2. However, given the fact that all Eva2 results show shorter average passenger waiting times than Eva1 in Table 5-10, we believe that there are some slight biases in the simulated travel time northbound. Comparing the control system results against Eva2, we still found consistent reduction in the average passenger waiting time.

4. We find that our dispatching control system is robust with respect to errors in southbound travel time forecasts. In Sim2, we used the exact arrival time at Braintree to choose the

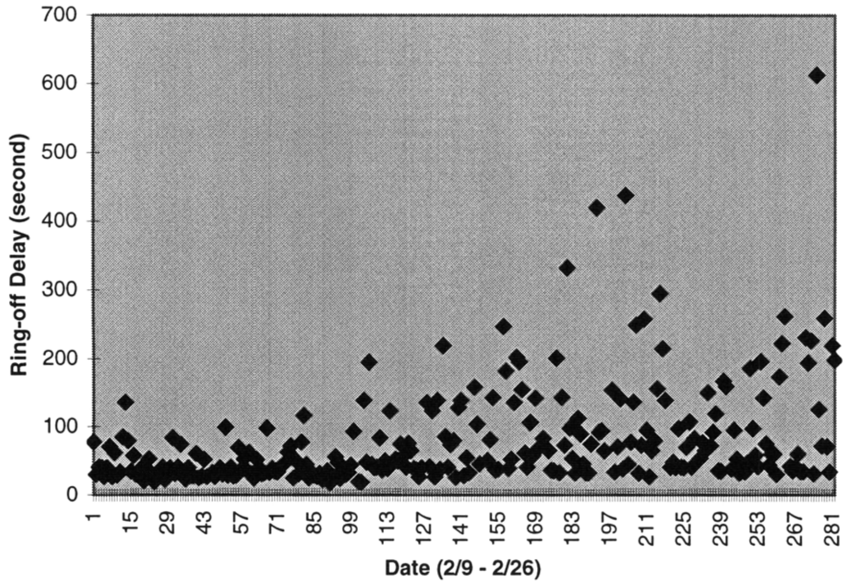
control strategy. Therefore, the result should give us an upper bound on the benefit of the control model. In Sim1, we use the historical mean to estimate the running time and thus choose the control strategy. The result should give us a lower bound on the benefit of our heuristic dispatching control model. We find that the results of these simulations are almost the same, except on Feb 24^{th} and 25^{th}, when Sim2 led to different departure times from Sim1. That means we can gain some small benefits if we can estimate EPDT better.

5. By comparing actual performance with that if the ring off times had been followed perfectly, we found that sometimes the dispatcher's decision might be similar to ours. However the train operator may not follow the ring off time precisely and this can affect the entire operation. For example from Table 5-10, if the operator follows the ring off time exactly, the average passenger waiting time would be 3.07 minutes (Ring off) on Feb. 12, however, the actual waiting time was 3.13 minutes (Eva2).

   There are two possible explanations for the difference between the ring off time and the actual departure time: mechanical problems and ring off discipline. We never expect operators to follow the ring off time exactly because of occasional mechanical problems. For example, a door may malfunction, causing a delay in departure. We can not develop a system to prevent this kind of problems; however, if a late departure resuls, the following ring off times can be adjusted to minimize the impacts.

Figure 5-6 plots the ring-off delay for each Braintree departure for these 11 days. This figure suggests that the operator discipline is poor at Braintree in the AM peak. From the analysis, we find that the average ring-off delay is 78 seconds, but with a standard deviation of 74 seconds.

Figure 5- 6 Difference between Ring off and Actual Departure



Some large delays might be caused by changes in the control system and should be eliminated from our analysis. However, the average and standard deviation of delay will still be large even after excluding those outliers.

Chapter 6

Summary and Conclusion

This chapter will briefly summarize the thesis and suggest directions for future research.

6.1 Summary

Rail transit operations are always subject to various disruptions, which can be categorized as major disruptions and minor disturbances. If these disruptions are not dealt with effectively, they can seriously affect service quality and substantially reduce the attraction of transit service. In this thesis, we developed different strategies to deal with both kinds of problems.

6.1.1 Major Disruption

When an incident lasts more than 20 or 30 minutes, we classify it as a major disruption. Generally, in a major disruption, we will lose part of the track and/or a station. Because we cannot afford to wait until normal operation can be restored, as when a minor disruption occurs, we have to reschedule or redesign the operating plan to run a single-track operation where possible. The objective of the rescheduling in major disruption is to maximize the reduced capacity of the rail system at the most constrained point and to see whether this can carry the volume of passengers traveling. If not, then substitute busing is needed which will typically be costly and ineffective. In Chapter 2, we focused on the single track operation plan redesign for disruptions on the MBTA Red Line to try to avoid busing whenever possible.

Through a series of computations, we find that single track operation is feasible in many Red Line major disruption cases, especially during the off-peak. We can organize two loops and a shuttle, or two overlapping loops to deal with these major disruptions without calling for buses.

However, single track operation will generally not be feasible during the peak period, and we will typically have to call for buses. The results of major disruption analysis might be considered as basis for the major disruption operation control systems at the MBTA. Such a system should eventually be able to assist the supervisors and managers in determining when a single track operation plan will be feasible and which strategy will be most effective.

6.1.2 Dwell Time

When a train arrives at a station, the doors must be opened for a sufficient time to allow all passengers who want to alight and board to do so. This is called the dwell time. Due to the high speed of rail transit systems, the running time between stations is relatively short; however, the dwell time can be relatively long particularly if there is high passenger flow. Moreover, the high variation of the dwell time can be a critical element in increasing headway variation. Therefore, the dwell time is very important for real-time control, simulation and line capacity, and thus is studied in Chapter 4.

Focusing on real-time dispatching control on the MBTA Red Line, we believe that the dwell time is one of the key functions in our entire model, and it will also play an important role in the disruption control system. Our objective was to develop a good dwell time function, based on data obtained from the MBTA. We also expect that the same methodology employed here can be used in other situations.

The dwell time depends on many complex factors, such as passenger flow, congestion in the train, passenger behavior, doorman behavior, and car type, etc. Because of a lack of the data, we could not find any good expression to reflect passenger flow conflicts. However, our simple model $T_{dwell} = T_{open+close} + \alpha \cdot P_{off} + \beta \cdot P_{on}$ produced good regression results based on the 1989 and 1997 CTPS data and 1997 checker data.

6.1.3 Dispatching Control Model

When delays occur at, or around, the terminals, the normal dispatching operation at the terminus will need to be modified. One typical problem is that a train might not arrive at the terminal early enough to be dispatched on schedule. This problem, often referred as the dispatching problem, is critical to the quality of the transit service since it causes uneven headways and increases passenger waiting times, especially when the dispatching direction is the heavy passenger flow direction. Because dispatching problems are often caused by minor disturbances, the dispatching problem can be treated as a special case of the minor disturbance problem. Generally, we have several strategies, including holding, short turning, expressing or deadheading to deal with this kind of disturbance problem.

Based on the discussion in Chapter 3, we found that the expressing and deadheading is generally not attractive for the dispatching problem. Therefore, we focused on holding and short turning in the Chapters 3 and 5. Using the results of mathematical analysis in Chapter 3, we design a heuristic dispatching control model in Chapter 5, which can choose a near optimal strategy involving holding with or without short turning. We considered different objectives, including minimizing passenger waiting time or the number of overcrowded trains in our model.

To evaluate the performance of the dispatching control model, in Chapter 5 we developed a simulation model of the MBTA Red Line Braintree Branch as our case study. The model was used in evaluation mode to assess the performance of the current MBTA dispatching system, while the simulation model was used to predict the performance of the proposed dispatching strategies. Introducing both a travel time estimation function and dwell time function, the simulations were made as accurately as possible.

Results of the simulation experiments showed that our dispatching control model should result in lower passenger waiting time and higher service quality. This dispatching control model will also tend to minimize the number of overcrowded trains. The results also suggest that the passenger waiting time saving will become larger when the operation is subject to greater disturbances. The forecast average passenger waiting time saving ranged up to 14%. Analysis of OCS data also showed that the ring off discipline at Braintree is generally poor, which reduces the effectiveness of the current dispatching system.

6.2 Future Research

It is clear that there is a great amount of additional research needed both in terms of model expansion and exploration of related topics.

The heuristic model in this thesis only considers the dispatching problem, which focuses on control at a single, albeit very important, point. Next, we can extend it to the line level, which would certainly be more difficult but also likely more productive. We may design an operation control model for entire transit line, such as Red Line, including branches. We could eventually extend the model to the network level, to consider smooth transferring between transit lines in the urban network. For example, we may choose minimizing transferring passenger waiting time as an objective in our future model, or trying to link the arrival time of trains on Green Line and Red Line. We believe that the strategy and methods that are found in this thesis are a useful starting point for this research.

In this thesis, we considered only two strategies, holding and short turning. In the future research, other control strategies, such as expressing and deadheading, should also be addressed, especially when the future model deals with the operation on the entire line or network. This work would build on Eberlein's prior analysis of expressing and deadheading which provides valuable insight into these strategies.

Several important elements in the heuristic operation control model should be analyzed further. Dwell time is a very important function in any form of operations control. Its importance will become more significant as the dispatching control system evolves. The key element to estimate a good dwell time function is high quality data. In this thesis, we had to rely on many estimates for key variables, and a new data collection activity is essential if more robust dwell time models are to be developed. Such an effort is vital if more comprehensive and effective real time control strategies are to be developed.

Ring off time is another important element in the operation control model. We have to analyze it carefully to fully understand the determinants of train departure times. The actual departure time will have some distribution about the ring off time. Understanding the factors affecting this distribution is an important first step in finding ways to improve the ring off discipline and thus improve the service quality.

Finally, we may consider determining the schedule dynamically. Using advanced monitoring facilities, in the future we will be able to monitor the passenger flow at each station. A dynamic schedule could be helpful to reduce the number of vehicles and operators, at the same time as increasing the service quality.

References

[1] L. A. Bowman and M. A. Turnquist. Service Frequency, Schedule Reliability and Passenger Wait Times at Transit Stops. *Transportation Research*, Vol. 15A, No. 6, pp.465-471.

[2] Nigel H. M. Wilson and Tyh-Ming Lin. Dwell Time Relationships for Light Rail Systems. *Transportation Research Record 1361*, (1993) 296-304

[3] P. I. Welding. The Instability of a Close-Interval Service. *Opnl. Res. Quart. 8*, (1957)133-142.

[4] P. I. Welding and D. J. Day. Simulation of Underground Railway Operation, *Railway Gazette* 438-441, 4 June 1965.

[5] V. R. Vuchic. Propagation of Schedule Disturbances in Line-Haul Passenger Transportation. *UITP Revue* 18, 281-284 (1969.)

[6] A. Barnett. On Controlling Randomness in Transit Operations. *Transportation Science* Vol.8, 102-116 (1974.)

[7] Osuna E. E. & G. F. Newell. Control Strategies for an Idealized Public Transportation System. *Transportation Science*, Vol.6, 52-72 (1972.)

[8] Arnold Barnett and Daniel J. Kleitman. Optimal Scheduling Policies for Some Simple Transportation Systems. *Transportation Science*, Vol. 7 (85-99.)

[9] Mark A. Turnquist and Steven W. Blume. Evaluating Potential Effectiveness of Headway Control Strategies for Transit Systems. *Transportation Research Record* 746 (1980): 25-29.

[10] Mark Abkowitz and Israel Engelstein. Methods for Maintaining Transit Service Regularity. *Transportation Research Record* 961 (1984): 1-8.

[11] Xu Jun Eberlein. Real-Time Control Strategies in Transit Operations: Models and Analysis. Ph.D. dissertation, Civil Engineering, MIT, 1995.

[12] Susan W. O'Dell. Optimal control Strategies for a rail Transit Line. M. S. thesis, Electrical Engineering, MIT, 1997.

[13] Richard A. Macchi. Expressing Vehicles on the MBTA Green Line. M. S. thesis, Civil Engineering, MIT, 1990.

[14] Anthony A. Deckoff. The Short-Turn as a Real Time Transit Operation Strategy. M. S. thesis, Civil Engineering, MIT, 1990.

[15] David W. Soeldner. A Comparison of Control Options on the MBTA Green Line. M. S. thesis, Civil Engineering, MIT, 1993.

[16] Gordon T. Coor. Analysis of the Short-Turning Strategy on High-Frequency Transit Lines. M. S. thesis, Civil Engineering, MIT, 1997.

[17] G.F.Newell. Dispatching Policies for a Transportation Route. *Transportation Science*, Vol.5, 91-105 (1971.)

[18] Peter G. Furth. Alternating Deadheading in Bus Rout Operations. *Transportation Science*, Vol.19, 13-28 (1985.)

Appendix A

Major Disruption Analysis

A.I Disruptions at stations

Table A-1 Red Line Station Disruption Classification

Blocked station	NB	SB
ALEWIFE	(*)	(*)
DAVIS	(**)	(**)
PORTER	SC,*SF*	SC,*SF*
HARVARD	SC1,*SH1*	SD1,SG1,*SF1*
CENTRAL	SB	SA
KENDALL	SC,*SF*	SC,*SF*
CHARLES	SH	SG
PARK	SG	SH
DOWNTOWN	1SC	1SF
SOUTH	SC1	SF1
BROADWAY	SD,SF	SC
ANDREW	SF1	SC1
JFK—ASHMONT	SA	SB
SAVIN	SC	SE,SF
CORNER	SC	SF
SHAWMUT	(***)	(***)
ASHMONT	(*)	(****)
JFK—BRAINTREE	SA,SB	SA,SB
NORTH QUINCY	SF,SG	SC,*SH*
WOLLASTON	SC,*SF*	SC,*SF*
QUINCY CTR	SC1,*SA*,*SB*,*SF1*	SC1,*SA*,*SB*,*SF1*
QUINCY ADAMS	(**)	(**)
BRAINTREE	(*)	(*)

In Table 1:

X1: State X, but with one more station north of the blocked station;

1X: State X, but with one more station south of the blocked station.

(*): We can only run a long loop including the station that has the blockage;

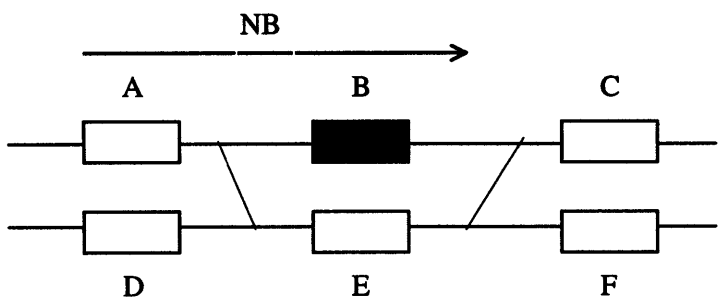
(**): We can run two independent loops;

(***): We can run one loop and one shuttle, which interact; or two loops + shuttle.

(****): We can run one loop and one shuttle, independently;

State SA:

Figure A-1 State SA (disruption happens at a station)



Strategy SA1: Loop1: D-E-A

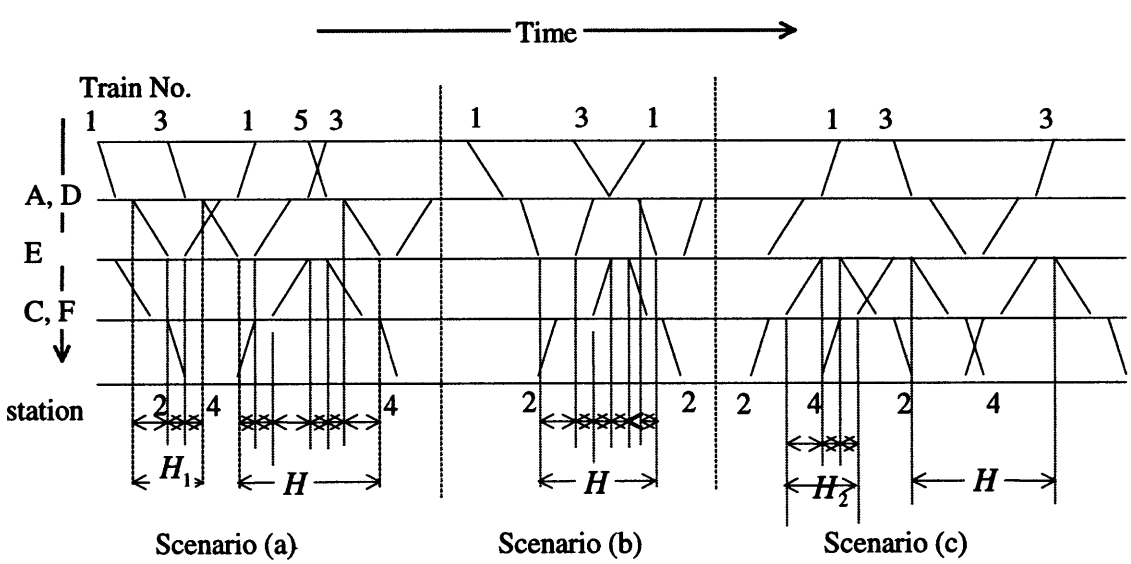
Loop2: C-E-F

Strategy SA2: Loop: D—E—F—...—C—E—A

In State SA, there is a disruption at platform B. There are two possible strategies to deal with this. Strategy SA1 is based on two overlapping loops while SA2 uses a single loop with single track operation through station E. We compute the headway for each strategy as follows:

Strategy SA1 Two Overlapped Loops

Figure A-2 Scenarios of Strategy SA1



t_{dwell} : train's standard dwell time at station;

t'_{dwell} : train's extended dwell time at turning station;

t_{switch} : the time to change the switch at the crossover track;

$t_{x,y}$: Minimum running time from station x to station y based on the control lines;

$t^{signal}_{x,y}$: After the first train leaves station x (heading to station y), the time before the signal clears (including the time for the train to clear the station).

t^{signal}_{x} : The time of train clear station x.

Scenario (a)

In this scenario, the left loop AED has the heaviest passenger flow. We can operate more than one train in the heavy loop, (up to n trains), between consecutive trains running in the light loop. As shown in the diagram, we operate 2 successive trains in loop AED, between successive trains in loop CEF (that is n=2). Then H_1, the minimum headway between trains on loop AED, and H, the headway between trains on loop AED when a train a train on loop CEF intervenes, are as follows:

$$H_1 = t_{D,E} + t'_{dwell} + t^{signal}_{E,A} + t_{switch}$$

$$H = 2 \cdot t'_{dwell} + 2 \cdot t^{signal}_{E} + t_{C,E} + t_{D,E}$$

$H_{AED} = H_1$ or H (alternating for n=2)

$$H_{CEF} = (n-1)H_1 + H$$

$$H_{CEF} = H_1 + H \qquad \text{when n=2}$$

Scenario (b)

In this scenario, the heaviest passenger flow is through station B/E. We must guarantee that the train movements through station E are balanced.

$$H = 2 \cdot t'_{dwell} + 2 \cdot t_E^{signal} + t_{C,E} + t_{D,E}$$

Scenario (c)

In this scenario, loop CEF has the heaviest passenger flow. Similar to scenario (a,) but H_1 is replaced by H_2:

$$H_2 = t_{C,E} + t'_{dwell} + t_{E-F}^{signal} + t_{switch}$$

$$H = 2 \cdot t'_{dwell} + 2 \cdot t_E^{signal} + t_{C,E} + t_{D,E}$$

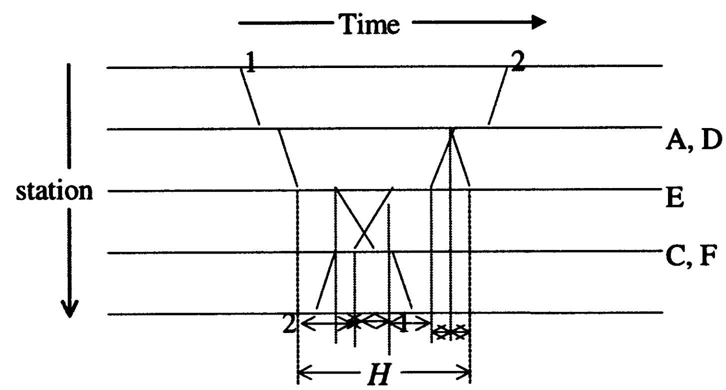
$$H_{CEF} = H_2 \text{ or } H \qquad (\text{alternating for n=2})$$

$$H_{AED} = (n-1)H_2 + H$$

$$H_{AED} = H_2 + H \qquad \text{when n=2}$$

Strategy SA2 One Single Loop

Figure A-3 Strategy SA2

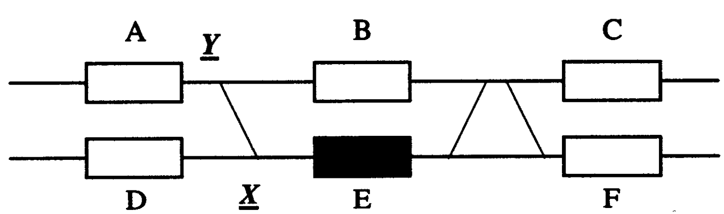


No matter where the heaviest passenger flow is, strategy SA2 has just a single headway.

$$H = 2 \cdot t_{dwell}^{E-NB} + t_{E,F}^{signal} + 2 \cdot t_{switch} + t_{C,E} + t_{E,A}^{signal} + t_{D,E}$$

State SB

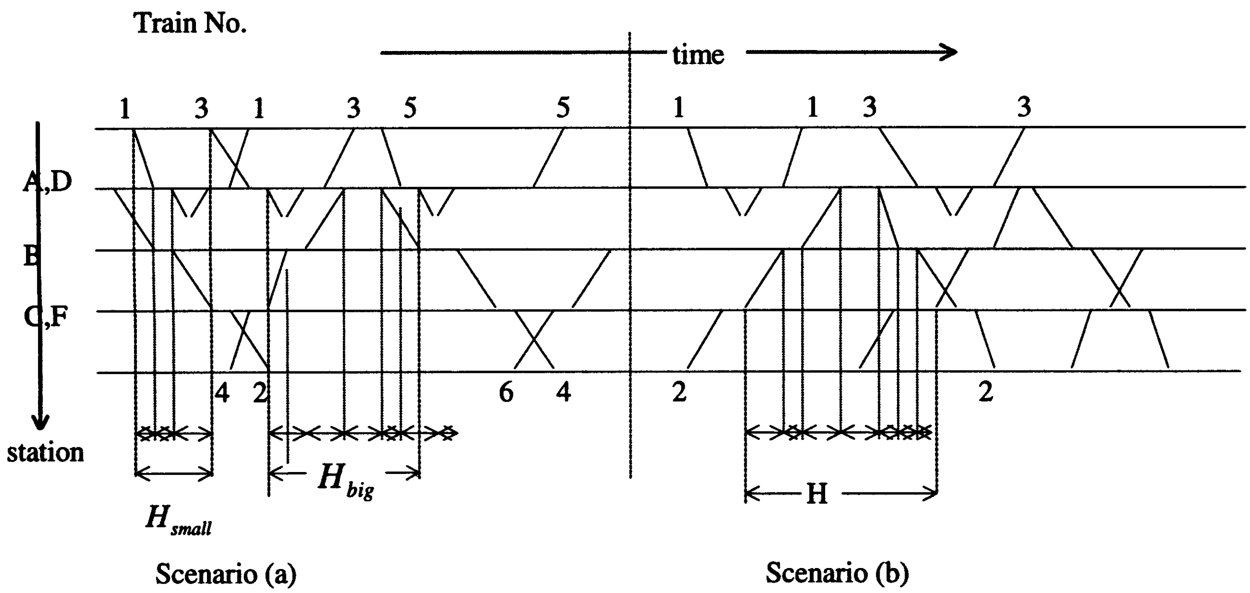
Figure A-4 State SB (disruption happens at station)



Strategy SB1: Loop1: D—X—Y—A

Loop2: C—B—A—B—F

Figure A-5 Scenarios of Strategy SB1



Scenario (a)

In this scenario, the loop DXYA has the heaviest passenger flow. We will operate more trains in the heavy flow loop at the headway H_{small}, for example n trains, then we run a train in the light flow loop, this will intervenes the headway of the loop DXYA from H_{small} to H_{bi}.

$$H_{small} = t_{,D} + t'_{dwell} + t^{signal}_{D-X-A} + t_{switch}$$

$$H_{big} = \max(t^{signal}_{D-X-A} + t_{switch} + t'_{dwell}, t_{B,A} + t^{signal}_{D,Y}) + t'_{dwell} + t^{signal}_{A,Y} + t_{switch}$$

$H_{DXYA} = H_{small}$ or H_{big}. (alternating for n=2)

$$H_{CBABF} = (n-1)H_{small} + H_{big}$$

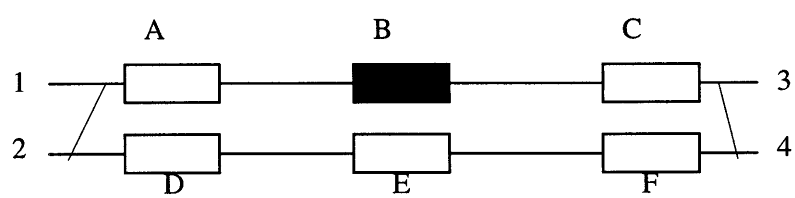
Scenario (b)

In this scenario, the blockage station or the right loop has the heaviest passenger flow. In these two cases, we must guarantee that we can operate the maximum number of trains that run on the right loop.

$$H = t_{C,B} + 2 \cdot t_{dwell} + t_{B,A} + t_{A,B} + t'_{dwell} + t^{signal}_{B-F} + t_{switch}$$

State SC

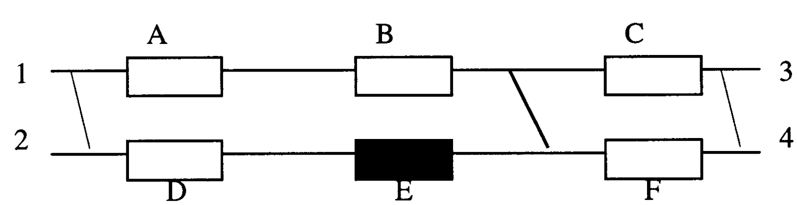
Figure A-6 State SC (disruption happens at station)



Strategy SC1: Loop1: 2—A—1
Loop2: 3—C—4
Shuttle: D—E—F

State SD

Figure A-7 State SD (disruptions happens at station)

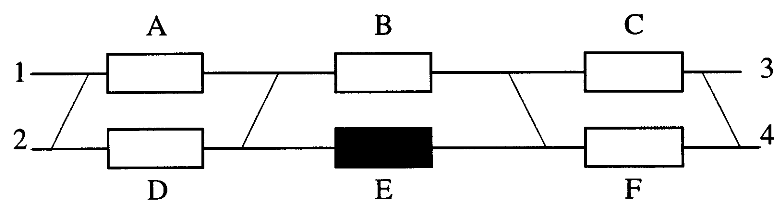


Strategy SD1: Loop1: 2—D—1
Loop2: 3—C—4
Shuttle: A—B—F

Strategy SD2: Loop1: 2—D—1
Loop2: C—B—A—B—F
This strategy is as same as state SG1

State SE

Figure A-8 State SE (disruption happens at station)



Strategy SE1: Loop1: 2—A—1

Loop2: 3—C—4

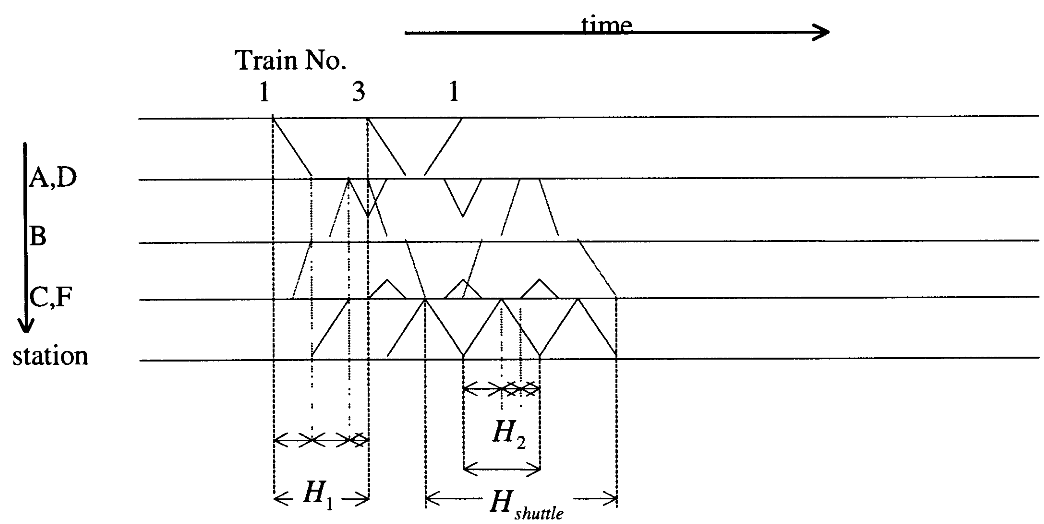
Shuttle: D—B—F

Strategy SE2: Loop1: D—B—A

Loop2: C—B—F

This is similar to the state SA1

Figure A-9 Strategy SC1, SD1 and SE1



In above situations, we can operate two loops at two side of the blockage, and run a shuttle through the blockage station on single track. The headways of loops and shuttle are as follows:

$$H_1 = t_{,A} + t'_{dwell} + t^{signal}_{A,} + t_{switch}$$

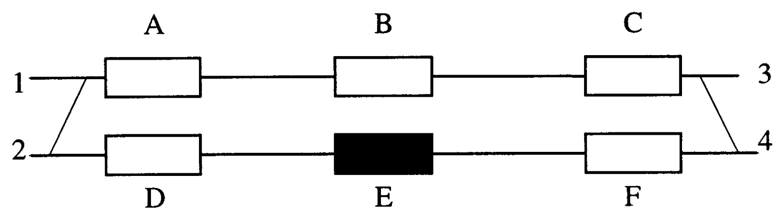
$$H_2 = t_{,C} + t'_{dwell} + t^{signal}_{C,} + t_{switch}$$

$$H_{shuttle} = \sum_{z=D}^{E} t_{z,z+1} + \sum_{z=D}^{E} t_{z+1,z} + 2 \cdot t_{dwell}^{'} + 2 \cdot t_{dwell}$$

In this two loops + shuttle operation plan, each route (loop or shuttle) operates independently, so the solution applies to all passenger flow configurations.

State SF

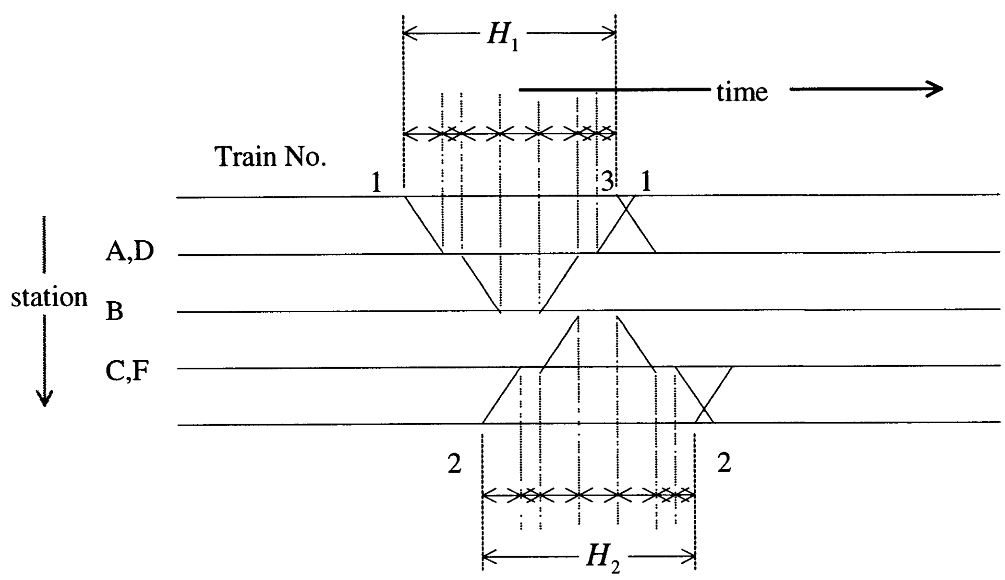
Figure A-10 State SF (disruption happens at station)



Strategy SF1: Loop1: 2—A—B—A—1

Loop2: 3—C—B—C—4

Figure A-11 Strategy SF1



In this case, we might also have different passenger flow levels, but because these two headways are relatively long, we will not run two consecutive trains in one direction in general. The headways are as follows:

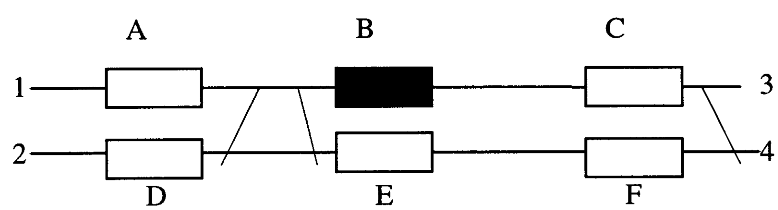
$$H_1 = t_{,A} + 2 \cdot t_{dwell} + t_{A,B} + t'_{dwell} + t_{B,A} + t^{signal}_{A,}$$

$$H_2 = t_{,C} + 2 \cdot t_{dwell} + t_{C,B} + t'_{dwell} + t_{B,C} + t^{signal}_{C,}$$

For this state, we also can run a big loop. However, the headway will be very long. Therefore, we eliminate this strategy.

State SG

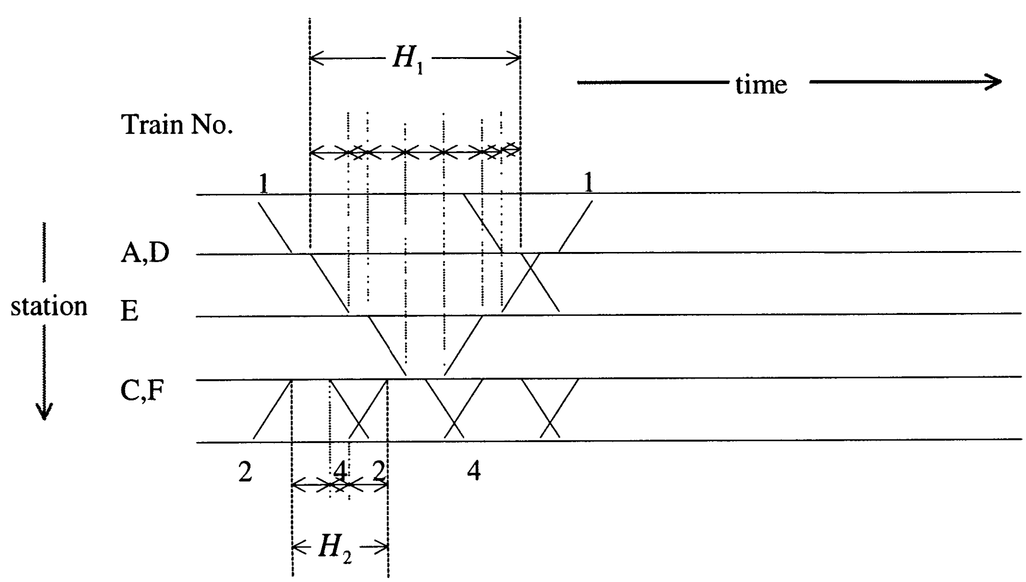
Figure A-12 State SG (disruption happens at station)



Strategy SG1: Loop1: D—E—F—E—A

Loop2: 3—C—4

Figure A-13 Strategy SG1

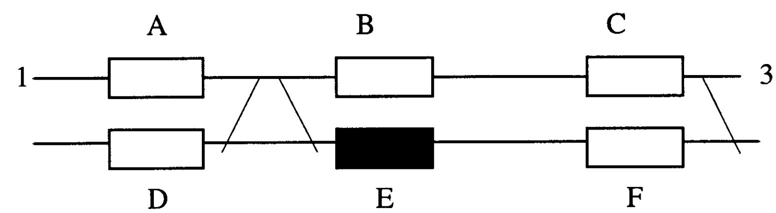


$$H_1 = t_{A,E} + 2\cdot t_{dwell} + t_{E,F} + t'_{dwell} + t_{F,E} + t^{signal}_{E,A} + t_{switch}$$

$$H_2 = t'_{dwell} + t^{signal}_{C,} + t_{switch} + t_{,C}$$

State SH

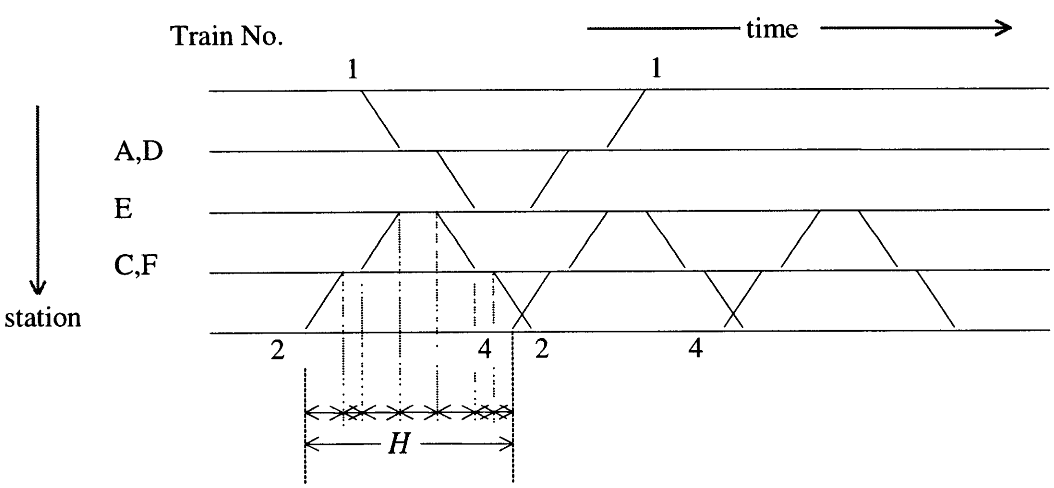
Figure A-14 State SH (disruption happens at station)



Strategy SH1: Loop1: 2—D—B—A—1

Loop2: 3—C—B—C—4

Figure A-15 Strategy SH1



$$H = t_{,C} + 2\cdot t_{dwell} + t_{C,B} + t'_{dwell} + t_{Bi,C} + t^{signal}_{C,} + t_{switch}$$

A.II Disruptions between stations

Table A-2 Red Line Inter-Station Disruption Classification

Blocking Location	**NB**	**SB**
ALEWIFE—DAVIS	(*)	(*)
DAVIS—PORTER	1TA	1TB,TC
PORTER—HARVARD	TA1,T*B1*	TA1,T*B1,TC*
HARVARD—CENTRAL	TA2	TB2
CENTRAL—KENDALL	1TA,*1TB*	1TA,*1TB*
KENDALL—CHARLES	1TA,TB,TC	TA,1TB,T*C,TD*
CHARLES—PARK	1TA2	1TB2,1TC1
PARK—DOWNTOWN	TA	TB,1TC
DOWNTOWN—SOUTH	1TA1	1TB1
SOUTH—BROADWAY	TA2	TB2,TD
BROADWAY—ANDREW	TB1,TD1	TA1
ANDREW—JFK	TB	TA
JFK—SAVIN	TA	TB,TD
SAVIN—CORNER	TA	TB
CORNER—SHAWMUT	TA1	TB1
SHAWMUT—ASHMONT	(**)	(*)
JFK—NORTH QUINCY	1TA,*1TB*	1TA,*1TB*
NORTH QUINCY—WOLLASTON	TA,T*B,TD*	TA,T*B,TD*
WOLLASTON—QUINCY CTR	TA1,T*B1*	TA1,T*B1*
QUINCY CTR—QUINCY ADAMS	TA,T*B,TD*	TA,T*B,TD*
QUINCY ADAMS—BRAINTREE	(*)	(*)

In Table 2:

Xn: State X, but with n more stations north of the disruption;

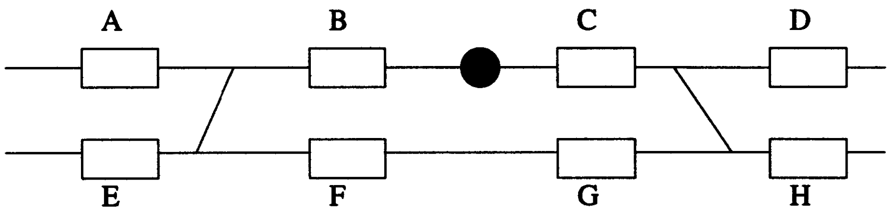
nX: State X, but with n more stations south of the disruption;

(*): We can run one loop and one shuttle, independently;

(**): We can run one loop and one shuttle, which interact;

State TA

Figure A-16 State TA (disruption happens between stations)

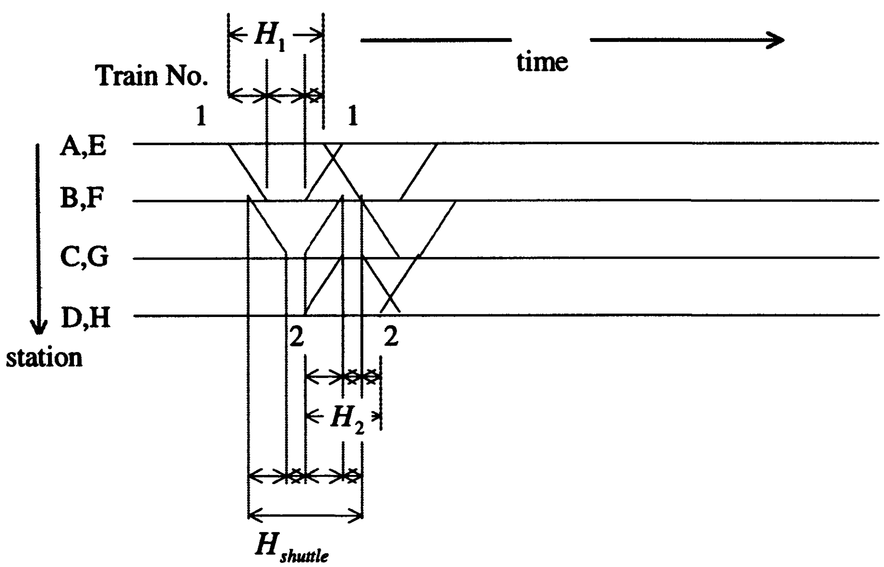


Strategy TA1: Loop1: E—B—A

Loop2: D—C—H

Shuttle: F—G

Figure A-17 Strategy TA1



$$H_1 = t_{A,B} + t_{dwell}^{'} + t_{B,A}^{signal} + t_{switch}$$

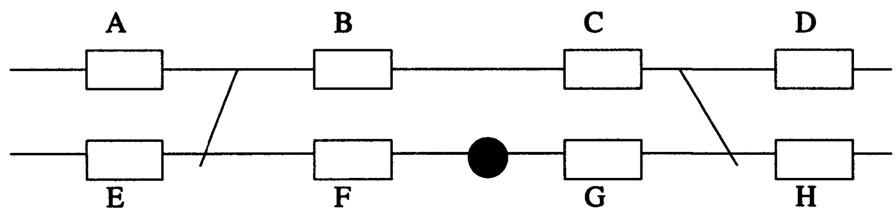
$$H_2 = t_{D,C} + t_{dwell}^{'} + t_{C,D}^{signal} + t_{switch}$$

$$H_{shuttle} = t_{F,G} + 2 \cdot t_{dwell}^{'} + t_{G,F}$$

In State TA, each route (loop or shuttle) operates independently, so the solution applies to all passenger flow configurations.

State TB

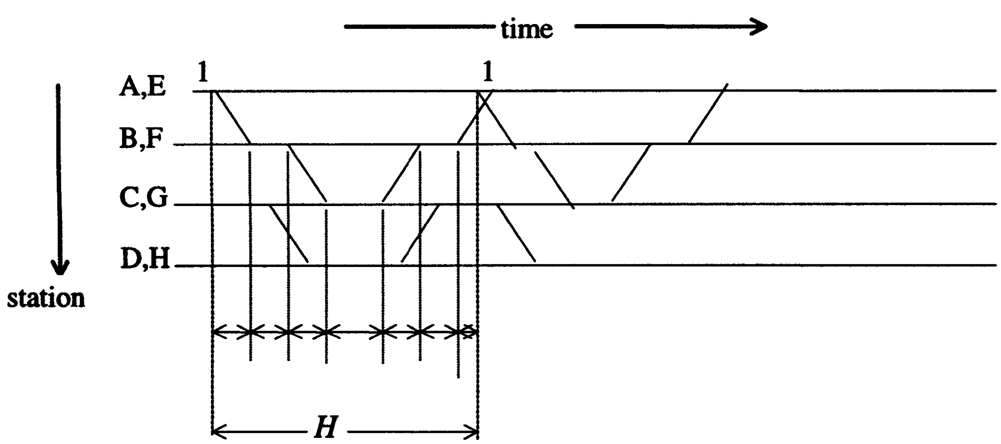
Figure A-18 State TB (disruption happens between stations)



Strategy TB1: Loop1: E—B—C—B—A

Loop2: D—C—H

Figure A-19 Strategy TB1



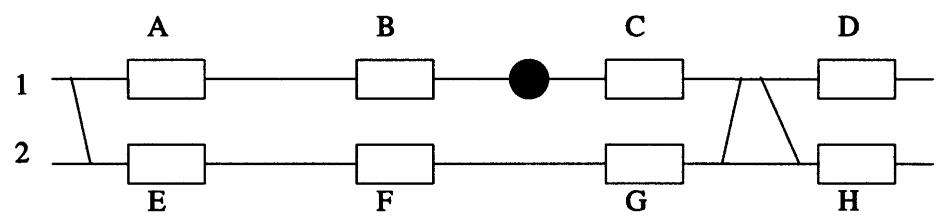
$$H = t_{A,B} + t_{dwell} + t_{B,C} + t'_{dwell} + t_{C,B} + t^{signal}_{B,E} + t_{switch}$$

Comparing to this headway, the headway of another loop DCH is relatively short, then we can run trains on the space when no train occupies the station, unless the train is not controlled by the another loop train. The headway formula is as follows:

$$H_{small} = t_{D,C} + t'_{dwell} + t^{signal}_{C,H} + t_{switch}$$

State TC

Figure A-20 State TC (disruption happens between stations)



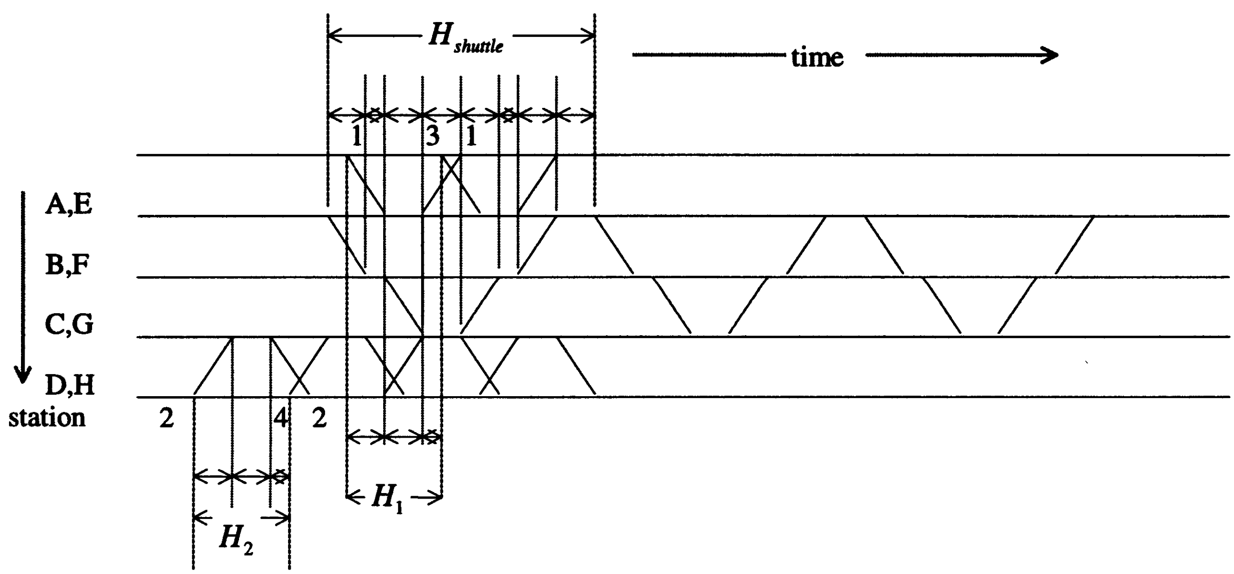
Strategy TC1: Loop1: 2—E—1

Loop2: D—C—H

Shuttle: E—F—G

Figure A-21 Strategy TC1

Two loops + overlapped shuttle



$$H_1 = t_{,E} + t'_{dwell} + t^{signal}_{E,} + t_{switch}$$

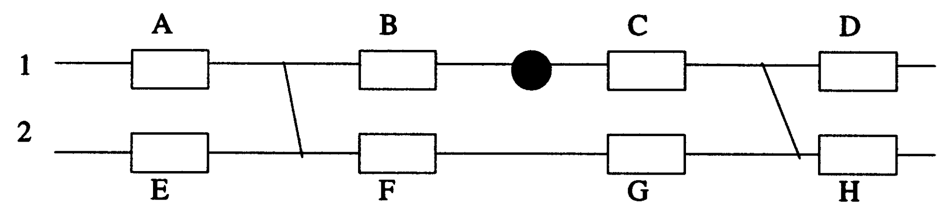
$$H_2 = t_{D,C} + t'_{dwell} + t^{signal}_{C,H} + t_{switch}$$

$$H_{shuttle} = t_{E,F} + 2 \cdot t_{dwell} + t_{F,G} + 2 \cdot t'_{dwell} + t_{G,F} + t_{F,E}$$

The headway of the shuttle is very vulnerable, because the train on loop 2E1 can control it very frequently. If the control lines that are involved in the loop 2E1 are very long, the headway of the shuttle will be very long.

State TD

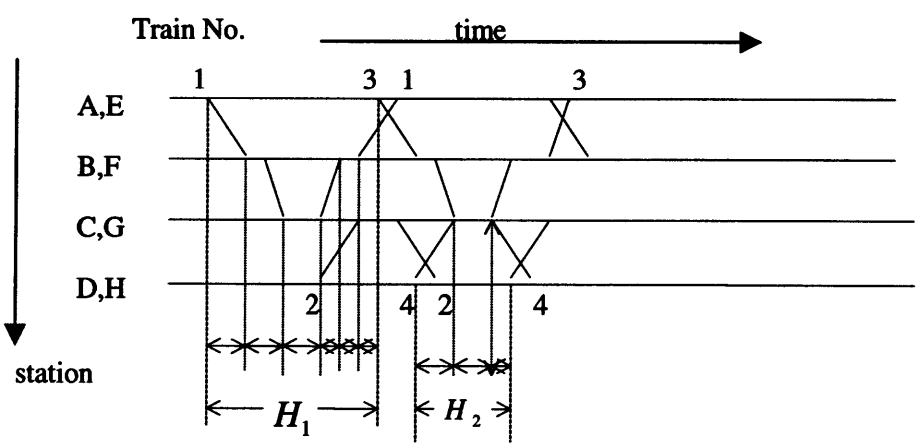
Figure A-22 State TD (disruption happens between stations)



Strategy TD1: Loop1: E—F—G—F—A
Loop2: D—C—H

In this situation, loop EFGFA and loop DCH overlapped but do not intervene with each other

Figure A-22 Strategy TD1



$$H_1 = t_{A,F} + 2 \cdot t_{dwell} + t_{F,G} + t'_{dwell} + t_{G,F} + t_{F,A}^{signal} + t_{switch}$$

$$H_2 = t_{D,C} + t'_{dwell} + t^{signal}_{C,H} + t_{switch}$$

When the disruption happens between the stations, it is not so important where the heaviest passenger flow is for two reasons:

1. We can find the best solution independent of where the heave flow is, for example two loops + shuttle.

2. When the headways of both loops are long, it is hard to believe that we can run two consecutive trains in one direction, because people on the other loop would have to wait for a very long time.

Appendix B

Su Shen's Travel Time Estimation Report

1. Procedure to compute EPDT (in seconds)

   1.1 Estimate the departure time at Quincy Adams southbound when train arrives at station k (k may be JFK, NQ, Wol, QC southbound)

   a. If the headway is smaller than the thresholds in the following table, train[I].EDT(QA) = train[I-1].EDT(QA)+180

   SMALL HEADWAY: headway less than:

Type \ Station	JFK	NQ	WOL	QC
Regular train	180	180	180	180
Pull-out train	270	245	220	200

   b. otherwise

   Using the formula for station k to estimate the travel time TT(k~QA) for current train from station k to Quincy Adams (in seconds.)

Station k	Formula (in seconds)	
JFK	$376.6 + 0.12 H_l + 0.539 ST - 108 Out + 22.4 \Pr e$	($\overline{R}^2 = 0.73$)
North Quincy	$216.5 + 0.084 H_l + 0.530 ST - 89.2 Out + 19.0 \Pr e$	($\overline{R}^2 = 0.75$)
Wollaston	$145.2 + 0.059 H_l + 0.576 ST - 62.1 Out + 14.8 \Pr e$	($\overline{R}^2 = 0.67$)
Quincy Center	$138.4 + 0.027 H_l + 0.31 ST - 49.2 Out + 6.8 \Pr e$	($\overline{R}^2 = 0.56$)

   And train[I].EDT(QA) = train[I].AT(k)+ train[I].TT(k~QA)

   i.

$$H_l : \begin{cases} H(k)-360, & \text{if } H(k)>360 \\ 0, & \text{otherwise} \end{cases}$$

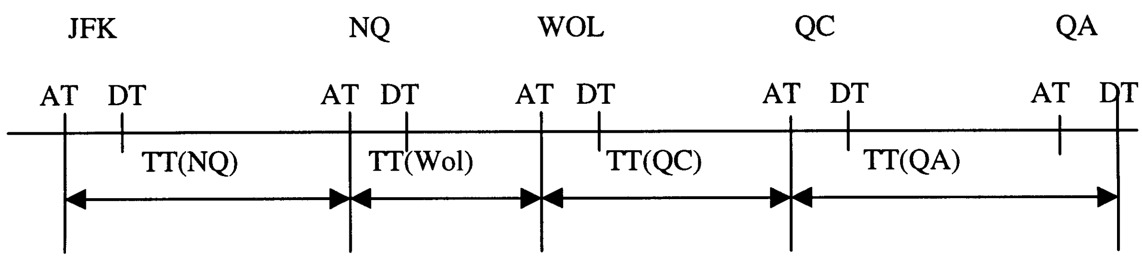
ii.

$$Out: \begin{cases} 1, \text{ if train i is pullout train. Specifically, if DW(k)<15 seconds} \\ 0, \text{ otherwise} \end{cases}$$

iii.

$$\text{Pre:} \begin{cases} 1, \text{ if the preceding train of train i is a pull out train.} \\ 0, \text{ otherwise} \end{cases}$$

iv. ST:



The initial value for smoothed travel time $\tilde{TT}$ is based on the historical mean (see the following table).

Station	NQ	Wol	QC	QA
Value	**361**	**122**	**162**	**200**

Smoothing the travel time:

① If trip i is made by a pull-out train or a small-headway train, we set station[k]. $\tilde{TT}$ =station[k]. $\tilde{TT}$.

② Otherwise, station[k]. $\tilde{TT}$ =0.6train[I].TT(k) + 0.4station[k]. $\tilde{TT}$

$$ST = \sum_{k=NQ,Wol,QC,QA} \tilde{TT}$$

1.2 Estimate the arrival time of train i at BR based upon queuing model

a. If train[I].EDT(QA) + 80 < train[I-2].DDT, train[I].EPDT = train[I-2].DDT + 135 + minimum recovery time.

b. Otherwise, train[I].EPDT = train[I].DT(QA) + 190 + minimum recovery time.

2. From each station to QC southbound (arrival time). Follow the similar procedure to compute the departure time at QA

   Here, the small headway definition is:

Type \ Station	JFK	NQ	WOL
Regular train	**180**	**180**	**180**
Pull-out train	**240**	**215**	**200**

The TT(k~QC) can be gotten by the following formulas.

Station	Formula (in seconds)
JFK	$238 + 0.077 H_l + 0.617\, ST2 - 66.4\, Out + 14.2\, Pre$
NQ	$84.8 + 0.043 H_l + 0.678\, ST2 - 44.2\, Out + 10.8\, Pre$
Wol	$23.3 + 0.020 H_l + 0.832\, ST2 - 18.7\, Out + 6.5\, Pre$

$ST = \sum_{k=NQ,Wol,QC} \tilde{TT}$. We only exclude the travel time between Quincy Center and Quincy Adams.

Appendix C

CTPS RED LINE SOUTH SHORE BRANCH PASSENGER COUNTS AND ANALYSIS

The accompanying spreadsheets provide composite one-day counts of train-by-train passenger boardings and alightings on all Red Line South Shore Branch trains scheduled to leave or arrive at Braintree from about 6:30 a.m. to 9:30 p.m. on a weekday. (Depending on checker availability, results for some stations show earlier starting or later ending times.)

Rows with data but with no time shown in the scheduled departure column are extra Run-As-Directed (RAD) trips. Since these do not operate at the same time every day, they do not have passenger counts for all stations. Blank entries for ons and offs for a scheduled trip at an individual station indicate that the trip either did not operate or bypassed that station on the count day.

Line volumes between stations on northbound trips were calculated by adding ons and subtracting offs at all preceding stations. Line volumes on southbound trips were calculated by subtracting total ons from total offs at all following stations.

Counts at stations from JFK/UMass through South Station in these tables are for Braintree Branch trains only. Counts for Ashmont Branch trains are contained in separate tables which are still being revised.

The count dates were as follows:

Station	Start to 2:00 p.m.	2:00 to 9:30 p.m.	After 9:30 p.m.
Braintree	May 2, 1997	April 28, 1997	

Quincy Adams	May 2, 1997	April 28, 1997	
Quincy Center	May 2, 1997	April 28, 1997	Oct. 29, 1997
Wollaston	May 2, 1997	May 1, 1997	
North Quincy	May 2, 1997	April 28, 1997	April 28, 1997
JFK/UMass	May 8, 1997	May 5, 1997	Oct. 23, 1997
Andrew	May 7, 1997	April 30, 1997	
Broadway			
Northbound	May 8, 1997	May 5, 1997	
Southbound	May 15, 1997	May 8, 1997	
South Station	May 7, 1997	May 6, 1997	Oct. 23 or 30, 1997

Note that most of these counts were taken prior to the opening of the Old Colony commuter rail lines, which was expected to divert some riders from Red Line stations.

Result of CTPS Data Analysis

CTPS data on May 2, 1997

Sch. Bra NB	Obs Bra NB	Bra LV out	QA LV out	QC LV out	Wol LV out	NQ LV out	Obs NQ NB	DH Bra	AH NQ	Aver. Headway	Pass/min.	After smoothing	Obs Time Braintree N
650	652	208	280	388	514	611	705	9	10	9.5	**64**	**68**	652
658	659	160	245	322	450	532	711	7	6	6.5	**82**	**92**	659
706	703	229	362	579	706	829	718	4	7	5.5	**151**	**90**	703
713	714	181	328	402	524	581	726	11	8	9.5	**61**	**107**	714
719	720	208	356	607	745	834	732	6	6	6	**139**	**94**	720
725	726	149	322	393	534	647	739	6	7	6.5	**100**	**121**	726
731	732	149	320	486	617	692	744	6	5	5.5	**126**	**111**	732
737	738	154	366	455	564	664	750	6	6	6	**111**	**121**	738
743	744	124	386	517	625	759	756	6	6	6	**127**	**109**	744
749	750	49	249	397	484	648	804	6	8	7	**93**	**120**	750
755	755	41	261	363	494	568	806	5	2	3.5	**162**	**102**	755
801	801	45	235	311	386	513	813	6	7	6.5	**79**	**113**	801
807	808	83	364	633	746	845	820	7	7	7	**121**	**101**	808
813	814	59	284	397	518	659	827	6	7	6.5	**101**	**106**	814
818	819	0	144	273	385	511	833	5	6	5.5	**93**	**106**	819
826	827	127	238	384	484	593	839	8	6	7	**85**	**70**	827
834	836	50	124	257	351	466	849	9	10	9.5	**49**	**57**	836
842	843	3	49	149	192	255	855	7	6	6.5	**39**	**45**	843
	846	1	23	116	116	116	857	3	2	2.5	**46**	**40**	846
850	852	18	41	86	108	164	900	6	3	4.5	**36**	**53**	852
902	903	14	84	213	291	386	916	11	16	13.5	**29**		903

- In the above table, the values in 1st shaded column show the cumulative number of passengers/min on Braintree branch

- For example, the second value 82 means from 6:52 to 6:59, there are 82 passengers/min from Braintree branch to Trunk part

- The 2nd shaded column shows the values of passengers/min after smoothing over the trains ahead and behind the current train.