A Real-Time Train Holding Model for Rail Transit Systems

by

André Puong

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 2001

Author ....
Department of Civil and Environmental Engineering
June 27, 2001

Certified by......   ....
Nigel H.M. Wilson, Ph.D.
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by ...   .
Oral Buyukozturk
Chairman, Department Committee on Graduate Students

A Real-Time Train Holding Model for Rail Transit Systems

by

André Puong

Submitted to the Department of Civil and Environmental Engineering
on June 27, 2001, in partial fulfillment of the
requirements for the degree of
Master of Science in Transportation

Abstract

Rail transit systems are subject to frequent disruptions caused by a variety of disturbances such as a disabled train or a door jam. Such disruptions often last for 10 to 20 minutes and can severely impact passenger level of service for long after the blockage has been cleared. Transit agencies usually employ various train control strategies such as train holding and short-turning to respond to these disruptions. The efficiency of these controls strategies relies upon the system-wide impacts of any control action. Unfortunately, it is extremely difficult for a dispatcher to assess such impacts and derive the best control actions in-real time.

This thesis focuses upon the development of a real-time disruption control model for a rail transit loop line. Holding and short-turning are studied as means to minimize the sum of passenger in-platform waiting time and in-vehicle delay time. A simple holding strategy is first introduced for a simplified subway system and a deterministic mathematical programming formulation is derived. The formulation is rewritten in matrix form, providing insights into the behavior of a more realistic system. The original formulation is next extended to a more realistic, albeit deterministic, holding problem for a general transit loop line. A quadratic and mixed integer programming formulation is obtained and a procedure to efficiently solve it is presented.

The formulation is applied to two disruption scenarios on a simplified system based on the MBTA Red Line. The sensitivity of the optimal holding strategy to the assumption of finite train capacity and the cost of in-vehicle time is also investigated. Results confirm that evening headway sequences at stations is generally equivalent to minimizing passenger waiting time but this goal is constrained by train capacities and limited by the cost of holding. It is also shown that accounting for holding costs leads to simple optimal strategies wherein a few early control actions are exerted on a few trains and terminal holding is preferred, with significant associated time savings (19-51%). All the problem instances are solved in less than ten seconds using a two-step solution procedure.

Finally, the short-turn strategy is studied in more detail. Guidelines are given for determining efficient short-turning options and exactly assessing their benefits using the developed holding model in some cases.

Thesis Supervisor: Nigel H.M. Wilson, Ph.D.
Title: Professor of Civil and Environmental Engineering

Acknowledgments

First, I would like to thank my advisor, Prof. Nigel Wilson, for his help, guidance and motivation throughout my graduate studies at MIT. I admire his intelligence, rigor, and the high standards that he has - both for himself and his students. His suggestions, criticisms and careful review of my work have made for a better overall thesis.

I would also like to thank the Office of International Studies at the Ecole Centrale Paris and especially Ms. Elinor Sigler, for their support while completing my studies at MIT. Thanks also to the MBTA for funding my research.

My sincere and deepest gratitude also go to my friends without whom my time at MIT would have been less enjoyable. I want to thank my friends from CTS who played pinball games, shared meals at nearby restaurants and pulled all-nighters with me, making the time spent in the computer lab more bearable. From the Ashdown residence where I lived in my second year, I want to thank friends who invited me on countless occasions for a drink or a chat in their room. Thanks also to all my friends who did not waver in their friendship despite the distance and kept distracting me from work with funny emails, instant messages or telephone calls.

Most of all, I am greatly indebted to my parents for their continuous support. They have encouraged me in all my choices and endeavors: my achievements so far would not have been possible without their love and sacrifices. To them I dedicate this piece of work.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

High-frequency rail transit systems have been playing an increasing role in urban transportation, reducing traffic congestion, creating livable communities and meeting the need for basic mobility. Nonetheless, they are not immune to disturbances that can impact their reliability as an alternate urban transport mode. Especially during peak periods, disturbances on a transit line may result in overcrowded trains, long delays and severe degradation in the level of service provided.

Transit agencies have made increasing use of better information technologies including automatic vehicle location (AVL) and automatic vehicle monitoring (AVM) systems to monitor and regulate train operations. While advanced train control systems can automatically recover from minor disturbances, longer delays must still be handled using dispatchers' experience and judgment. Recent research to improve the train control decision making process has not yet led to implementable models, even when assuming that all the disruption parameters are known deterministically.

This thesis reports on real-time control strategies for rail transit systems to deal with disruptions. It presents a model that considers holding and that can be compatible with a real-time decision-making process. The model is tested on a modified system based on the Massachusetts Bay Transportation Authority (MBTA), and findings from the test disruption scenarios are presented. In particular, the sensitivity of the optimal holding strategy to vehicle capacity and the cost of in-vehicle delay is thorougly examined. It is also shown that this model can serve as a basis for evaluating some types of short-turning strategies.

1.1 Motivation

Heavy rail transit lines are subject to frequent disruptions or delays that can severely impact passenger level of service. According to Song [31], there were approximately 323 incidents or disruptions of at least 10 minutes duration on the MBTA Red Line during a two-year period ending in 1996. This averages three disruptions per week, with most of the recorded delay durations between 10 and 20 minutes.

Disruption durations longer than 20 minutes are generally caused by disabled trains, but can also be associated with bomb threats, fires on trains or other severe technical problems. In such disruption cases, part of a track and/or a station is lost and changes in operating plan are needed such as introducing substitute bus service.

Disruptions with shorter durations, such as 10 to 20 minutes, are more frequent and are typically caused by a disabled train, door jam or malfunctioning signal. For these types of disruption, substitute bus service would be costly and non-responsive. Thus, transit agencies usually use real-time control strategies to deal with these types of delays in the case of both bus and rail. Common strategies include holding a vehicle at a station, expressing a vehicle over a segment of the line, and short-turning a vehicle. For a better understanding of the model presented in this thesis and what it purports to do, a more detailed description and analysis of these control strategies is presented in Section 1.3.

Existing automatic train regulation systems control vehicle speeds and/or spacings between vehicles in order to achieve adherence to trains' dispatching schedule and to maintain regular headways between trains. These systems strive to address minor deviations from the operating plans and the scheduled headway sequences between trains. Yet, none of them specifically address the cases of longer disruptions that require train control strategies such as the ones mentioned above.

This lack is compounded by the large amount of information to consider and the difficulty of quickly assessing the system-wide costs and benefits of any train control action. Not only must dispatchers know the location of the trains relative to the disruption location, but they also need to know the train loads and the track configuration in order to select appropriate control strategies to minimize the service impacts. Making effective decisions in a short time is extremely difficult given the unpredictable and stochastic nature of transit systems and because any decision can have major repercussions down the line and on future operations. Hence, train control decisions are usually made based on judgment and experience.

Therefore, there is a need for automatic train control DSS that would determine in real-time the optimal train controls to apply to minimize the system-wide effects of a disruption. With the increasing use and improvement of automatic vehicle location (AVL) and automatic vehicle monitoring (AVM) systems, there are great opportunities for the development and integration of such tools into existing train regulation systems.

The research presented in this thesis is motivated by recent research (Shen [30] and O'Dell [26, 27]) that formulated the disruption recovery problem as deterministic mathematical programs in the case of heavy rail transit. The objective was to minimize passenger waiting time (and on-board delay in [30]), given a known fixed delay duration. Although these formulations led to interesting findings and recommendations for heavy rail control, the required times to solve the control problem may not be compatible with real-time implementation of the models. Compounding this was the assumption of a fixed delay duration, as this parameter is generally unknown. Moreover, relaxing this assumption would yield a more complex problem and almost certainly much larger solution times, thus affecting the real-time tractability of any non-deterministic train control problem.

The model developed in this research determines the optimal train holding strategy. The holding model is implemented using commercial optimization software and yields solution times that are compatible with real-time decision making. It can be applied on any non-branching rail transit line with real-time train location information available. The model formulation can also serve as a basis for a real-time stochastic train control model that considers the uncertainty of the delay duration.

1.2 Review of Prior Work

1.2.1 Literature Review

The earliest research in the field of transit control strategy stemmed from mathematical studies of simplified transportation systems. These early studies drew on fields of applied mathematics such as queuing theory, Markov chains, game theory, stochastic processes and *Monte Carlo* simulation (see [1, 5, 4, 23, 28, 35]). One consistent finding from this early work was the complexity of the mathematical models involved and the difficulty of obtaining exact analytical results. Moreover, because of the lack of AVL systems at the time of those studies, only a few parameters of the system could be assumed to be known in real-time and available to analyze disruptions and consequently derive optimal train control strategies. The solution to these train control problems usually depended upon macro-level parameters such as the mean and variance of the headway distribution, the correlation between successive arrivals, or the ratio of passengers benefited to passengers disbenefited by the control strategy (see [4]). Of particular interest is the work presented by Barnett [5] who considers a rudimentary system with only one infinite-capacity vehicle, and all passengers boarding at the same stop and having the same travel cost structure. Based on this simple system and other economic constraints, Barnett derived a simple holding strategy (holding the vehicle until a threshold headway) that solves an apparently complex mathematical problem. Although his result was purely analytical, he suspected that any strong analytical results for larger and more complicated systems would be most difficult to obtain.

Far from being discouraged by those results, researchers further investigated the opportunity of using real-time information in modeling complex transportation systems. Thus, the locus of interest shifted from seeking exact analytical results to the application of mathematical optimization methods.

Turnquist [34] first described an advanced schedule control model which illustrates how real time data on vehicle location and passenger loads for vehicles along a transit route can be used to reduce schedule deviations. Turnquist suggested that the model could be used as the basis for a control strategy that uses multiple control points and which regulates vehicle speeds to restore vehicles to schedule at the succeeding stops, rather than simply reacting to deviations from schedule by holding vehicles. The strategy was designed to take advantage of the broad capabilities for real time data acquisition represented by automatic vehicle location (AVL) and automatic passenger counting (APC) systems. The work was applicable to both rail and bus transit operations.

Furth [14] investigated optimal headway adjustment strategies to recover from delays, taking into account both ride time and waiting time at stations. Contrary to Turnquist's approach, Furth considered using holding strategies at stations only. He compared benefits provided by spreading the recovery over a large number of trains (minimizing platform waiting time) and by immediate recovery from the delay (optimizing ride time). Depending on the location and the duration of the initial delay, recovery spreading was shown to yield larger benefits. Furth's approach was also interesting as he limited the control scheme to the trains located behind the blockage. He found that benefits were not very large but merited being included in automatic train control systems. O'Dell [26, 27] later showed that much larger benefits come from controlling the trains ahead of the blockage (see below).

More recent research (Eberlein [12], O'Dell [26, 27] and Shen [30]) has focused on the evaluation of control strategies under different disruption duration scenarios and using mathematical programming formulations of the problem. Holding, expressing and short-turning of trains were considered in these rail transit studies.

Eberlein [11, 12] first considered the problems of deadheading, expressing and holding with vehicle locations as real-time input parameters. The objective was to minimize the total passenger waiting time in a single-loop system with no costs associated with the control actions. Delay incurred due to holding by passengers on-board the held train was omitted. In her more general model, running times between stations and passenger flow rates were station specific but deterministic. Moreover, as the model focused on routine control and not on significant disruptions, train capacity and short-turning strategies were not considered. The resulting mathematical formulation had a non-linear cost function and non-linear constraints. Heuristics were used to solve the problem as fast methods to solve non-linear mathematical programs were not yet available. Yet, Eberlein's formulations provided a solid foundation for further applications of operations research methods in this area.

Based upon the models developed by Eberlein, O'Dell [26, 27] formulated train holding and short-turning as disruption recovery strategies for a transit line with branches, namely the MBTA Red Line. Train capacity was considered and O'Dell also introduced the symplifying assumption that the delay duration was known. In-vehicle delay time was also omitted. The resulting formulation was a linearly constrained mathematical program with non linear costs. The quadratic objective function was approximated by piece-wise linear functions so that the resulting linear mixed integer program (MIP) could be solved by the existing linear optimizers.

For the holding problem, O'Dell considered and formulated three different types of holding strategies: holding trains at any station, holding trains at the first station they reach after the disruption occurs, and holding each train at only one optimally chosen station. Holding at any station clearly yielded greater benefits as the two other options were constrained versions of the general holding problem. Also, since in-vehicle delay time was not considered by the model, the resulting optimal holding strategies overestimated the benefits of holding and considered holding trains at multiple stations, which can be more difficult for dispatchers to implement. Nonetheless, she found that most of benefits of holding could be achieved by controlling a small set of trains ahead of the blockage, at a few stations. The short-turning problem was also formulated, assuming the final train order after the short-turn is given, and extending this formulation to an undetermined train order was also discussed.

The formulations were tested on two disruption scenarios with two different blockage durations. This yielded reductions in total passenger waiting-time of 15-40% in the case of the holding strategy, with no consideration of the increased on-board time for held passengers. The implementation of the short-turning model led to the conclusion that greater benefits are achieved by short-turning when the blockage duration is longer, as the time spent on short-turning trains is a relatively less significant factor.

O'Dell implemented her model formulations on a Sun SPARC 20 workstation with CPlex 3.0. By reducing the set of trains deemed impacted by the optimal control strategy, the problem size was reduced so that solution times under 30 seconds were achieved. Although these values seemed compatible with real-time implementation, the presented formulations were still too slow to solve if the model were to be extended to a stochastic formulation of the problem.

This model served later as a basis for Shen [30] in designing a more general model that considered holding, short-turning and expressing. Shen's model also assumed the disruption duration to be known and used a linear approximation of the cost function however the objective function combined the total passenger in-platform time *and* the on-board delay time due to holding.

Shen showed that, when accounting for the cost of holding, the benefits of holding might be less than suggested by O'Dell. Decreases less than 20% in weighted waiting time were observed in the only disruption scenario used to test the model formulation. Nonetheless, Shen also showed that short-turning combined with holding could achieve tremendous savings in waiting time: up to 57% decrease in weighted waiting timed was observed.

Two other major findings from the model implementation were also emphasized. First, expressing was shown to yield marginal additional benefits compared to the other two control strategies. Second, it was concluded from the optimal solutions that benefits of holding primarily stem from holding trains ahead of the blockage, which was consistent with O'Dell's findings.

In addition, Shen investigated the consequences of mis-estimating the disruption duration. By comparing the optimal solutions obtained with the hypothesized and the true durations, Shen showed that the effectiveness of short-turning is quite sensitive to the estimate of the disruption duration. In contrast, holding and expressing were robust strategies.

Furthermore, even after improving the solution search procedure with CPLEX 4.0, the solution times needed to solve the MIP problem formulation were not compatible with use in a stochastic formulation of the problem: it took more than 60 seconds to solve the combined holding, expressing and short-turning problems for a 20-minute disruption scenario. The large solution times were the consequences of numerous binary variables used to model each of the control strategies. The obtained range was nevertheless deemed acceptable for real-time use, assuming a deterministic delay duration.

1.2.2 Shortcomings and Limitations

While the abovementioned research has provided valuable insights and comparisons of the respective benefits of each control strategy under certain conditions, none of them have been demonstrated to be reliably implementable within transit agencies. All the prior work still has serious shortcomings and impediments to practical use:

- First, all the studies presented above used a deterministic representation of train control parameters such as running times, dwell-times and delay duration. While the consequences of using deterministic dwell-times and running times for a certain time period might be acceptable, the assumption of a predetermined delay duration is almost certainly too strong. The sensitivity analysis by Shen [30] showed that while the benefits of certain control strategies can be significant, they may be sensitive to the disruption duration variability. Given that dispatchers usually cannot estimate this duration accurately, these methods might lead to ineffective control strategies.

- Second, the solution times [26, 30] may still be too large for real-time use in transit operations, at least in the short term. It is suspected that the model formulations are overly complex to yield reliable fast solution times with the current state-of-art in optimization solvers. It is almost certain that, in the short run, these formulations could not form the core mathematical representation for the stochastic version of the train holding problem. Moreover, all those models seek optimality: it is unclear that this goal is worth the extra solution time given the simplifications made at the modeling stage. For real world applications, strict optimality is not a necessary goal.

- Third, some important practical issues have been overlooked for the sake of model tractability. For example, constraints due to terminal operations practices were not taken into account. Nonetheless, some important realistic assumptions were made in developing the optimization framework. For instance train capacity and short-turning locations were incorporated in the optimization models as well as line branching.

- Fourth, a very limited number of instances of the problem have been tested using real-time information in the optimization formulation. While findings were in general consistent among the different applications, the small number of such applications has yielded only general conclusions on the efficiency and the relative performance of train control strategies. In particular, the sensitivity of the optimal solutions to the relative weights of in-vehicle delay time and in-platform waiting time was not investigated.

None of the research using optimization techniques has lead to methodologies or tools which are readily usable by the transit industry. Actually, only studies based on simple heuristics or incremental recovery strategies have been of some success in that respect (Song[31]).

1.3 Existing Real-Time Control Strategies

Three types of control strategies are available to transit agencies to respond to disruptions that do not require substitute bus service, but introduce changes to the operating plans. They are holding, short-turning and expressing.

1.3.1 Holding

The holding strategy consists of delaying trains' departures at stations. It is the easiest strategy to implement and therefore the most frequently used within transit agencies. Disruptions generally result in a gap in front of the train immediately affected by the blockage. The objective of holding is to even out the sequence of headways preceding the blocked train. This goal is mainly driven by two rationales.

First, the original rationale for transit practitioners to even out headways stems from the ease of train and crew scheduling that a regular schedule entails. Indeed, rail transit operations include not only train and service scheduling but also crew assignment. Each of these problems is a rather complicated task on its own, so that optimally solving these problems together has been long considered impossible. Thus, operating a regular schedule was likely to be the best option for transit operators to conduct operations efficiently while providing an adequate level of service to their passengers.

Second, it has been demonstrated that, under some simplifying assumptions, the average waiting time at a station is minimized when the variance of the vehicles' (train or bus) headway is minimized (Welding [29]). While this result is simple and consistent with transit operators' practices, it depends on two important assumptions:

i) For the considered observation period, the passenger arrival process has a constant rate at the observation station.

ii) Vehicles have (effectively) infinite capacity, i.e. *any capacity constraint* is not binding.

Clearly, assumption *i)* is difficult to argue since, at any station, the passenger arrival process is likely to vary over time and generally shows a high level of randomness (see Eberlein [11]). Assumption *ii)* is most questionable when train loads are close to the train capacity. This situation is most likely to arise when a disruption occurs, resulting in a long preceding headway for the blocked train and a consequently larger number of passengers waiting for this train down the line.

Although the goal of holding is simple, different holding strategies can differ greatly in their implementation complexity, depending on the extent of required control actions. Indeed, a holding strategy is defined by the choice of the trains to be held and the stations where the holding occurs. Thus, as described by O'Dell [26], holding strategies can be classified according to the restrictions in selecting:

1. The set of trains that might be held and,

2. The set of stations where each of these trains might be held.

Below we review each of these possible restrictions along with results of the effectiveness of each type of control strategy.

Choice of The Controlled Trains

We distinguish here two sets of trains:

- the set of trains located behind the blockage and,

- the set of trains located ahead of the blockage (both downline from the disruption and in the reverse direction).

O'Dell [26] showed from her model implementation that negligible savings in waiting time are achieved through active holding[1] of trains behind the blockage (less than 1%).

[1]Following O'Dell [26], we make the distinction between passive and active train holding. A train is said to be *passively* held at a station if it dwells beyond the necessary dwell-time due to some physical constraints or as a consequence of the operating plan. For instance, trains queuing just behind the blockage are said to be passively held during the blockage duration. In the same fashion, a train that waits for a ring-off to be dispatched from the terminal station is also said to be passively held. Train holding for reasons other than the aforementioned ones is said to be *active*.

This result is intuitive when we consider in more detail the dynamics created by a disruption. We refer to Figure 1-1 for illustration. As the delay duration increases, trains behind the blockage (referred to as -1 and -2 in the figure) queue up behind the blockage. Meanwhile, a gap is created between train 0 and train 1 if no active control is exerted. These two distinct dynamics result in an increasing number of passengers affected by the growing delay, at stations both behind and ahead of the blockage. Behind the blockage, passengers who board trains queuing up behind the blockage are passively held and see their travel time increase. Ahead of the blockage (downline from the disruption and in the reverse direction), passengers arriving at stations already served by train 1 are affected by the long headway and accumulate at those stations.

Thus, train capacity becomes an issue for trains behind the blockage as they must serve the growing number of passengers arriving at stations behind and ahead of the disruption. It follows that actively holding these trains would only increase the waiting time and the number of affected passengers, with no benefits derived. Thus, trains located behind the blockage are not considered for active holding.

Nevertheless, one must evaluate the waiting time for these trains in assessing a holding strategy when delays grow and train capacity becomes a concern. For example, trains which are held ahead of the blockage not only change the headway sequence for trains *in front* of the blockage, but also help reduce the gap in front of the blocked train (train 0). This leads to fewer passengers waiting for train 0 (and following trains) and an alleviation of on-board congestion for these trains. Hence, the beneficial impacts of holding trains ahead of the blockage on trains behind the blockage should be captured in the assessment of any holding strategy.

Furthermore, the number of trains behind the blockage needed to clear all the passengers left behind (once the blockage is removed) not only depends on the train controls exerted ahead of the blockage, but also on the disruption duration and location. For instance, in Fig. 1-1, there are few stations and trains behind the blockage in the disruption direction. As the delay grows, trains in the reverse direction (e.g.,

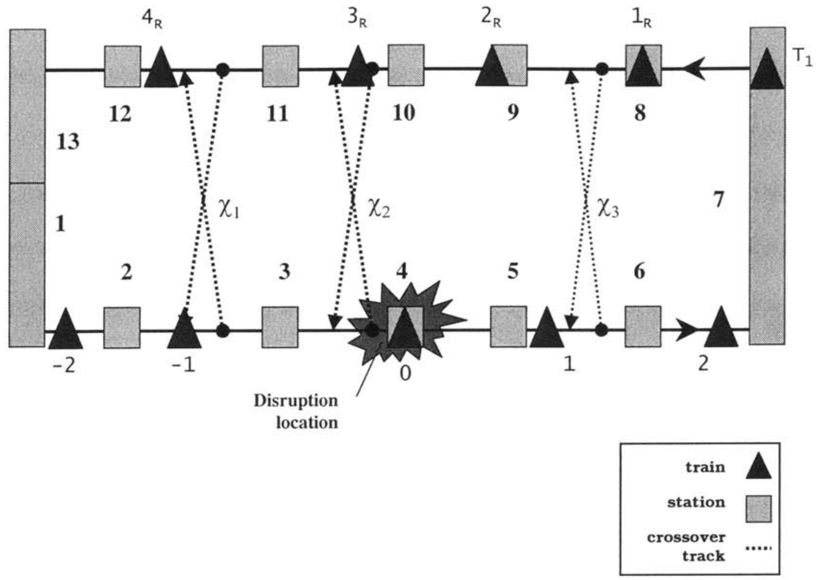

Figure 1-1: A disruption on a simple loop line

3_R and 4_R) would become part of the queue forming behind the blockage. Also, once the blockage is removed, they may need to reach congested stations as fast as possible in order to alleviate platform crowding.

One needs to be careful in drawing the line between the trains considered "behind the blockage" (not considered for holding) and those ahead of the blockage (considered for holding). The model presented later in this thesis considers *trains behind the blockage* to include only those trains needed (in the disruption direction and the reverse direction) to clear all the passengers left behind when no control action is taken[2].

Furthermore, we note that holding a full train at a station yields no benefits. This is because the immediate benefits of a train's hold consist of a decrease in total

[2]Trains in the reverse direction that are considered as "behind trains" are not actively held and are dispatched at the minimum safe headway after they reach the beginning terminal in the disruption direction. They are affected by a holding strategy only through the number of passengers left behind at stations after the terminal. Thus, these trains can be modeled as trains sitting at the terminal with a dispatching headway equal to the minimum safe headway.

passenger waiting time for the following train, as the following departure headway is reduced. In the case of a train loaded to capacity, arriving passengers cannot board the full train and must await the next train, so that the total passenger waiting time for the following train is not affected by the additional hold. In addition, these passengers might be frustrated by the sight of a full train they cannot board. Compounding this is the additional waiting time for departure incurred by passengers who boarded this train before it reached capacity. Finally, through-passengers[3] on-board the held train would be affected by the extra riding time due to the holding action. In consequence, once capacity is reached, trains should not be held any longer.

Choice of The Holding Stations

Differentiating holding strategies through selection of the holding station leads to three different types of strategies. The first strategy is to hold trains at the first station they arrive at after the disruption starts ("Hold First"). The second strategy is to hold each train at only one optimally chosen station ("Hold Once"). The third strategy allows trains to be held sequentially at more than one station on the line.

Intuitively and as confirmed by model experiments in [26], the fewer the constraints placed on the choice of the holding stations, the more effective the holding strategy will be. Specifically, the "Hold All" strategy is more effective than the "Hold Once" strategy, which is in turn more effective than the "Hold First" strategy. However, O'Dell showed that the benefits of "Hold All" were only slightly greater than for the two other strategies (the reduction in waiting time was less than 1%).

Nevertheless, it is expected that the difference between these strategies' respective effectiveness will be more pronounced when on-board delay time is included in the objective function. To illustrate this, one can consider train 1_R in Fig 1-1, arriving at station 8 where many passengers board to travel to the following station 9. In this case, holding at station 9 might turn out to be more beneficial than at station 8. This would happen if delaying the downstream benefits of the hold has a cost that is less

[3]Through-passengers of a train at a station are defined as passengers who are on-board the train both when it enters and leaves the station.

than holding passengers at station 8.

Conclusion

In consequence, an effective holding strategy should only consider actively holding trains ahead of the blockage at multiple stations. Trains behind the blockage are passively held but they are clearly affected by the choice of the holding strategy, as active holding in front of the blockage impacts both passenger waiting time for these trains and their loads. Hence, trains behind the blockage, albeit not controlled, should be included in assessing the effectiveness of a holding strategy.

1.3.2 Short-Turning

Short-turning is another strategy that is often employed by transit agencies. It consists of turning a train to the reverse direction before it reaches the terminus. It is considered because simply holding trains might not bring sufficient gains in recovering from some disruptions. Especially for longer disruptions, spreading the delay over a limited number of trains ahead of the blockage will still yield long headways and can be of limited value. Similarly, when a blockage is located in a high-demand portion of the line, the situation can very quickly become critical in terms of train and platform congestion. In these situations, dispatchers could make beneficial use of the less crowded trains and additional capacity that may be available on trains serving other parts of the line.

For instance, we could short-turn trains from the reverse direction to "fill the gap" developing in front of the blockage. Alternatively, when there exists high levels of travel demand in the reverse direction, trains could also be short-turned from behind the blockage to supply service in the other direction. We review below with the help of Figure 1-1 the different types of service impacts due to a blockage and estimate the effectiveness of diverse short-turning strategies in response.

Analysis

We consider the simple disruption case shown in Fig. 1-1. We consider different demand scenarios in parts of the line and we investigate the potential for short-turning in each of these scenarios. We first consider the case of high demand for travel between stations downline from the disruption, which might lead to short-turning trains to a location ahead of the blockage. We then consider the case of high demand for service between stations 11 and 13, which might lead to short-turning trains behind the blockage to the reverse direction.

We first assume that many passengers are traveling from station 6 to station 7 (terminal). In this case, only train 1 services these passengers during the duration of the blockage. Holding train 1 at station 6 might yield benefits but they are limited by the train capacity constraint and the additional travel-time incurred. Therefore, a short-turn loop might be considered, consisting of stations 6, 7, 8 and crossover χ_3. Short-turning trains 1_R and T_1 might then be considered at crossover χ_3, but we would then remove the service provided by these trains in the reverse direction to the disruption. This service removal results in a headway increase following train 2_R and consequently larger waiting times at stations located after station 9, once serviced by train 2_R. *One of the tradeoffs in choosing an effective short-turning strategy is between the potential waiting time savings in the served areas versus service degradation in areas where service is reduced.*

We next consider that many passengers are traveling from station 11 to station 13. In this case, trains -1 and -2 could be short-turned from *behind the blockage*, using crossover χ_2. In this case, trains outside the short-turn loop (consisting of stations 1, 2, 3, 11, 12, 13 and crossover χ_2) can be held in order to reduce the gap developing in front of the blocked train 0. These holds can result in a decrease in total waiting time at stations 5 through 10 once the blockage is removed but additional in-platform and in-vehicle time are incurred as in any holding strategy. Most of these savings may also be achieved by holding only a subset of the trains 1 through 3_R, so that train 3_R might "enter" the short-turn loop. This would provide additional service for passengers traveling from stations 11 and 12 to station 13. The benefits of these short-turns have associated costs as, within the short-turn loop, people "dumped" by short-turned trains -1 and -2 at station 3 create additional demand at this station for train 4_R. Nevertheless, if there are few passengers on-board these trains, queuing these trains behind the blockage instead of using this additional capacity might be sup-optimal. Hence, the decision here is clearly based on *the tradeoff between the potential waiting time savings at stations outside the short-turn loop and the costs of the short-turning actions.*

We now suppose that crossover χ_2 is not available, and trains -1 and -2 are short-turned using χ_1 and there is also high demand for travel between stations 2 and 3. In this case, short-turning train -1 might not be beneficial within the short-turn loop, as many passengers would be denied reaching station 3. Here, *one tradeoff in making an effective short-turning decision is between the waiting time savings in the fully served areas versus service degradation in skipped areas.*

From the above analysis, we can also conclude that the choice of a short-turn location depends on the availability and the location of the crossover tracks, but more importantly on the location of the disruption. During peak hours, most heavy rail transit lines serve the CBD that is located at the middle of the line and to which many riders travel. Since passengers are dumped by the short-turned trains, short turns *ideally* take place near the terminal, where trains' passenger loads are low. If the disruption is located beyond the CBD and close to the terminal, blocked trains are likely to have additional capacity that can be redirected to the reverse (peak) direction. In this case, trains that are short-turned would use a crossover track that is located behind the blockage, if available. If the blockage is located before the CBD and close to the terminal, then an appropriate strategy would short-turn trains at a crossover located between the blockage and the CBD.

Another important point is that short-turning a train requires reverse track operations, for which running speeds are lower for safety reasons: on MBTA lines, normal running speeds are between 20 and 40 mph, compared to a maximum speed of 25 mph on reverse tracks. As crossover operations are time-consuming (six minutes is typical on MBTA lines), short-turning strategies are only appropriate when we have longer disturbances or unusual circumstances. Therefore, based on train location information, inter-station running times, cross-over locations and the duration of the delay, there are *limited sets of candidate trains for short-turning at any available cross-over.* Consequently, not only must the choice of a short-turn location consider the location of the disruption and the aforementioned tradeoffs (based on the characteristics of travel demand along the line), but it is also restricted by the availability, configuration and ease of use of crossovers.

Conclusion

The potential benefits of short-turning depend mainly on the location of the disruption relative to the CBD (which determines the type of short-turn to be considered), and also on the choice of the short-turned trains and the achieved sequence of trains in the direction where trains are short-turned to. These critical factors cannot be treated separately as they all affect the choice and the effectiveness of a short-turning strategy.

Also, for a given short-turn strategy, train holding (passive and active) must be considered both to enable trains to be short-turned and to maximize benefits. Hence, the holding problem can be seen as the core sub-problem in any short-turning strategy. The difficulty of the short-turning strategy then lies in choosing a small set of feasible and sound short-turning strategies to assess.

Chapter 5 will provide a more detail discussion of these critical factors of the short-turn decision making process. It will also present how a holding model might be used to determine the effectiveness of a short-turn decision.

1.3.3 Expressing

Expressing is a third control strategy that is often used by transit agencies in response to service disruptions. An expressed train skips selected stops during its trip and hence has a reduced running time and preceding headway beyond the express segment. Shen showed in [30] that the expressing strategy provides only marginal benefit beyond holding and short-turning. Rather, expressing can be seen as a post-processing decision once holding and short-turning decisions have been made.

Therefore, we will focus on holding and short-turning as the primary strategies to deal with disruptions and treat expressing as a post-processor option.

1.4 Thesis Contents

1.4.1 Research Approach

Based on the discussion above, we first model the holding problem since it is the key sub-problem of the overall train control problem. The goal at this stage is to obtain a model that not only includes the features in the Shen and O'Dell models as well as terminal operations, but also has reduced solution times. These two seemingly contradictory goals can be achieved through the use of a more effective formulation of the problem, based on findings in O'Dell [26] and Shen [30].

Second, the short-turning problem is discussed in more detail and guidelines will be developed to help dispatchers make sound and efficient short-turn decisions. We will show how any given short-turning strategy can be formulated as a corresponding holding problem. Hence, the goal here is to give guidelines to choose a small set of short-turn strategies and assess each of these with the use of the previously developed holding model. These guidelines, combined with the holding model, could be eventually included in an automatic decision support system to help dispatchers make efficient short-turn decisions in real time.

1.4.2 Thesis Plan

In Chapter 2, we analyze a simple holding strategy that considers holding trains ahead of the blockage at the first station each train arrives at after the blockage. The resulting model serves as a basis for the more complex train control problems addressed in the remainder of the thesis.

In Chapter 3, we develop a more realistic, general system model, which considers holding trains at multiple stations in the disruption direction, ahead of the blockage, as well as in the reverse direction. Terminal operations and terminal capacity issues are analyzed and incorporated into the train control model.

In Chapter 4, we test the model formulation on several disruption scenarios, analyze the control strategies output by the model and perform sensitivity analysis with respect to various problem parameters. Results from the test scenarios are also studied to derive useful holding recommendations for manual control as well as for potential use in a DSS. To decrease solution times, a simple two-step solution procedure is presented and tested.

In Chapter 5, the short-turn strategy is discussed and studied in more detail. The study will focus on providing guidance to dispatchers in order to make useful short-turn decisions. In particular, it is shown how, for a common type of short-turn, the holding model developed in Chapter 4 can be easily modified and used to assess the impacts and the efficiency of a given short-turn strategy.

Finally, Chapter 6 summarizes the findings and offers suggestions for future research. In particular, the assumption of a known delay duration is discussed and suggestions for developing models that relax this assumption in future research are presented.

Chapter 2

Formulation and Analysis of a Simple Control Strategy

We first analyze a simple train control strategy for a simplified subway system. The system to be considered consists of a set of trains operating on a non-branching loop line, on which a disruption occurs. We assume that layover times at the terminal station are sufficiently large to ensure that delays caused by both the disruption and any control strategy will not impact line operations in the opposite direction. Thus, holding trains in the reverse direction from the disruption is not considered in the problem tackled in this chapter.

Although such a subway system seems *a priori* unrealistic, its study will provide us a better understanding of control strategy mechanisms in a simple system. The control strategy studied in this chapter is the "Hold First" strategy, wherein controlled trains are held at the first station they arrive at after the disruption occurs. The model does not consider capacity constraints for trains behind the blockage and the impacts of passengers left behind by these trains.

As mentioned in section 1.2.1, prior work by Furth [14] first focused on holding trains behind the blockage. O'Dell [26, 27] later showed that significantly greater benefits stem from holding trains at stations ahead of the disruption location. Hence, our study will focus on controlling trains ahead of the blockage, which are not delayed by the disruption.

This chapter, as well as the next one, aims to formulate and solve the above holding problems as optimization programs. A careful formulation and analysis is necessary to derive suitable numerical methods for solving this problem efficiently. It would also give us insight into the seeming complexity of the train control problem: one interesting aspect for instance is the impact of a train being held on other trains' movements on the line. Numerical methods will be used to solve the optimization program and obtain implementable control suggestions.

2.1 The "Hold First" Strategy

As previously stated, the "Hold First" control strategy consists of holding trains ahead of the blockage, when we can hold a controlled train only at the first station arrived at after the disruption starts. Even though this strategy may not be optimal[1], such a holding scheme is not unrealistic in some cases. For instance, during off-peak periods, we may want to recover quickly from minor service disruptions by immediately holding a small set of trains ahead of the blockage.

The assumptions and the notation used for the "Hold First" problem formulation are provided below and illustrated in Figure 2-1.

2.1.1 Assumptions and Notation

Assumptions

We make the following assumptions for our "Hold First" problem.

- *The duration of the delay is a known fixed parameter.* As discussed before (Section 1.2.2), this is probably the most questionable assumption.

[1] e.g., holding at one particular station might be undesirable when many on-board passengers alight at the following station. In this case, holding at the next station would cause less inconvenience.

- *Passenger arrival rates and alighting fractions are constant and station-specific.* This assumption is reasonable if the system's characteristics do not show much variation throughout the analysis period. This assumption also implies that passengers arrive randomly at stations, which is questionable especially in the case of transfer stations. At those stations, bulk arrivals are more likely due to transfers from other lines, but the characteristics of this process are complex and will not affect the expected waiting time. Therefore, a random arrival process will also be assumed at these stations to estimate expected passenger waiting times.

- *Train dwell-times are constant and station-specific.* Dwell-time is generally a function of platform boardings and alightings (see Lin [20, 21]). Therefore, since the number of passengers waiting for a given train depends on its preceding headway, holding the train ahead will affect its dwell-time. Nonetheless, dwell-time standard deviations at a station are in general under half a minute, which is a small fraction of the mean passenger waiting time. Thus, simplifying the dwell-time component may not be critical in developing holding strategies that seek to minimize passenger waiting time.

- *Train movements between stations are deterministic.* This assumption seems *a priori* questionable as train movements are stochastic in nature: they are function of many factors such as weather, track conditions and the train control system. Yet, we note again that for a limited analysis period, the standard deviation of train running time between two stations is likely small compared with other times of interest such as train headways and holding times, whose variations more significantly affect the level of service provided.

- *Full trains ahead of the blockage are not allowed to be held.* As discussed in the presentation of the various holding strategies in Section 1.3.1, trains ahead of the blockage that are loaded to capacity before departing are not held beyond the necessary dwell-time.

Notation

The following notation is used in the holding problem formulations:

d_0	$=$	delay duration
N	$=$	number of trains located between the blockage and the terminal in the disruption direction
M	$=$	number of stations on the line
M_0	$=$	station where the blockage occurs, or the station immediately ahead of the blockage if the disruption occurs between stations
C_i	$=$	capacity of train i
r_i	$=$	holding time of train i at its control station
H_i	$=$	uncontrolled departure headway of the i^{th} train ahead of the blocked train
H_s	$=$	minimum safe headway
$m(i)$	$=$	the current station of train $i = 1, \dots, N$ or the next arrived at –if traveling between two stations– after the disruption starts ($m(0) = M_0$)
G_i	$=$	station group of train i (see below).
$L_{i,m}$	$=$	load of train i upon arriving at station m
L_i^0	$=$	$L_{i,m(i)}$ for $i = 0, \dots, N$
λ_m	$=$	passenger arrival rate at station m
α_m	$=$	passenger alighting fraction at station m

We define for each controlled train i its *station group*, denoted G_i, as the set of stations between its control station $m(i)$ and the preceding train $i+1$'s control station $m(i+1)$[2] (non-inclusive). For train 0, $G_0 = \{M_0, M_0+1, \dots, m(1)-1\}$. For train N, the station group G_N includes its control station and all stations down the

[2]For the sake of simplicity, we adopt here the unusual notation: train $i+1$ is the train *ahead* of train i.

line until the station before the terminal: $G_N = \{m(N), m(N)+1, \dots, M-1\}$. For instance, the time-space diagram (Fig. 2-1) shows that $G_1 = \{M_0+3, M_0+4, M_0+5\}$ and $m(1) = \{M_0+3\}$.

M, λ_m, α_m and H_s are parameters of the system whose values are known; also, H_i, G_i, M_0 and L_i^0 can be determined in real-time[3], given basic train location information.

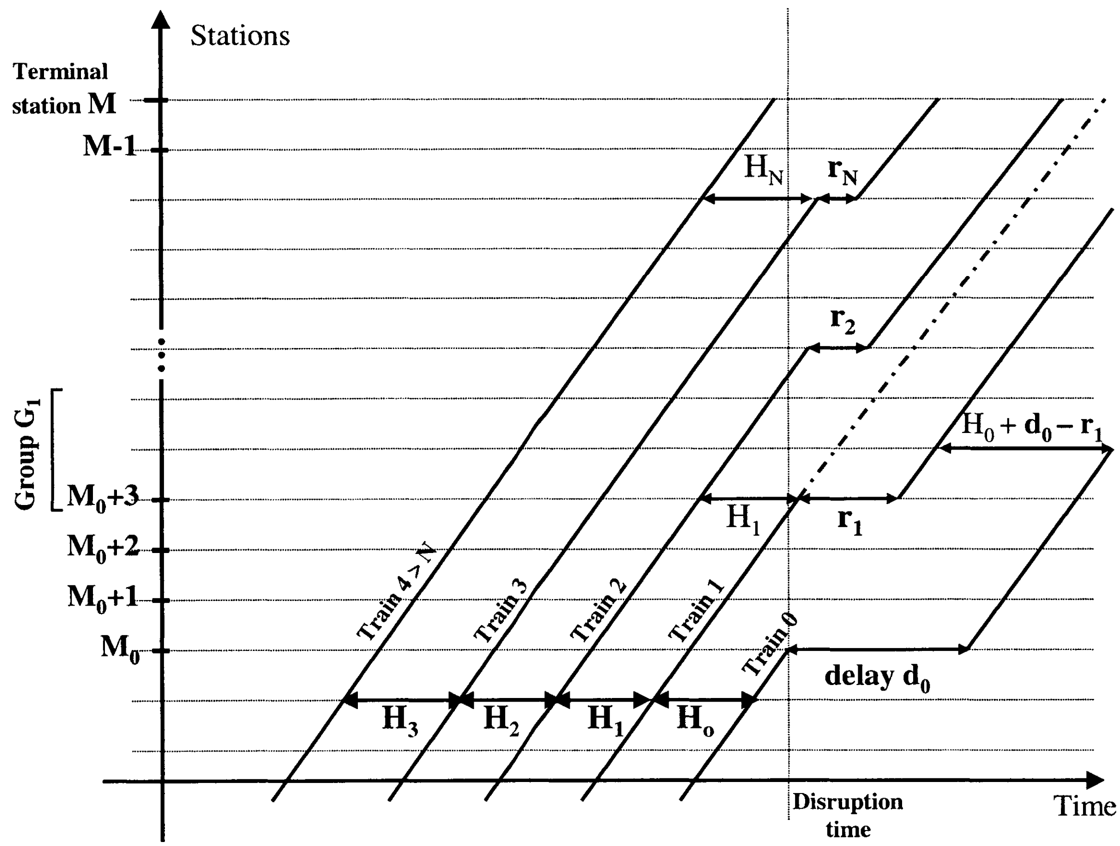

Figure 2-1: "Hold first" strategy with $N = 3$ trains held

2.1.2 Problem Formulation

Choice of a Cost Function

As stated by Barnett in [5],

[3]We can estimate L_i^0 from knowledge of train i 's departure time (and hence headways) at previous stations and the passenger arrival rate at those stations.

> [s]ince the primary effect of randomness is to create irregular intervals, it is appropriate that the waiting-time distribution for passengers be a major focus of efforts to improve service.[... Thus], the [transit] company chooses the waiting-time distribution that minimizes a group cost function [...].

The discussion provided in Barnett [5, pages 119–122] serves as a basis for the formulation of our problem and above all for the selection of an appropriate objective for our train control strategy.

In [5], Barnett assumes that the time-related "cost" the passenger associates with his trip is a function of three variables $F(a, u, v) = Q(a) + R(u) + S(v)$, where a represents the amount of time the rider has allotted for his trip ; u his waiting time for the vehicle, and v the difference between his actual arrival time and the time he wanted to arrive. The cost component $Q(a)$ is based on the amount of time the rider considers necessary for his trip. The $R(u)$ cost component relates solely to the inconvenience of waiting. $S(v)$ measures the rider's "disutility" of arriving at his destination v units late or, if v is negative, $|v|$ units early. Through variations of the "holding pattern", Barnett considers what control can be exerted on the waiting-time distribution to produce the lowest average cost per rider.

Although realistic and analytically tractable, this form of the objective function is not suitable for a model formulation to be used in a DSS. Indeed, as acknowledged in [5], the use of $Q(a)$ and $S(v)$ ignores

> [...] the fact that waiting has different "disutility" for different passengers–who face varying consequences of reaching destinations early or late–[... as] waiting costs are assessed identically for all users.

The lack of information to overcome this difficulty is tantalizing and leads us to adopt a cost function that is the least subjective possible and can be easily calculated from available real-time data (which would be $R(u)$ in [5]).

Hence, we choose our cost function as the *total* passenger time, that is waiting and on-board time. The choice of a *total* value is mainly motivated by mathematical tractability concerns. Indeed, an intuitive alternative cost function would be a weighted sum of *average* waiting times at all stations of the form $\sum_{\substack{\text{train } i \\ \text{station } m}} \lambda_{i,m}\bar{W}_{i,m}$, where $\bar{W}_{i,m}$ is the average waiting time for train i at station m and $\lambda_{i,m}$ is defined as the ratio of number of passengers waiting at station m for train i to the total number of passengers waiting in the system. Clearly, the value of those weights is dependent on the adopted control strategy, inasmuch as they include passengers arriving at the system during a train's hold. Specifically, additional passengers arriving downstream of the holding station during the hold must be considered in the cost function. Since those arrivals appear in both the numerator and the denominator of the ratios, the cost function can turn out to be quite difficult to compute and lack some necessary properties for optimization (e.g. convexity).

In the same fashion, capturing the *average* negative effects of holding on on-board passengers is also complex, since these effects depend on the proposed solution to the holding problem. Therefore, minimizing total passenger time seems a more tractable and reasonable goal to target for the train control problem.

Formulation

In the scope of the "Hold First" strategy, we have already determined the number of trains to be held and their control stations. Thus, the control problem is viewed as finding the optimal holding times $\{r_i\}_{i=1,\dots,N}$ so that the benefits of our holding strategy are maximized while the inconvenience it causes is minimized. This problem is equivalent to minimizing an objective function of the form $F(r) = F_1(r) + \mu F_2(r)$ with respect to $r = (r_1, r_2, \dots, r_N)$, where:

- F_1 represents the total platform waiting time for both the held trains and the delayed train

- F_2 accounts for the total extra riding time due to holding

- μ is a positive coefficient that weighs the negative effects of extra riding time against the total waiting time

We can derive the expression for those functions by inspection from the time–space diagram (Fig. 2-1) based on the headways that are modified by holding trains. We obtain the following equations:

$$F_1(r) = \sum_{i=}^{N} \sum_{m \in G_i} \frac{\lambda_m}{2} \left\{ (H_0 + d_0 - r_1)^2 + \sum_{j=1}^{i-1} \left(H_j + r_j - r_{j+1}\right)^2 + (H_i + r_i)^2 \right\}$$

$$F_2(r) = \sum_{i=1}^{N} L_i^0 (1 - \alpha_{m(i)}) r_i$$

In the first equation, the first term corresponds to the delayed train 0's headway at a given station m. The second and third terms correspond to headways for the preceding trains up to the train that is held at this station. We also consider that passengers who board a train during a station hold experience extra waiting time (for the train to depart), rather than extra ride time. Thus, in the second equation, we only account for the held train's passengers who remain on board at the control station, which is represented by the expression $L_i^0(1 - \alpha_{m(i)})$.

Hence, we must solve the following standard constrained optimization problem:

$$(HF) \qquad \min F(r) = \sum_{i=1}^{N} \sum_{m \in G_i} \frac{\lambda_m}{2} \left\{ (H_0 + d_0 - r_1)^2 + (H_i + r_i)^2 + \sum_{j=1}^{i-1} \left(H_j + r_j - r_{j+1}\right)^2 \right\} + \mu \times \sum_{i=1}^{N} L_i^0 (1 - \alpha_{m(i)}) r_i \quad (2.1)$$

subject to:

Load constraints

$$L_{0,m+1} = L_{0,m}(1 - \alpha_m) + \lambda_m \left(H_0 + d_0\right), \qquad \forall m : M_0 \le m < m(1) \qquad (2.2a)$$

$$L_{0,m+1} = L_{0,m}(1-\alpha_m) + \lambda_m (H_0 + d_0 - r_1), \qquad \forall m : m(1) \le m \le M-1 \tag{2.2b}$$

$$L_{i,m+1} = L_{i,m}(1-\alpha_m) + \lambda_m (H_i + r_i), \qquad \forall m : m(i) \le m < m(i+1)\,, \; \forall i = 1, \ldots, N-1 \tag{2.2c}$$

$$L_{i,m+1} = L_{i,m}(1-\alpha_m) + \lambda_m (H_i + r_i - r_{i+1}), \qquad \forall m : m(i+1) \le m \le M-1\,, \; \forall i = 1, \ldots, N-1 \tag{2.2d}$$

$$L_{N,m+1} = L_{N,m}(1-\alpha_m) + \lambda_m (H_N + r_N), \qquad \forall m : m(N) \le m \le M-1 \tag{2.2e}$$

Capacity constraints

$$L_{i,m+1} \le C_i, \qquad \forall m : m(i) \le m \le M-1, \; \forall i = 0, \ldots, N \tag{2.3}$$

Minimum safe headway constraints

$$H_0 + d_0 - r_1 \ge H_s \tag{2.4a}$$

$$H_i + r_i - r_{i+1} \ge H_s, \qquad \forall i = 1, \ldots, N-1 \tag{2.4b}$$

$$r_i \ge 0, \qquad \forall i = 1, \ldots, N \tag{2.4c}$$

The first group of equality constraints calculate train loads at stations. The first equation of this group calculates entering loads at stations for train 0. The second equation has the same calculations for the trains ahead of the blockage at the stations contained in the train station group while the third equation considers loads at stations further down the line. The second set of inequalities constrain those loads to be bounded by the train capacity. Indeed, in the scope of this strategy, which is applicable to disruptions of limited duration, we are likely to hold trains ahead of the blockage so that no capacity constraint is violated. Furthermore, the capacity constraint applies only to trains 1 through N as the disabled train 0 is likely to be overloaded. The third set of inequality constraints simply state that train headways must be greater than the minimum safe headway and that holding times must be positive.

We recognize this as a constrained quadratic program with linear constraints. To write the objective function in a more standard form, we first note that

$$
\begin{aligned}
&\sum_{i=1}^{N} \sum_{m \in G_i} \lambda_m \sum_{j=1}^{i-1} \left(H_j + r_j - r_{j+1}\right)^2 \\
&= \sum_{i=1}^{N} \sum_{j=1}^{i-1} \left(H_j + r_j - r_{j+1}\right)^2 \sum_{m \in G_i} \lambda_m \\
&= \sum_{i=1}^{N} \sum_{j=1}^{i-1} \left(H_j + r_j - r_{j+1}\right)^2 \lambda(i), \qquad \text{where } \lambda(i) = \sum_{m \in G_i} \lambda_m \\
&= \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left(H_j + r_j - r_{j+1}\right)^2 \lambda(i) \\
&= \sum_{j=1}^{N-1} (H_j + r_j - r_{j+1})^2 \sum_{i=j+1}^{N} \lambda(i) \\
&= \sum_{j=1}^{N-1} \Lambda(j+1) \left(H_j + r_j - r_{j+1}\right)^2, \qquad \text{where } \Lambda(j) = \sum_{i=j}^{N} \lambda(i)
\end{aligned}
$$

Thus,

$$
\begin{aligned}
F(r) = \frac{1}{2}\Lambda(1)\left(H_0 + d_0 - r_1\right)^2 + \frac{1}{2}\sum_{i=1}^{N-1} \Lambda(i+1)\left(H_i + r_i - r_{i+1}\right)^2 \\
+ \frac{1}{2}\sum_{i=1}^{N} \lambda(i)\left(H_i + r_i\right)^2 + \mu \times \sum_{i=1}^{N} L_i^0 \left(1 - \alpha_{m(i)}\right) r_i
\end{aligned}
\tag{2.5}
$$

and in standardized matrix form,

$$F(r) = \frac{1}{2} r' Q r + c' r + f \tag{2.6}$$

where

$$Q = \begin{pmatrix} 2\Lambda(1) & -\Lambda(2) & 0 & \dots & 0 \\ -\Lambda(2) & 2\Lambda(2) & -\Lambda(3) & & \vdots \\ 0 & -\Lambda(3) & \ddots & \ddots & 0 \\ \vdots & & \ddots & 2\Lambda(N-1) & -\Lambda(N) \\ 0 & \dots & 0 & -\Lambda(N) & 2\Lambda(N) \end{pmatrix} \in \mathbb{R}^{N \times N}, \tag{2.7a}$$

$$c = \begin{bmatrix} \Lambda(1)\big(H_1 - H_0 - d_0\big) \\ \Lambda(2)\big(H_2 - H_1\big) \\ \vdots \\ \Lambda(i)\big(H_i - H_{i-1}\big) \\ \vdots \\ \Lambda(N)\big(H_N - H_{N-1}\big) \end{bmatrix} + \mu \begin{bmatrix} L_1^0\big(1 - \alpha_{m(1)}\big) \\ L_2^0\big(1 - \alpha_{m(2)}\big) \\ \vdots \\ L_i^0\big(1 - \alpha_{m(i)}\big) \\ \vdots \\ L_N^0\big(1 - \alpha_{m(N)}\big) \end{bmatrix} \in \mathbb{R}^N \tag{2.7b}$$

and f is a constant term with respect to the decision variable r.

2.2 Model Analysis

2.2.1 Model Formulation

The structure of the quadratic program obtained in our simple "Hold First" strategy is appealing in several respects. First, the objective function has been formulated in a simple form in equations (2.5) through (2.7b), allowing for easy implementation. Also, the expression shown in equations (2.7a) and (2.7b) provides us with an interesting interpretation of the coefficients in our quadratic cost function.

The coefficients of the linear terms (with respect to the holding times) are shown in equation (2.7b). The coefficient of the holding variable r_i in the first term of the right-hand side is $\Lambda(i)\big(H_i - H_{i-1}\big)$. Given that $\Lambda(i)$ represents the cumulative arrival rate from station $m(i)$ through station $M-1$, we can interpret the coefficient $\Lambda(i)\big(H_i - H_{i-1}\big)$ as the difference in the number of boardings between trains i and $i-1$ from station $m(i)$ through station $M-1$. If this term is positive, this means that there will be more passengers waiting for train i than for train $i-1$, thus holding train i (i.e, $r_i > 0$) is likely to increase the value of our cost function. Otherwise, holding train i will provide us benefits by reducing waiting time for train $i-1$, which captures more ridership than its predecessor. The second part of the right-hand side in equation (2.7b) is simply associated with the extra-riding time.

The coefficients of the non-linear terms in equation (2.7a) consist of diagonal and cross-term coefficients, associated with the variables $r_i r_i$ and $r_i r_{i+1}$ respectively. The positive sign of the diagonal coefficient $\Lambda(i)$ is obviously related to the increased waiting time for train i if held at station $m(i)$. More interesting is the negative sign of the non-diagonal coefficients $\Lambda(i+1)$. This shows that the interaction between train i and $i+1$'s holds contributes to a decrease in waiting time for train i beyond station $m(i+1)$.

2.2.2 Solution Procedure

Provided that the matrix Q is definite positive[4], the objective function (2.6) is convex and thus has a minimum value. This is a standard optimization problem that can be solved in real-time for any disruption since the variables are continuous and are very small in number.

A disruption scenario with parameters $N = 5$ and $M = 13$ was solved using

[4]The proof of this conjecture is beyond the scope of this thesis.

MATLAB. The solution time was negligible. Results of this implementation are not presented in this thesis but this disruption scenario is solved using a more general model presented in Chapter 3. Therefore, results from this model implementation are presented and analyzed in Chapter 4.

2.3 Conclusion

This analysis has led us to the formulation of a simple model that is, under the assumptions made about the system, easy to implement and use relying on data readily available to the transit operator. Given the existence of AVL and adequate data storage and transfer technologies, this simple train control strategy could be easily implemented and may prove useful.

Nevertheless, we acknowledge that this simple control strategy (hold trains *only* at the first station and up to the terminal) is of limited use. We thus extend this problem formulation to derive a more general train control model in the next chapter.

Chapter 3

The General Train Holding Problem

Although insightful and implementable in some cases, the "Hold First" model is overly restrictive in regard to the choice of holding stations. In a general case, operators might hold trains at any station or at multiple stations on the line, which could yield more benefits. Indeed, we know from Section 1.3.1 that benefits can be greater when holding at multiple stations or at least beyond the first station reached[1]. Another drawback of this model is the non-consideration of trains located behind the blocked train and in particular capacity issues for these trains.

In the scope of our holding strategy study, we extend the model developed in Chapter 2 to address holding at multiple stations and include in our model holding trains at stations in the reverse direction to the disruption. Also, we model passengers left behind by trains located behind the blocked train as these trains might be overloaded.

[1]This general holding strategy is referred to as the "Hold All" strategy through the remainder of this thesis.

3.1 Problem Formulation

3.1.1 Multiple Holding

In order to model holding at multiple stations, we first define the following holding variables:

Definition 3.1. *For $i = 0, \dots, N$ and $m \geq m(i)$, we define $r_{i,m}$ as the holding time of train i at station m*[2]*.*

Definition 3.2 (Cumulative holding time). *For $i = 0, \dots, N$ and $m \geq m(i)$, we note $R_{i,m} = \sum_{p=m(i)}^{m} r_{i,p}$ the cumulative holding time of train i along the line, up to station m. By convention, $R_{i,m(i)-1} = 0$.*

Remark. We thus clearly have $r_{i,m} = R_{i,m} - R_{i,m-1}$ for all $m \geq m(i)$.

By inspecting the time-space diagram (Fig. 3-1), we can write an expression for the total waiting time for train 0 and the trains located beyond the blockage in the disruption direction:

$$\begin{aligned} F(R_{i,m}, L_{i,m}) = &\sum_{i=0}^{N} \sum_{m \in G_i} \frac{\lambda_m}{2} \left\{ \sum_{j=0}^{i-1} \big(H_j + R_{j,m} - R_{j+1,m}\big)^2 + (H_i + R_{i,m})^2 \right\} \\ &+ \mu \sum_{i=1}^{N} \sum_{m=m(i)}^{M-1} L_{i,m}\big(1 - \alpha_m\big)\big(R_{i,m} - R_{i,m-1}\big) \end{aligned}$$

We can reformulate this expression as a function of the cumulative holding times by first noting that

$$\begin{aligned} &\sum_{i=0}^{N} \sum_{m \in G_i} \lambda_m \sum_{j=0}^{i-1} \big(H_j + R_{j,m} - R_{j+1,m}\big)^2 \\ =& \sum_{j=0}^{N-1} \sum_{i=j+1}^{N} \sum_{m \in G_i} \lambda_m \big(H_j + R_{j,m} - R_{j+1,m}\big)^2 \\ =& \sum_{j=0}^{N-1} \sum_{m=m(j+1)}^{M-1} \lambda_m \big(H_j + R_{j,m} - R_{j+1,m}\big)^2 \end{aligned}$$

[2] Active and passive holdings are considered. Thus the holding time of delayed train 0 includes the delay duration d_0.

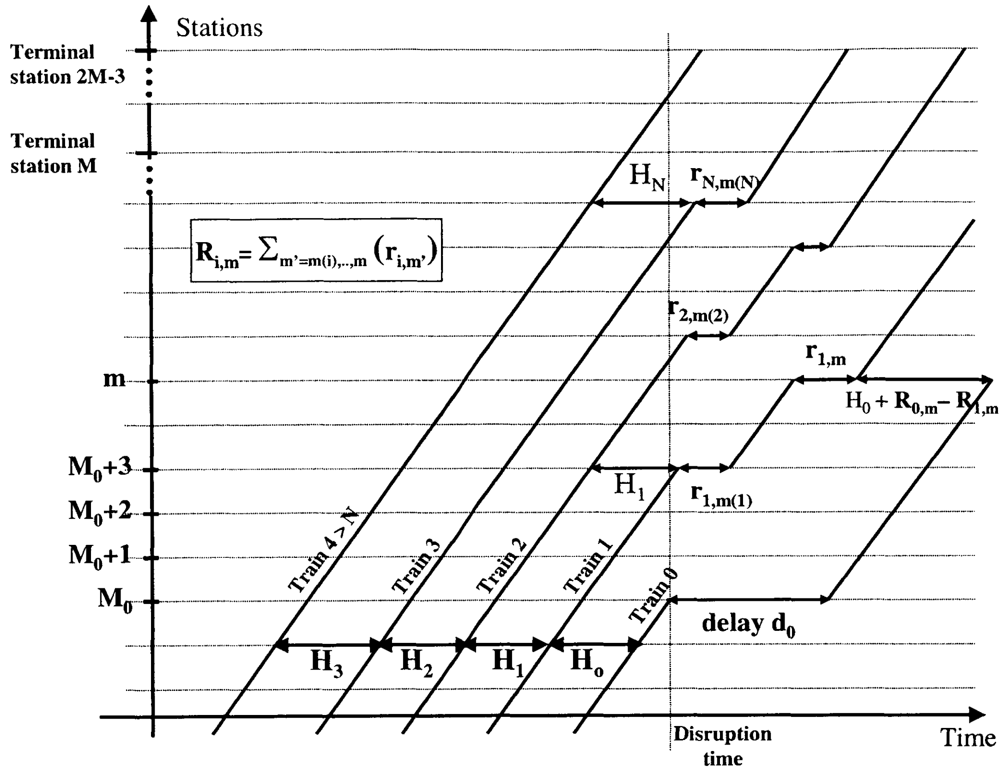

Figure 3-1: "Hold all" strategy with $N = 3$ trains held

Thus,

$$F(R_{i,m}, L_{i,m}) = \frac{1}{2}\sum_{i=0}^{N}\sum_{m=m(i)}^{m(i+1)-1} \lambda_m\big(H_i+R_{i,m}\big)^2+\frac{1}{2}\sum_{i=0}^{N-1}\sum_{m=m(i+1)}^{M-1} \lambda_m\big(H_i+R_{i,m}-R_{i+1,m}\big)^2$$
$$+\,\mu \times \sum_{i=1}^{N}\sum_{m=m(i)}^{M-1} L_{i,m}\big(1-\alpha_m\big)\big(R_{i,m}-R_{i,m-1}\big) \tag{3.1}$$

Clearly, we can easily extend this model to holding at stations in the reverse direction–as opposed to the disruption direction– for trains ahead of the blocked train[3]. Nevertheless, some modeling issues deserve careful consideration:

1. *Terminal capacity* should be appropriately modeled so that we allow two trains

[3]This is equivalent to "unfolding" the line.

to sit at the terminal station in a standard stub-end configuration. When both platforms are occupied and another train is about to arrive at this terminal as well, this train would then need to wait until a platform is cleared. In case the corresponding queuing location is not a station, we would then model it using a *virtual* station $M-1$ with no passenger arrivals ($\lambda_{M-1}=0$) or alightings ($\alpha_{M-1}=0$) associated with it[4].

The model shown hereafter considers both active and passive holding at this queuing location. That is, trains can be held at the queuing location rather than at station $M-2$. Such an action might seem irrational as queuing at station $M-2$ before the terminal could benefit passengers traveling from this station to the terminal as they can board the train delayed there. Also, passengers held at a station are usually less concerned by the holding action if it occurs at a station.

Nonetheless, there might be cases where queuing between stations would be beneficial. Indeed, a train queuing after station $M-2$ frees this station so that passengers from following trains can alight at this station without additional delay. If this benefit outweighs the benefits of train queuing at station $M-2$ and servicing more passengers traveling from station $M-2$ to the terminal, queuing between stations will be appropriate. The benefits of such situations are captured by the objective function defined above.

2. *Queuing behind the blockage* must also be considered as queued trains can incur different waiting times depending on the holding strategies ahead of the blockage.

3. *Train capacity* is to be modeled and the load calculation constraints consequently modified for trains behind the blockage. The extra waiting-time incurred by *passengers left behind* by fully loaded trains should be properly in-

[4]In this case, the total number of stations included in the model is $2M-3$ when we unfold the line. We have two modeled stations for every physical station (one platform in each direction), except for the terminal station and the queuing location which are each modeled by one station.

cluded in the objective function as well.

3.1.2 Notation

We introduce some new notation to describe the general holding model:

T_i		is a subscript used for trains currently at the terminal ahead of the disruption. Here, $i = 1, 2$ as we consider two platforms at the terminal
i_R		is a subscript used for the i^{th} train in the *reverse* direction
δ	=	a binary parameter that is equal to one if a *second* train is present at the terminal ahead of the disruption. We assume that there is always at least one train at this terminal.
N_B	=	number of trains located behind the first blocked train in both the disruption and the reverse directions, needed to clear the passengers left behind. Trains in the reverse direction are modeled as sitting at the terminal behind the disruption, with a dispatching headway equal to the minimum safe headway
N_R	=	number of trains operating in the reverse direction (and not considered as "trains behind"), not including any train sitting at any terminal
S_B	=	index set of trains "behind the blockage" ($S_B = \{-N_B, \ldots, -2, -1, 0\}$)
S_A	=	index set of trains ahead of the blockage in the disruption direction. ($S_A = \{1, 2, \ldots, N\}$)
S_T	=	index set of trains at the terminal ahead of the blockage. ($S_T = \{T_1\}$ or $S_T = \{T_1, T_2\}$)
S_R	=	index set of trains in the reverse direction ($S_R = \{1_R, 2_R, \ldots, N_R\}$)
S	=	index set of all trains in the system, i.e., $S = S_B \cup S_A \cup S_T \cup S_R$

$h_{i,m}$ = departure headway of train i at station m

$P_{i,m}$ = number of passengers left behind by train i at station m

Ω_i = remaining layover time of train i at the terminal following the disruption, if the train is not held before this station

Ψ = maximum dispatching time deviation from schedule at the terminal following the disruption

Ξ = minimum turnaround time at terminal stations

H_N^p is defined as the *projected headway* of train N. It is defined as the difference between the preceding train's departure time at station $M-1$ and the projected departure time of train N if this train is not held between stations $m(N)$ and $M-1$

The minimum turnaround time Ξ is the time required after arrival for the train to be ready to depart in the reverse direction. This typically includes the time for the train crew to switch ends of the train and routine checking.

Note that the holding variable $R_{i,m}$ associated with train i at the terminal station in our model includes the layover time, which is considered to be a holding time. Also, the definition of train N's *projected headway* at station $M-1$ allows us to consider any holding actions already exerted on train N between stations $m(N)$ and $M-1$. For trains $i \in S_B$ $(i < 0)$ behind the blocked train, only passive holding is considered, as discussed in Section 1.3.1. These passive holds only occur at the queuing location, at the terminal and at queuing stations behind the blockage.

3.1.3 Definition of the objective function

We clearly have four groups of trains to consider: *1)* trains behind the blockage including train 0 (referred to as group I), *2)* trains ahead of the blockage in the disruption direction (referred to as group II), *3)* one or two trains at the terminal following the disruption (group III) and *4)* trains in the reverse direction (group IV).

Thus, we can write the cost function as the sum of four different functions associated with each of these train groups, plus the waiting cost incurred by any passengers left behind:

$$\boxed{F(h_{i,m}, R_{i,m}, L_{i,m}, P_{i,m}) = \sum_{g=I}^{IV} F_g(h_{i,m}, R_{i,m}, L_{i,m}) + F_P(h_{i,m}, P_{i,m})} \tag{3.2}$$

with

$$F_I(h, R, L) = \sum_{i=-N_B}^{0} \sum_{m=m(i)}^{2M-3} \frac{\lambda_m}{2} h_{i,m}^2 + \mu \times \sum_{i=-N_B}^{0} \sum_{m=m(i)}^{2M-3} L_{i,m}(1-\alpha_m)(R_{i,m} - R_{i,m-1}) \tag{3.3}$$

where we note $h = \{h_{i,m}\}$, $R = \{R_{i,m}\}$ and $L = \{L_{i,m}\}$.

$$F_{II}(h, R, L) = \sum_{i=1}^{N} \sum_{m=m(i)}^{2M-3} \frac{\lambda_m}{2} h_{i,m}^2 + \mu \times \sum_{i=1}^{N} \sum_{m=m(i)}^{2M-3} L_{i,m}(1-\alpha_m)(R_{i,m} - R_{i,m-1}) \ , \tag{3.4}$$

$$F_{III}(h, R, L) = \sum_{i=1}^{1+\delta} \sum_{m=M}^{2M-3} \frac{\lambda_m}{2} h_{T_i,m}^2 + \mu \times \sum_{i=1}^{1+\delta} \sum_{i=M}^{2M-3} L_{T_i,m}(1-\alpha_m)(R_{T_i,m} - R_{T_i,m-1}) \ , \tag{3.5}$$

$$F_{IV}(h, R, L) = \sum_{i=1_R}^{N_R} \sum_{m=m(i)}^{2M-3} \frac{\lambda_m}{2} h_{i,m}^2 + \mu \times \sum_{i=1_R}^{N_R} \sum_{m=m(i)}^{2M-3} L_{i,m}(1-\alpha_m)(R_{i,m} - R_{i,m-1}) \tag{3.6}$$

$$\text{and } F_P(h,P) = \sum_{i=-N_B}^{0} \sum_{m=m(i)}^{2M-3} P_{i,m} h_{i-1,m} \tag{3.7}$$

Equation 3.5 shows how we can simply model a second train at the terminal ahead of the disruption with the use of the binary parameter δ.

3.1.4 Formulation of the constraints

Below we present the constraints to which the holding time, load and passenger-left-behind variables are subject. For the sake of readability, we assume that the train preceding train i is also denoted $i+1$ for trains $i \notin \{-N_B, \ldots, 0, \ldots N-1\}$. For instance, the train preceding train $i = N$ is train $i+1 \overset{def}{=} T_1$ and the train preceding train $i = T_2$ is train $i+1 \overset{def}{=} 1_R$.

Headway calculation constraints

$$h_{i,m} = H_i + R_{i,m},\ \forall m : m(i) \le m < m(i+1),\ \forall i \in S - \{N_R\} \tag{3.8a}$$

$$h_{i,m} = H_i + R_{i,m} - R_{i+1,m},$$
$$\forall m : m(i+1) \le m \le 2M-3,\ \forall i \in S - \{N_R\} \tag{3.8b}$$

$$h_{N_R,m} = H_{N_R} + R_{N_R,m},\ \forall m : m(N_R) \le m \le 2M-3 \tag{3.8c}$$

Load/capacity constraints for trains ahead of the blockage

$$L_{i,m+1} = (1-\alpha_m) L_{i,m} + \lambda_m\, h_{i,m},$$
$$\forall m : m(i) \le m \le 2M-3,\ \forall i \in S_A \cup S_T \cup S_R \tag{3.9a}$$

$$L_{i,m+1} \le C_i, \qquad \forall m : m(i) \le m \le 2M-3,\ \forall i \in S_A \cup S_T \cup S_R \tag{3.9b}$$

Load/capacity constraints for trains behind the blockage

$$L_{i,m+1} \leq P_{i+1,m} + (1-\alpha_m)L_{i,m} + \lambda_m h_{i,m},$$
$$\forall m : m(i) \leq m \leq 2M-3,\ \forall i \in S_B \tag{3.10a}$$

$$L_{i,m+1} \leq C_i, \qquad \forall m : m(i) \leq m \leq 2M-3,\ \forall i \in S_B \tag{3.10b}$$

$$L_{i,m+1} \geq P_{i+1,m} + (1-\alpha_m)L_{i,m} + \lambda_m h_{i,m} - K\nu_{i,m},$$
$$\forall m : m(i) \leq m \leq 2M-3,\ \forall i \in S_B \tag{3.10c}$$

$$L_{i,m+1} \geq C_i - K(1-\nu_{i,m}), \qquad \forall m : m(i) \leq m \leq 2M-3,\ \forall i \in S_B \tag{3.10d}$$

, where K is a large constant

Left-behind-passenger constraints

$$P_{i,m} = P_{i+1,m} + L_{i,m}(1-\alpha_m) + \lambda_m h_{i,m} - L_{i,m+1}\,,$$
$$\forall m : m(i) \leq m \leq 2M-3,\ \forall i \in S_B \tag{3.11a}$$

Minimum safe headways

$$H_i + R_{i,m} \geq H_s,\ \forall m : m(i) \leq m < m(i+1),\ \forall i \in S - \{N_R\} \tag{3.12a}$$

$$H_i + R_{i,m-1} - R_{i+1,m} \geq H_s,\ \forall m : m(i) \leq m < m(i+1),\ \forall i \in S - \{N_R\} \tag{3.12b}$$

Terminal capacity queuing constraints

$$R_{i+2,M} \leq H_{i+1} + H_i + R_{i,M-1},\ \forall i \leq N-2 \tag{3.13a}$$

$$R_{T_1,M} - R_{T_1,M-1} \leq H_N^P + H_{N-1} + R_{N-1,M-1} \tag{3.13b}$$

$$R_{T_2,M} - R_{T_2,M-1} \leq H_{T_1} + H_N^P + R_{N,M-1} + K(1-\delta) \tag{3.13c}$$

Queuing constraints and no-active holding for trains behind the blockage

$$R_{i,M_0+i-1} = 0, \quad \forall i \in S_B : M_0 + i - 1 \geq 2 \tag{3.14a}$$

$$R_{i,m} - R_{i,m-1} = 0, \quad \forall m \notin \{M-1, M\} \text{ and } m \geq 2,\ \forall i \in S_B \tag{3.14b}$$

Layover constraints at terminal

$$R_{i,M} \geq \Omega_i, \quad \forall i \in S_B \cup S_A \cup S_T \tag{3.15}$$

Turn-around constraints at terminal

$$R_{i,M} - R_{i,M-1} \geq \Xi, \quad \forall i \in S_B \cup S_A \cup S_T \tag{3.16}$$

Maximal deviation from schedule constraints

$$R_{i,M} - \Omega_i \leq \Psi, \quad \forall i \in S_B \cup S_A \cup S_T \tag{3.17}$$

Passenger left-behind pickup constraints (for trains behind)

$$\nu_{i,m} \geq \nu_{i-1,m}, \quad \forall i \in S_B, \forall m \leq 2M-3 \tag{3.18}$$

Disruption duration constraint

$$R_{0,M_0} \geq d_0 \tag{3.19}$$

Cumulative holding times are monotonically increasing

$$R_{i,m} - R_{i,m-1} \geq 0, \quad \forall i \in S \tag{3.20}$$

$$R_{i,m}, L_{i,m}, P_{i,m} \geq 0 \text{ and } \nu_{i,m} \in \{0,1\}$$

The use of the large constant K and the binary parameter δ in the constraints involving a second train at a terminal are straightforward. For a second train at a terminal, the large constant K enables the slack variable of the corresponding inequality to be very large and thus the constraint is effectively "dropped" if this second train is not present.

Equations (3.8) simply calculate the modified preceding headway of train i. Equations (3.10) are equivalent to the following set of load/capacity constraints for trains $i \in S_B$ $(i < 0)$ behind the blockage:

$$L_{i,m+1} = \min\left(L_{i,m}(1 - \alpha_m) + \lambda_m\, h_{i,m}, C_i\right),\ \forall m : m(i) \leq m \leq 2M - 3 \tag{3.21}$$

where the *min* function has been modeled through the use of the large constant K and the binary variables $\nu_{i,m}$ ($\nu_{i,m} = 1$ iff passengers are left behind by train i at station m, 0 otherwise).

Consequently, our holding problem is a 0-1 Mixed Integer Program (0-1 MIP) that might be hard to solve in a reasonable amount of time. In fact, this combinatorial problem is greatly simplified by constraint (3.18). This constraint states that, if a train $i < 0$ leaves no passengers behind, then the following trains do not deny boarding either. Given the description of the train capacity issue presented in Section 1.3.1, this simply means that trains behind the blockage gradually pick up passengers left behind at stations where the train capacity issue arises.

As in the case of the "Hold First" model, Equations (3.9) state that trains ahead of the blockage, and in the reverse direction, are not held after they reach capacity. Thus, no passengers are left behind by these trains.

On the contrary, trains behind the blockage are likely to experience overcrowding as passengers boarding these trains accumulate both ahead of, and behind, the blockage during the disruption, as discussed in Section 1.3.1. The load/capacity constraints for trains behind the blockage (Eq. (3.10)) state that these trains may be overloaded and thus deny passengers boarding.

The terminal capacity constraints (3.13) are derived directly from inspection of the headways in Figure 3-2. In this figure, distinction is made between trains i with $i \leq N - 2$ and trains $N - 1$ and N. We focus on the case $i \leq N - 2$ as the two other cases are easily derived from it. In this case, we can see from Figure 3-2 that the gap between a train i and its second predecessor at station $M-1$ is equal to $(H_{i+1}+R_{i+1,M-1}-R_{i+2,M-1})+(H_i+R_{i,M-1}-R_{i+1,M-1}) = H_i+H_{i+1}+R_{i,M-1}-R_{i+2,M-1}$. This value must be greater than train $i+2$'s holding time at the terminal $(R_{i+2,M}-R_{i+2,M-1})$ so that a platform is free for train i to enter, which yields (3.13a).

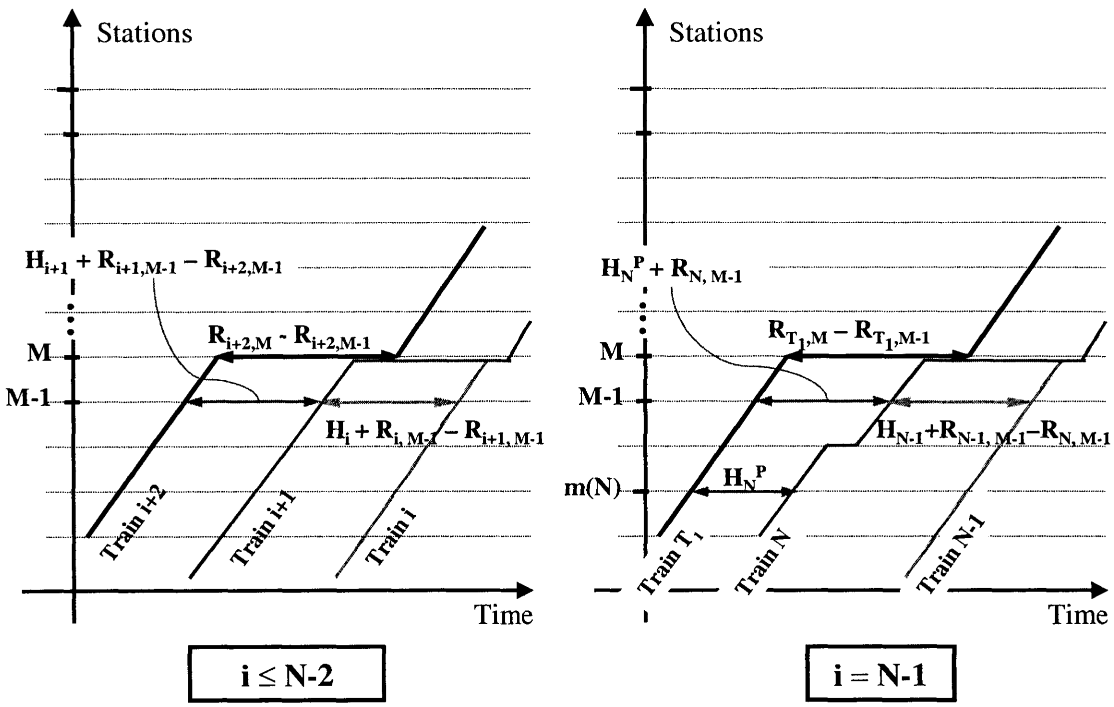

Figure 3-2: Train queuing before the terminal

Constraints (3.14) constrain trains behind the blocked train not to be held until they reach the closest station to the disruption where they can queue. In this case, the queueing time is included in the holding variable R_{i,M_0+i} for $i \in S_B$ as queuing or holding has the same effect on trains' headway (passive hold). In the same fashion, delay d_0 is incorporated into the cumulative holding time R_{0,M_0}.

The meaning of constraint (3.15) is clearer if we rewrite it in the form $R_{i,M} - R_{i,M-1} \geq \Omega_i - R_{i,M-1}$, $\forall i \leq N$. In the previous equation, the left-hand side represents the (passive) holding time of train i at terminal M, and the right-hand side represents train i's remaining layover time when it reaches station M. Thus, the inequality states that train i must be passively held at the terminal until it is dispatched according to the schedule.

Finally, equations (3.15), (3.16) and (3.17) ensure that operational constraints are respected at the terminal station.

3.2 Model Size and Boundary Effects

Clearly, the model used here is limited by the set of stations modeled and included in the evaluation of the objective function ($2M-3$). Including solely stations in the disruption direction and the reverse direction can be considered unsatisfactory in the case of very long disruptions and/or very short recovery times. For such cases, the number of trips needed to recover from the delay (those trips incurring holding) might be well beyond the one trip in the reverse direction implied by our model. Also, as the downstream effects of holding at station $2M-3$ are not evaluated in our objective function, unnecessary or unreasonable holds might be considered at this station.

One could attempt to correct this limitation by "unfolding" the line more than once and setting the boundary of our system to a station with index greater than $2M-3$. This is equivalent to considering the trains in the system as traveling on a very long one-way line (with every other segment of the line being identical). The limit of this line could depend on the delay duration. Nevertheless, this approach clearly expands the size of the model and increases the solution time as the delay duration increases. This is a major impediment to the real-time implementability of our model. In addition, difficulties arise from the longer duration of the observation period, resulting in variations of the system parameters (passenger arrival rates and alighting fractions).

Another approach to overcome the boundary effects would be assuming that, beyond station $2M-3$, long delays are recovered through the use of terminal layover times. One would then restrict the cumulative holding times in the other direction ($R_{i,2M-3} - R_{i,M}$) in order to make this strategy effective. The difficulty here is to appropriately evaluate the benefits derived from the use of terminal layover time beyond station $2M-3$, given the holding pattern exerted on the system before station $2M-3$. Variations of the problem parameters also make this approach less attractive.

3.3 A Two-Step Solution Procedure

We have here a 0-1 MIP formulation of the holding problem that can be solved by current commercial integer program solvers. These solvers usually tackle this type of problem by intelligently enumerating the feasible solutions to the problem. Nonetheless, we can improve this enumeration procedure by using *a priori* knowledge about the values of the binary variables. Constraint (3.18) already provided a way to prune many branches of the solution tree. Indeed, if for given i and m, the subtree below the branch $\nu_{i,m}=0$ is searched during the solution procedure, constraint (3.18) prunes all the branches $\nu_{j,m}=1$, $j<i$. We provide below another means to further prune the tree.

It is clear that, if a train at a station was not fully loaded when no control strategy was considered, then an effective control strategy will not "create" a capacity issue for this train at the same station.

This leads to the two-step solution procedure shown below (we denote our holding program $\mathcal{H}$):

Step 1. Constrain all trains' active holding at stations to be null and seek a feasible solution R^0, L^0, P^0, ν^0 to the corresponding linear system of constraints.

Step 2. Solve $\mathcal{H}(R, L, P, \nu)$ with the variables $\nu_{i,m}$ for $\nu^0_{i,m} \neq 0$, and constrain the other $\nu_{i,m}$ to be zero.

The underlying rationale of this approach is straightforward. We first locate in *Step 1* the locations where the train capacity issue arises in the absence of any holding. Given this information (the $\nu^0_{i,m}$'s) from this worst-case scenario, we attempt to come up with a better solution in *Step 2*. At this stage, we will not have train capacity issues at stations where trains were not fully loaded before. That is, if $\nu^0_{i,m} = 0$ after *Step 1*, then $\nu_{i,m}$ is set to zero. This means that, during the solution procedure, none of the branches $\nu_{i,m} = 1$ (for $\nu^0_{i,m} = 0$) in the solution tree are explored.

This procedure can dramatically reduce the number of free binary variables if few trains and/or few stations are affected by the train capacity issue without holding. This can occur when the delay is not long enough to lead to capacity issues at many stations. In this case, train capacity issues arise mainly at stations with large passenger arrival rates and/or small passenger alighting fractions. Similarly, the problem size is reduced if a small set of trains are overloaded at these stations. In both cases, the smaller number of train i/station m combinations affected by the capacity issue ($\nu^0_{i,m} = 1$) leads to the smaller numbers of binary variables present in *Step 2* and branches which must be searched during the solution procedure.

3.4 Conclusion

In this chapter, we developed a formulation of the train holding model, based on train headways and cumulative holding times. All the model variables are continuous, except for the binary variables associated with train capacities. The model considers actively holding trains ahead of the blockage (downline from the disruption and in the reverse direction) at stations up to the end of the reverse direction. The objective function is chosen as the weighted sum of the total passenger waiting time and in-vehicle time. The exact non-linear form of this function was used in solving the problem, as distinct from Shen [30] and O'Dell [26].

We also discussed the possible impacts of the limited set of stations where we evaluate the effects of a holding strategy. These impacts are shown through results of model application in Chapter 4. A simple two-step solution procedure was also proposed to solve this problem.

Chapter 4

Application of the General Holding Model

In this chapter, we present results from the general holding model ("Hold All") application. Two problem instances on the MBTA Red Line are treated, both involving a disruption on the line during the morning peak period, as even minor disruptions at this time of day can lead to serious consequences if effective control actions are not taken.

Results from both disruption cases are presented and analyzed in this chapter. Detailed implementation results from both disruption cases are presented in Appendices A and B. The analysis of the implementation results will focus on several points:

- The benefits achieved by the optimal holding strategy as well as its structure
- The impact of train capacity on the optimal holding strategy
- The sensitivity of the holding strategy to the cost associated with in-vehicle delay due to holding (i.e., to μ).
- The viability of the resulting holding strategies for use by dispatchers

Moreover, we will assess the efficiency of the model and the two-step solution procedure from a real-time implementation point of view. Execution times with and without the two-step solution procedure will be examined and compared with the ones obtained by O'Dell in [26].

4.1 Problem Setting

4.1.1 The MBTA Red Line

We applied the "Hold All" model to the different disruption scenarios on the Massachusetts Bay Transportation Authority (MBTA) Red Line, which also served as an instance of model application in Shen [30] and O'Dell[26].

The MBTA Red Line is a heavy rail system with two branches and a common trunk portion. The junction point of the branching structure is JFK Station (see Figure 4-1). Each train has a capacity of about 960 passengers and are typically dispatched onto the line from two terminal stations, Ashmont and Braintree, each being located at the end of a branch. Trains are scheduled to be dispatched every eight and six minutes from Ashmont and Braintree respectively. This results in a mean scheduled headway of three to four minutes on the trunk portion of the line. In addition, there are layovers of approximately six minutes at Alewife Station, which is located at the northern end of the trunk portion of the line.

Although our model is not generally applicable to transit systems with a branching structure such as the Red Line, we use the MBTA Red Line so we can compare our results with results presented in O'Dell [26], who solved an instance of disruption located on the trunk portion of the line (20-minute disruption at Harvard Northbound). To apply the model developed in Chapter 3, we modeled the line as a single loop line with two terminal stations (Alewife and JFK), in place of a branching structure. Trains dispatched from the two branch terminals are thus not considered until they reach JFK so that the passenger arrival rate at JFK station is modified to include both passengers boarding at JFK and passengers arriving from the two branches. Also, trains arriving at JFK are next considered dispatched in the reverse direction, in the same fashion as at any regular terminal.

The scenarios presented in this section are based on disruptions occurring on the trunk portion of the line at 8:15AM, during the morning peak period. Since the model formulation was tested off-line, train location and passenger load information were derived from the dispatching schedule and nominal running times instead of using real-time information.

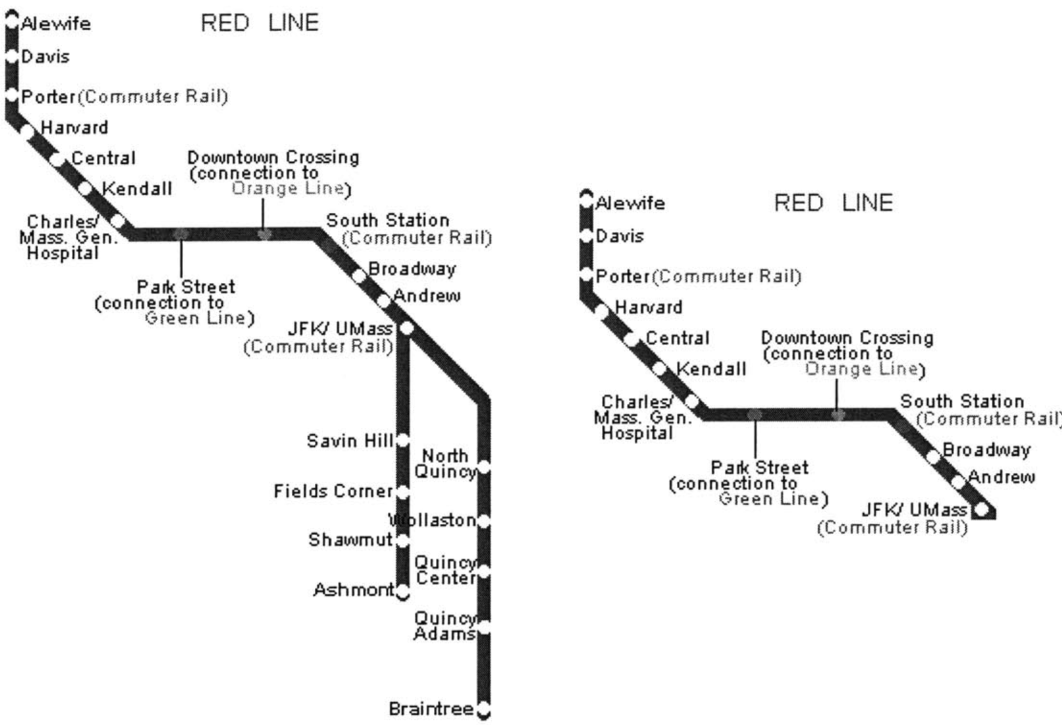

Figure 4-1: The MBTA Red Line

4.1.2 Input Data

Passenger arrival rates and alighting fractions at each station (see Table 4.1) were estimated from data collected by the Massachusetts' Central Transportation Planning Staff (CTPS) [32]. The CTPS data consist of detailed counts of passengers arriving and alighting at each station, for fifteen minute intervals throughout the day.

Station Name	Station Acronym	Passenger Arrival Rate (pax/min)	# Alightings per Train	Departing Loads of Trains	Alighting Fraction
JFK	JFK	147.6	0	633	0.00
Andrew	AND	10.5	11	657	0.02
Broadway	BRW	6.3	15	671	0.02
South Station	STA	24.3	198	572	0.30
Downtown Crossing	DTX	19.6	272	408	0.48
Park Street	PKS	18.1	170	286	0.42
Charles MGH	MGH	4.7	56	253	0.20
Kendall	KEN	1.3	96	162	0.38
Central	CEN	2.6	45	130	0.28
Harvard	HAR	4.3	87	57	0.67
Porter	POR	1.0	13	47	0.23
Davis	DAV	0.8	14	37	0.29
Queuing Location	QUE	0.0	0	37	0.00
Alewife	ALW	38.7	37	153	1.00
Davis	DAV	44.3	2	270	0.01
Porter	POR	30.1	4	382	0.01
Harvard	HAR	37.3	53	475	0.14
Central	CEN	27.2	25	487	0.05
Kendall	KEN	5.6	72	468	0.15
Charles MGH	MGH	3.7	42	442	0.09
Park Street	PKS	21.2	134	399	0.30
Downtown Crossing	DTX	18.3	171	306	0.43
South Station	STA	3.6	190	132	0.62
Broadway	BRW	0.5	12	124	0.09
Andrew	AND	1.3	8	120	0.07

Table 4.1: Station-specific parameters

4.2 Disruption Description

The problems analyzed here are a blockage on the Northbound tracks at the Harvard Square Station and a blockage located at Porter Square Southbound. The disruption durations are assumed to be twenty minutes in the case of Harvard Northbound and fifteen minutes at Porter Square Southbound. In both cases, all initial train headways are equal to four minutes. In both disruption scenarios, one major consequence of the blockage is an increased headway in front of the blocked train, that is propagated in the southbound direction if control actions are not taken. As many passengers travel from the Alewife terminal to the central business district (Park Street-South Station) during the morning peak, this would result in increased passenger waiting time, large passenger accumulation and overloaded trains at stations in the Southbound direction. Thus, appropriate holding actions should show significant benefits in this case, both reducing in-platform waiting time at stations and the number of passenger left behind by the blocked trains.

In the two disruption cases considered, train locations[1] (see Tables 4.2 and 4.3) are derived from the knowledge of train running times between stations[2] and assuming four-minute headways between trains. Trains' initial loads $L_{i,m(i)}$ are calculated using passenger arrival rates and alighting fractions at stations (we assume that the preceding headways of all trains at all stations are also four minutes). We take ten minutes as the maximal deviation from the scheduled dispatched time for all trains (referred to as Ψ in the model).

Sensitivity analysis of the model solution is studied by solving the Harvard disruption case for different values of the model parameters. We first applied the holding model assuming infinite train capacity and without considering the effects of holding on on-board passengers. We then applied the model using a finite train capacity and different values for the weight μ of the in-vehicle delay time versus in-platform waiting time: μ =0, 0.1 and 0.5. The Porter Square disruption case was solved with a finite train capacity and $\mu = 0.5$. Results are presented and discussed in the following section with more details provided in Appendices A and B.

[1]In the Porter Southbound disruption, train -1 just departed from Alewife and train -2 is sitting at this station. We included the remaining layover time of train -2 (four minutes) in its holding time and thus, the current headway is 0 minute. In the same fashion, train -3 is currently traveling in the reverse direction with an actual preceding headway of four minutes. We model it as sitting at Alewife with $H_{-3} = 0$ and $m(-3) = 1$.

[2]Taken from Shen[30].

Station	JFK	AND	BRW	STA	DTX	PKS	MGH	KEN	CEN	HAR	POR	DAV
Train -6	*											
Train -5		*										
Train -4			*									
Train -3				*								
Train -2							*					
Train -1									*			
Train 0										Blockage		
Train 1											*	
Train 2												*

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Terminal Train T_1	*											
Terminal Train T_2	*											
Reverse Train 1_R			*									
Reverse Train 2_R					*							
Reverse Train 3_R							*					
Reverse Train 4_R									*			
Reverse Train 5_R											*	

Table 4.2: Initial train locations: Harvard Northbound disruption case

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train - 2	*												
Train - 1		*											
Train 0			Blockage										
Train 1					*								
Train 2							*						
Train 3									*				
Train 4											*		
Terminal Train T_1													*
Terminal Train T_2													*

Station	AND	BRW	STA	DTX	PKS	MGH	KEN	CEN	HAR	POR	DAV
Reverse Train 1_R	*										
Reverse Train 2_R		*									
Reverse Train 3_R			*								
Reverse Train 4_R						*					
Reverse Train 5_R								*			
Reverse Train 6_R									*		
Reverse Train 7_R										*	
Reverse Train -3											*

Table 4.3: Initial train locations: Porter Square Southbound disruption case

4.3 Model Results

4.3.1 Minimizing In-Platform Waiting Time ($\mu = 0$) with Infinite Train Capacity

We first applied the general holding model for the Harvard disruption case with the assumption of infinite train capacity and without considering the impact of holding trains on on-board passengers ($\mu = 0$). The resulting optimal holding times and headways are summarized in Tables 4.4 and 4.5[3] respectively.

Headway Distribution and Holding Actions Pattern

From Table 4.4, the optimal holding pattern produced by our model results in nearly perfectly even headways (at each station, across all trains). This observation is consistent with the result derived by Welding in [29], which states that passenger waiting time at a given station is minimized when the variance of headways between trains is minimized:

$$\overline{WT} = \frac{\bar{h}}{2}\left(1 + \frac{Var(h)}{\bar{h}^2}\right) \tag{4.1}$$

where:

$\overline{WT} =$	average waiting time of passenger
$\bar{h} =$	mean headway of trains arriving at this station
$Var(h)$	variance of train headways

[3]No holding action is taken for trains/stations that are not shown in the tables. Blocked train 0 and trains queue behind the blockage and, after the blockage is cleared, are not held at stations except at the terminal where they are held for the minimum turn-around time.

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 1		10.0	3.3	0.4	2.3	2.8	0.5	0.0	0.4	0.0	0.3	0.0	0.2	0.0	0.2	0.0
Train 2			6.7	2.3	4.5	2.1	1.1	0.0	0.8	0.0	0.6	0.0	0.4	0.0	0.3	0.0
Train T_1					11.0	1.4	1.6	0.0	1.1	0.0	0.8	0.0	0.7	0.0	0.5	0.0
Train T_2					4.5	0.7	2.1	0.0	1.5	0.0	1.1	0.0	0.9	0.0	0.7	0.0
Train 1_R							2.7	0.0	1.9	0.0	1.4	0.0	1.1	0.0	0.9	0.0
Train 2_R									2.3	0.0	1.7	0.0	1.3	0.0	1.0	0.0
Train 3_R											2.0	0.0	1.6	0.0	1.2	0.0
Train 4_R													1.8	0.0	1.4	0.0
Train 5_R															1.6	0.0

Table 4.4: Holding times (min): Harvard Northbound disruption; $\mu = 0$, infinite capacity

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	24.0	14.0	10.7	10.3	10.0	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 1		14.0	10.7	8.7	6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 2			10.7	13.0	6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train T_1					6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train T_2					6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 1_R							6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 2_R									6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 3_R											6.0	6.0	5.8	5.8	5.6	5.6
Train 4_R													5.8	5.8	5.6	5.6
Train 5_R															5.6	5.6

Table 4.5: Preceding departing headway: Harvard Northbound disruption; $\mu = 0$, infinite capacity

By inspecting the locations and values of the holds in Table 4.4, along with the headway sequences across stations, the following properties of the optimal holding strategy are found:

i) No train is considered for (active) holding at a station located between two stations $m(i)$ and $m(i+1)$

ii) The value of the constant headway sequence decreases, as we move down the line

iii) At a given station, a train's holding time is smaller than its preceding train's holding time.

iv) For a given train traveling in a given direction, its holding time (at holding stations) is monotically decreasing

To understand these properties, we must first note that at a station m only the headways of trains 0 and ahead are affected by holding. If we denote by h_i train i's departure headway at the preceding station and by r_i its holding time at station m, the sequence of departing headways at station m is as follows:

$$\begin{array}{l} h_p + r_p \\ h_{p-1} + r_{p-1} - r_p \\ h_{p-2} + r_{p-2} - r_{p-1} \\ \vdots \\ h_1 + r_1 - r_2 \\ h_0 - r_1 \qquad \text{(train 0 is not held)} \end{array} \tag{4.2}$$

where train p is assumed to be the first controlled train at station m so that $p+1$ headways departing station m are affected by the holding actions.

Property *i)* suggests that, at a station m located between two stations $m(i)$ and $m(i+1)$, there may not be holding actions such that $\overline{WT} = \bar{h}/2 \times (1 + Var(h)/\bar{h}^2)$ is decreased. Indeed, we first note that at stations m such that $m(i) \leq m < m(i+1)$, the same sequence of headways is to be evened out and is of the form shown in Equations (4.2). By summing the $p+1$ equations in (4.2), we observe that the holding terms cancel out so that the average headway $\bar{h}$ is constant across stations $m \in [m(i), \ldots, m(i+1)-1]$. Thus, according to Welding's formula, waiting time at these stations is minimized only if $Var(h)$ is minimized. This leads to evening out the headway sequence by holding trains at station $m(i)$, but no control should be considered at a station $m \in [m(i)+1, \ldots, m(i+1)-1]$: otherwise $Var(h)$ would be increased and the waiting time at this station would not be minimal.

On the other hand, a new train $p+1$'s headway is added to the sequence of headways coming out of station $m(i+1)$. If this headway is not equal to the following trains' even headway entering this station, control actions should then be considered to achieve a new even headway sequence at station $m(i+1)$. We refer to Table 4.5 to illustrate this: the departing headway sequence is even and equal to six minutes at Charles MGH and no train is held at the following station (Park Street) –thus keeping a six-minute headway sequence. At Downtown Crossing, train 4_R's departing headway is added to the sequence of headways to be evened out. The value of this headway without control is four minutes, which is different from the six-minute headway sequence at Park Street. Thus trains are held at Downtown Crossing to yield a 5.8-minute departing headway sequence[4].

Since in our scenario, the additional uncontrolled headway to be added is always equal to four minutes, it also implies that evening out a new headway sequence at a later station (where a four-minute headway is added) yields a lower even headway: this is the property *ii)* above.

Property *iii)* can also be demonstrated by using expression (4.2). At a given con-

[4]In light of Equation (4.2) and the aforementioned cancellation of the holding times in summing headways in (4.2), one can verify that the average entering headway ((8×6 min.+4 min.)/9=5.77 min.) is equal to the average headway departing this station (5.8 min.).

trol station the corresponding sequence of headways is evened out and we can write $h_1 + r_1 - r_2 = h_0 - r_1$, which is equivalent to $r_2 - r_1 = r_1$ since $h_0 = h_1$ (since the entering headways are assumed to be even). As $r_1 \geq 0$, we obtain $r_2 \geq r_1$. A similar argument by induction can be used to prove that train $i + 1$'s hold is longer than train i's hold (for $i \leq p - 1$): we again write that $h_i + r_i - r_{i+1} = h_{i-1} + r_{i-1} - r_i$ and $h_i = h_{i-1}$, which yields $r_{i+1} - r_i = r_i - r_{i-1} \geq 0$, the last inequality being implied by the induction hypothesis. From a more intuitive point of view, we can also observe that train i's departure headway at station m is $h_i + r_i - r_{i+1}$ for $i = 0, \ldots, p$, so that train i's headway is changed by the difference between its holding time and that of its preceding train, namely $r_i - r_{i+1}$. Since a smaller headway h_{p+1} is introduced in the headway sequence, this sequence is evened out by reducing train i's headway. This implies that $r_i - r_{i+1}$ must be negative, that is train $i+1$ is held longer than train i.

We give here an intuitive interpretation of property *iv)*: for a given train, we know that its controlled headway at a control station is equal to the value of the achieved even headway sequence. From properties *i)* and *ii)*, we know that for two control stations, the value of the headway sequence is smaller (and closer to the normal headway) at the second station downline. Also, there are more trains controlled there as well. Hence, the amount of holding for a given train should be less at the second station in order to achieve even headways.

Maximal Deviation from Schedule Constraint

Nevertheless, we note from Table 4.4 that the abovementioned properties do not hold for all trains at all stations. In particular, trains are held at Davis Square Inbound (which is not a station $m(i)$) and the corresponding holding times are not decreasing. Also, even headways are not achieved at either the queuing location or at Alewife.

Uneven headways are permissible at the queuing location as no in-platform waiting time is associated with headways here: the objective value is not a function of the headway distribution at this "virtual" station.

The two other points are explained by observing from Table 4.4 that the cumulative holding time of train 1 at Alewife is 16 minutes[5]. Since train 1's layover time at the beginning of the disruption is six minutes and the maximal deviation from schedule is ten minutes, this means that the maximal deviation from schedule constraint is binding for train 1, which forces it to be dispatched from Alewife after being held for only 2.3 minutes. Limiting the hold at Alewife results in an uneven departure headway sequence at Alewife: train 0's headway is ten minutes while preceding trains left this station with six-minute headways. As the headway sequence "entering" Davis is uneven, trains are held at this station to achieve even departure headways and smaller waiting time even though this is not a station $m(i)$. Also, the holding time sequence at Davis is increasing, as distinct from other stations: this can be derived mathematically following the same argument as for decreasing holding times[6].

Conclusion

We conclude that, under some "ideal" conditions (no consideration of the holding cost, no maximal deviation from schedule constraints, and infinite train capacities), minimizing passenger waiting time is equivalent to achieving perfectly even headway sequences at stations. The control actions needed to achieve this regularity follow a special pattern, resulting in smaller headways ahead of the blocked trains. It is noted that these conditions are necessary for this equivalence to hold strictly.

Moreover, the holding strategy required to achieve this goal is complicated, as shown in Table 4.4: trains are held at multiple stations along their trip, which might be difficult for dispatchers to implement.

Finally, it is clear that not considering the effects of holding on on-board passengers is unrealistic as passengers held at multiple stations are likely to be disgruntled by such delays. Therefore, such an assumption and the derived strategy are not appropriate for real application.

[5]Train 1 is held 10 minutes at Porter Square, 3.3 minutes at Davis Square, 0.4 minutes at the queuing location and 2.3 at Alewife.

[6]We have here one long headway followed by an even sequence of smaller headways at Davis but an even sequence of headways followed by one smaller headway at other stations.

4.3.2 Minimizing In-Platform Waiting Time ($\mu = 0$) with Finite Train Capacity

a) Analysis

Headway Distribution and Holding Pattern Solving the same problem with *finite train capacity* yields a much different holding pattern as shown in Tables 4.6 and 4.7. Here, trains from train 0 through the second terminal train are fully loaded when departing from Kendall Square Southbound, thus limiting the holding actions earlier on the line and the possibility of achieving even headway sequences at stations (see Tables 4.8 and 4.9[7]). Also, train 0 is fully loaded and leaves passengers behind at Porter, Harvard and Central stations. Therefore, the train capacity constrains the length of the holds on trains T_1 and T_2 up to Kendall Station, which results in uneven headway sequences.

This limitation is better understood if we separate the trains into two different groups composed of trains 0 through T_2, and the reverse trains. Within each of these groups, the headway distribution is regular (except for train 0) so there are clearly different impacts of the holding actions for each group. For the group of reverse trains, the train capacity constraint is not binding at any station, so that perfectly even headways between these trains are achieved. Trains 0 through T_1 are only affected by the holding actions on themselves and on train T_2 (which is also constrained by the train capacity issue arising at Kendall Station). This results in a perfectly even headway distribution for train 0 through T_2 at all stations except at stations between the queuing location and Kendall (inclusive), where train 0's headway is different.

[7]Trains in the Table 4.9 are shown in a reverse order to better show the "spillover" effect of passengers left behind.

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 1		10.0	3.3	0.3	2.3	1.5	0.6	0.0	0.3	1.7	0.0	0.0	0.2	0.0	0.2	0.0
Train 2			6.7	2.3	4.6	1.1	1.4	0.1	0.8	1.2	0.0	0.0	0.3	0.0	0.3	0.0
Train T_1					11.0	0.7	2.2	0.1	1.3	0.8	0.0	0.0	0.5	0.0	0.5	0.0
Train T_2					4.5	0.4	3.0	0.2	1.8	0.4	0.0	0.0	0.6	0.0	0.7	0.0
Train 1_R							3.8	0.2	2.3	0.0	0.0	0.0	0.8	0.0	0.9	0.0
Train 2_R									3.2	0.0	1.1	0.0	1.1	0.0	1.0	0.0
Train 3_R											2.1	0.0	1.4	0.0	1.2	0.0
Train 4_R													1.8	0.0	1.4	0.0
Train 5_R															1.6	0.0

Table 4.6: Holding times (min): Harvard Northbound disruption;
$\mu = 0$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	24.0	14.0	10.7	10.3	10.0	8.5	7.9	7.9	7.6	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train 1		14.0	10.7	8.7	6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train 2			10.7	12.9	6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train T_1					6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train T_2					6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train 1_R							7.8	8.0	7.2	7.2	6.1	6.1	5.8	5.8	5.6	5.6
Train 2_R									7.2	7.2	6.1	6.1	5.8	5.8	5.6	5.6
Train 3_R											6.1	6.1	5.8	5.8	5.6	5.6
Train 4_R													5.8	5.8	5.6	5.6
Train 5_R															5.6	5.6

Table 4.7: Preceding departing headway: Harvard Northbound disruption;
$\mu = 0$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train -2	104	43	35	26	26	77	165	224	267	308	273	256	221	163	69	64	62
Train -1	113	46	37	28	28	77	165	255	454	645	559	516	404	267	109	100	96
Train 0	23	110	98	78	78	387	761	960	960	960	849	795	682	495	209	193	187
Train 1		23	31	31	31	251	553	730	852	960	849	795	682	495	209	193	187
Train 2			14	18	18	251	553	730	852	960	849	795	682	494	209	193	187
Terminal Train T_1					0	251	553	730	852	960	849	795	682	494	209	193	187
Terminal Train T_2					0	251	553	730	852	960	849	795	682	494	209	193	187
Reverse Train 1_R							331	562	781	938	837	784	678	493	208	192	186
Reverse Train 2_R									548	716	649	613	558	424	182	169	164
Reverse Train 3_R											566	537	506	394	171	158	155
Reverse Train 4_R													489	385	167	155	152
Reverse Train 5_R															234	216	208

Table 4.8: Entering train loads: Harvard Northbound disruption; $\mu = 0$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train 0	0	0	0	0	0	0	31	160	159	0	0	0	0	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
													Additional Wait = 700 pax-min				

Table 4.9: Passengers left behind: Harvard Northbound disruption; $\mu = 0$, capacity = 960 passengers/train

At the queuing location and Alewife, this is explained by the same reasons as in the case of infinite train capacity: no waiting time is associated with the queuing location while train 1's remaining layover time limits the evenness of train headways at Alewife. At Porter, Harvard and Central, the train capacity constraint is binding for train 0 and the additional waiting time incurred by passengers left behind by this train must be included in the objective function. Total in-platform waiting time *–which includes the extra wait due to denied boardings–* is minimized at these stations for headway sequences that are not even. This is permissible since the evenness of headways was an optimality condition for minimizing *in-platform waiting time incurred by passengers who are not left behind* (this quantity will be referred to as the *in-platform waiting time* throughout the remainder of the thesis, as opposed to the *total in-platform waiting time*, which includes the extra waiting time incurred by passengers left behind). We note here that the *total in-platform waiting time* is minimized for headway values that are smaller for trains 1, T_1 and T_2 but greater for the capacitated train 0. This suggests that the savings in waiting time for the former trains outweigh the additional waiting time experienced by passengers left behind by train 0.

At Davis Square, the headways are not even, even though no capacity issue arises at this station (from Table 4.8, train 0's departing load is 761). This result seems *a priori* counterintuitive but can be explained as follows. The departing headways achieved at this station not only impact the *in-platform waiting time* at this station, but also the train passenger loads at following stations, where the train capacity issue arises. Thus, the headway sequence achieved at this station also determines the value of the *total* in-platform waiting time at following stations and thus its unevenness is not a violation of the optimality conditions.

Moreover, the impacts of the holding strategy on each of the two groups of trains are clearly not independent, as holding train 1_R affects train T_1's headway. From Table 4.6, it is noted that headways among all trains are essentially equal at any station beyond Charles MGH. The optimal solution seeks to achieve some level of "continuity" in the headways between these two groups of trains although the number of stations where this can be achieved is limited because of the train capacity constraint.

Furthermore, we note that *the properties highlighted in the case of an infinite train capacity do not generally hold here.* Specifically, these properties do not necessarily hold in sections of the line affected by train capacity. For instance, train 1 is held longer than train 2 at Kendall, as the regularity of the headway sequence departing Central Square is limited by the train capacity constraint: two headways of 7.2 minutes are followed by four headways of 5.5 minutes, and then followed by the 7.6-minute headway of train 0.

Passengers Left Behind The number of passengers left behind is greatly reduced by holding train 1 in front of train 0 (see Tables 4.10 and 4.11). When no trains are controlled, trains 0 through −3 leave passengers behind at Davis, Porter, Harvard, Central and Park Street stations. This results in a total number of 5090 passengers left behind and an additional waiting time of 10177 passenger-minutes incurred by these passengers. These passengers are first denied boarding by train 0, which adds to the level of congestion of the following trains as overloads "spill over". When holding is applied, the reduced preceding headway of the blocked train 0 (less than ten minutes) reduces the number of passengers left behind (by train 0) at Porter, Harvard and Central stations to only 350 passengers, resulting in additional waiting time of only 700 passenger-minutes. This is important as denying passengers boarding is badly perceived and should be avoided as much as possible.

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train 0	0	0	0	0	0	691	592	612	497	0	0	106	0	0	0	0	0
Train -1	0	0	0	0	0	0	540	552	503	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	323	510	0	0	0	0	0	0	0	0
Train -3	0	0	0	0	0	0	0	0	164	0	0	0	0	0	0	0	0
												Additional Wait = 10177 pax-min					

Table 4.10: Passengers left behind: No hold, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train 0	0	0	0	0	0	0	31	160	159	0	0	0	0	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
												Additional Wait = 700 pax-min					

Table 4.11: Passengers left behind: $\mu = 0$, capacity = 960 passengers/train

b) Evaluation and Comparison of Different Control Strategies

In order to assess the efficiency of this control strategy and others–in comparison with the "No Hold" strategy, we must define a common set of trains and stations for which in-platform waiting times and in-vehicle delay times are evaluated and compared. For instance, in the case of the Harvard Northbound disruption, all trains located downline of the disruption, at the terminal and in the reverse direction are subject to control actions. Moreover, there are no benefits from the holding actions (exerted at any station) at stations located behind the blockage, since holding only affects headways at stations ahead of the blockage. Thus, passenger times associated with the controlled trains (and train 0) are evaluated at each train's first control station $m(i)$ through the end boundary of the modeled line, station $2M - 3$. For trains behind the blockage, we know that trains 0 through -3 leave passengers behind at stations located behind the blockage when no control actions are taken. Hence, passenger times associated with trains -1 through -4 are evaluated at stations from Porter Square through $2M - 3$, since control actions reduce the number of passengers boarding these trains at these stations. Trains -5 and -6 are not impacted by any control action and thus are not included in the set of trains: they queue behind the blockage and then run at maximal speed with a minimum safe headway once the blockage is removed.

The Harvard Square scenario was solved by O'Dell [26], with a holding model that considered the branching structure of the MBTA Red Line but did not consider in-vehicle delay time (that is, $\mu = 0$ for our model). Detailed implementation results are not shown in [26] but the passenger-time savings were indicated: O'Dell found that holding led to a passenger time savings of 46%, while we found a value of 49% using the train and station impact sets described above.

c) Conclusion

The previous analysis strongly supports the view that the headway distribution must have a high level of regularity to be optimal, but that this optimization goal is *con-*

strained by the train capacity. At stations where this constraint is binding, achieving perfectly even headway sequences does not necessarily lead to minimal waiting times since waiting time of passengers left behind must be accounted for. Rather, different headway distributions are achieved for different groups of trains, each of these experiencing different levels of congestion during the observation period.

Hence, we can only conclude that *minimizing total in-platform waiting time with finite train capacity is equivalent to achieving headway sequences at stations so that fewer passengers are left behind* and *headway variance is minimized within different groups of trains*. We also observed that the number of passengers left behind was substantially decreased.

4.3.3 Minimizing In-platform Waiting Time and Optimal Holding Structure

From the two previous sections, we observe that minimizing only in-platform waiting time leads to holding patterns that look both complex to implement and difficult to justify from the dispatcher's perspective. For instance, holding is considered at Central Square Southbound while departing loads are high from this station.

In the case of infinite train capacity, this apparent complexity can be simply interpreted if we have in mind that, minimizing *in-platform* waiting time is theoretically equivalent to achieving regular headways at each station, under some operational constraints. As achieving regular headways at one station impacts the headways at other stations for the same trains, we see the complex holding pattern as the constrained solution to obtaining a regular headway pattern under the specified constraints.

However, we have shown that this equivalence does not hold when a finite train capacity is considered and the corresponding constraint is binding. In this case, passengers left behind by overloaded trains incur additional waiting time. This additional time is included in the objective function, which cannot be written under the form shown in Equation 4.1. Consequently, even headways do not necessarily lead to minimal *in-platform waiting time*, and the structure of the holding solution shows a level of complexity that cannot always be explained as simply as in the case of infinite train capacity.

4.3.4 Minimizing Total Waiting Time

As mentioned above, not accounting for the in-vehicle delay due to holding might be inappropriate as the extra ride time incurred by passengers on-board held trains might not be negligible relative to in-platform waiting time. Nevertheless, we must differentiate between these two types of times as passengers on board a train are likely to perceive the time spent on board a train as being less onerous than on-platform waiting time. Thus, in-vehicle delay is generally perceived as less detrimental than additional waiting at platforms.

To explore this issue, we solved the Harvard Northbound disruption case with two non-zero values for the relative weight μ of in-vehicle delay against in-platform waiting time ($\mu = 0.1$ and 0.5). An additional disruption case (Porter Southbound with finite train capacity and $\mu = 0.5$) was also solved to further illustrate our findings. Next we present the results from each of these model applications focusing on:

- The impacts of considering in-vehicle delay time on the effectiveness of the holding strategy
- The sensitivity of the optimal headway distribution and holding actions to the weight μ

a) Minimizing Total Waiting Time with $\mu = 0.1$

We first solved the Harvard Northbound disruption case for a value of $\mu = 0.1$, which is based on one minute of in-vehicle delay time being valued the same as ten minutes of in-platform waiting time. Results of this application are presented in Tables 4.12 through 4.15 below.

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	20.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1		10.2	0.8	0.0	4.9	1.5	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
Train 2			4.9	0.0	9.4	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_1					12.1	0.5	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_2					5.4	0.2	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1_R							3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2_R									1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 3_R											0.0	0.5	0.0	0.0	0.0	0.0
Train 4_R													0.0	0.0	0.0	0.0
Train 5_R															0.0	0.0

Table 4.12: Holding times (min): Harvard Northbound disruption;
$\mu = 0.1$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	24.0	13.8	12.9	12.9	10.0	8.5	8.5	8.5	8.5	8.5	8.5	8.1	8.1	8.1	8.1	8.1
Train 1		14.2	10.1	10.1	5.7	6.4	6.4	6.4	6.4	6.4	6.4	6.8	6.8	6.8	6.8	6.8
Train 2			8.9	8.9	6.2	6.5	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train T_1					6.8	7.0	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7
Train T_2					7.4	7.6	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3
Train 1_R							7.9	7.9	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train 2_R									5.8	5.8	5.8	5.3	5.3	5.3	5.3	5.3
Train 3_R											4.0	4.5	4.5	4.5	4.5	4.5
Train 4_R													4.0	4.0	4.0	4.0
Train 5_R															4.0	4.0

Table 4.13: Preceding departing headway: Harvard Northbound disruption;
$\mu = 0.1$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train -2	104	43	35	26	26	77	165	224	267	308	273	256	221	163	69	64	62
Train -1	113	46	37	28	28	77	165	274	495	709	614	566	438	287	116	107	102
Train 0	23	110	98	80	80	387	761	960	960	960	864	817	743	572	247	229	223
Train 1		23	31	30	30	220	500	686	828	960	852	798	704	526	225	208	203
Train 2			14	17	17	239	523	703	834	960	850	797	688	505	214	198	193
Train T_1					0	262	571	737	847	960	848	793	676	490	207	191	186
Train T_2					0	284	618	771	859	960	846	789	664	475	200	184	179
Train 1_R							331	566	783	912	809	759	662	490	209	193	188
Train 2_R									548	678	609	575	515	391	168	155	152
Train 3_R											566	530	466	348	148	137	134
Train 4_R													489	352	148	137	133
Train 5_R															234	215	205

Table 4.14: Entering loads: Harvard Northbound disruption; $\mu = 0.1$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train 0	0	0	0	0	0	0	50	184	185	0	0	0	0	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Additional Wait = 839 pax-min

Table 4.15: Passengers left behind: Harvard Northbound disruption; $\mu = 0.1$, capacity = 960 passengers/train

Holding Pattern and Cost When the cost of holding is included in the objective function and even with the assumption that ten minutes of in-platform waiting time is equivalent to only one minute of in-vehicle time, the optimal holding pattern becomes very simple as only a few trains are held at only a few stations. As shown in Table 4.12, seven trains are actively held at seven stations in this case (resulting in a total active holding time of 39.3 minutes), as compared to nine trains held at eleven stations for a total holding time of 72 minutes when $\mu = 0$. This suggests that *the holding costs incurred by on-board passengers can be large* so that, even for small values of μ, this quantity must be considered.

Table 4.16 below provides support for this statement[8] (the time incurred by passengers left behind is indicated within parentheses). When in-vehicle delay time is not included in the objective function, this quantity is equal to 22361 pax-min, which is comparable to the value of 34816 pax-min of total in-platform waiting time. For $\mu = 0.1$, the in-vehicle delay time is decreased to 4740 pax-min, that is a 79% decrease at a cost of 600, or only 2%, extra pax-min of total in-platform waiting time.

Model Parameters	Total In-Platform Waiting Time (pax-min)	Unweighted In-Vehicle Time (pax-min)
No Hold	68794 (10177)	0
$\mu = 0$	34816 (700)	22361
$\mu = 0.1$	35418 (839)	4740

Table 4.16: In-platform waiting time and in-vehicle delay time

As expected, holding actions are exerted on trains which are not heavily loaded and preferably at stations without high passenger through volumes as holding such trains yields less in-vehicle delay time. We also note that no hold takes place at the queuing location. As discussed in Section 3.1.1, this is because holding at this location has an associated cost but no associated benefits.

[8]The waiting time associated with each train at each station and the cost associated with both passive and active holding are shown in detail in Appendix A.

Moreover, we can infer from these results that after the disruption starts, trains are mainly held in their travel direction at the next station reached $m(i)$ and/or a few stations beyond it. For instance, train 1 is held 10.2 minutes at Porter Square in the disruption direction and only 0.8 minutes at the following station. This suggests that an optimal strategy seeks to control trains at "early" stations and obtain benefits from these early actions further down the line, since holding a train at a station not only modifies its departure headway at this station but also at later stations[9]. *Hence, holding a train at one of the earliest stations arrived at can yield significant benefits down the line and avoid the cost of holds at later stations.*

Yet, there also exists a tradeoff between holding a train at the first station arrived at and holding it at a later station, which stems from the costs (in-vehicle delay) and benefits (smaller waiting time at following stations) achieved in each case. For instance, Train 3_R is not held at Charles MGH but at the following station (Park Street) as fewer people would be negatively affected by the hold at the latter station (466 versus 530 passengers) and few people would benefit from a hold at Charles MGH.

Furthermore, it is noted from Table 4.12 that, when in-vehicle delay time is taken into account, the locations of the active holds are shifted from trunk line stations to the terminal. Trains arriving at the terminal are held beyond the scheduled layover time but incur no (or few) holds at further stations. Indeed, terminal holding and use of the scheduled layover time to "buffer" against the delay are preferred as no extra ride-time cost is associated with terminal holding[10]. This implies that *delay recovery is preferably performed at the terminal* in order to minimize the negative impacts of the disruption in the reverse direction. For instance, terminal train T_1 is held 6.1 minutes more than its scheduled layover time of six minutes, and is held for only 0.5 and 0.3 minutes respectively at Davis and Porter.

[9]The preceding train's hold also modifies it.

[10]Holding has no associated costs other than the incurred additional waiting time for departure, since there are no through-standees at the terminal stations ($\alpha_M = 1$ and thus, $(1 - \alpha_M)L_{i,M} = 0$).

Headway and In-Platform Time Distribution Although holding costs are not negligible and passengers left behind incur extra-waiting time, in-platform waiting time is still the main component of the cost function to minimize. Thus, we focus here on the headway distribution and the associated waiting time distribution to assess the efficiency of holding with respect to in-platform waiting time when holding costs are accounted for.

From inspection of Table 4.13, it seems *a priori* that the headway sequences at stations show a high level of variability, which suggests that in-platform waiting time at stations is far from optimal. Nevertheless, Table 4.17 shows that in-platform waiting time is in fact only two percent above the minimum (34579 versus 33825), in spite of the uneven headway sequences: the waiting times at stations do not significantly deviate from the values in the case of infinite capacity and $\mu = 0$, for which all headways sequences were essentially even and the in-platform waiting time at stations was strictly minimized. This suggests that even though the smaller number of holds exerted at a few stations result in uneven headways, in-platform waiting time can still be significantly reduced and remain close to its minimal value. This counterintuitive result can be better understood if we note that at a given station, some headways are larger than the even headway value while other headways are smaller, so that the resulting waiting time increase for the former train is "compensated" by the decrease in waiting time for the latter ones[11].

Table 4.18 also shows that the gain in total passenger time (48%) is close to the one obtained for $\mu = 0$ ($\sim$ 49 %).

[11] Consider as an example the case of a two-headway sequence which is optimized for a common headway value of 4 minutes. In this instance, having an uneven headway sequence of three and five minutes results in a total waiting time of $0.5 * (3^2 + 5^2) = 17$ pax-min, assuming a hypothetical passenger arrival rate of one passenger per minute. This is only 6% above the minimal value of $0.5 * (4^2 + 4^2) = 16$ pax-min, reached for a common headway value of four minutes.

Station	Finite Capacity No Hold	Infinite Capacity $\mu = 0$	Finite Capacity $\mu = 0$	Finite Capacity $\mu = 0.1$
POR	294	197	197	197
DAV	250	143	143	146
ALW	9278	5509	5509	5540
DAV	10632	6096	6146	6167
POR	7455	4248	4311	4374
HAR	9250	5272	5364	5427
CEN	6972	3984	4061	4076
KEN	1434	819	825	838
MGH	968	557	557	578
PKS	5588	3217	3218	3307
DTX	4986	2900	2900	3011
STA	988	575	575	597
BRW	149	88	88	92
AND	373	220	220	230
Total	58617	33825	34115	34579

Table 4.17: Average in-plaform waiting time (pax-min) at stations

	Finite Cap. No Hold	Infinite Cap. $\mu = 0$	Finite Cap. $\mu = 0$	Finite Cap. $\mu = 0.1$
Total In-Platform Waiting Time (Left-Behind-Pax Time)	68794 (10177)	33825 (0)	34816 (701)	35418 (839)
Unweighted In-Vehicle Delay Time	0	21141	22361	4740
Weighted Total Time ($\mu = 0$)	68794	33825 (51%)	34816 (49%)	
Weighted Total Time ($\mu = 0.1$)	68794			35892 (48%)

Table 4.18: Benefits of control strategies

b) Minimizing Total Waiting Time with $\mu = 0.5$

We also solved the Harvard Northbound disruption case for a value of $\mu = 0.5$ accounting for a greater value of in-vehicle delay time relative to in-platform waiting time. Results of this application are presented in Tables 4.19 through 4.22. For the purpose of sensitivity analysis, we will highlight here the main similarities and differences with the case $\mu = 0.1$.

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	20.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1		7.9	0.0	0.0	8.1	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2			1.7	0.0	12.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_1					11.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_2					5.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1_R							2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2_R									0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 3_R											0.0	0.0	0.0	0.0	0.0	0.0
Train 4_R													0.0	0.0	0.0	0.0
Train 5_R															0.0	0.0

Table 4.19: Holding times (min): Harvard Northbound disruption;
$\mu = 0.5$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	24.0	16.1	16.1	16.1	10.0	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5
Train 1		11.9	10.2	10.2	5.8	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3
Train 2			5.7	5.7	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train T_1					6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train T_2					7.7	7.7	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1
Train 1_R							6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6
Train 2_R									4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 3_R											4.0	4.0	4.0	4.0	4.0	4.0
Train 4_R													4.0	4.0	4.0	4.0
Train 5_R															4.0	4.0

Table 4.20: Preceding departing headway: Harvard Northbound disruption;
$\mu = 0.5$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train -2	104	43	35	26	26	77	165	224	267	308	273	256	221	163	69	64	62
Train -1	113	46	37	28	28	77	165	344	590	826	713	656	502	323	130	119	114
Train 0	23	110	100	84	84	387	803	960	960	960	869	826	779	618	269	250	245
Train 1		23	29	29	29	225	504	689	829	960	851	798	693	511	217	201	195
Train 2			14	15	15	241	515	697	832	960	851	797	690	508	215	199	194
Train T_1					0	241	515	697	832	960	851	797	690	508	215	199	194
Train T_2					0	298	636	783	864	960	845	787	659	469	197	182	176
Train 1_R							331	526	699	843	754	710	637	484	208	193	188
Train 2_R									548	630	557	522	450	330	140	129	126
Train 3_R											566	530	455	333	141	130	127
Train 4_R													489	352	148	137	133
Train 5_R															234	215	205

Table 4.21: Entering loads: Harvard Northbound disruption;
$\mu = 0.5$, capacity = 960 passengers/train

	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Train 0	0	0	0	0	0	0	121	220	210	0	0	0	0	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
													Additional Wait = 1102 min				

Table 4.22: Passengers left behind: Harvard Northbound disruption;
$\mu = 0.5$, capacity = 960 passengers/train

Holding Pattern and the Cost of Holding From Table 4.19, we note that the simplicity of the holding pattern highlighted in the case of $\mu = 0.1$ is even more pronounced: five trains are held at five stations for a total holding time of 30.9 minutes. The reduced amount of holding simply reflects that holding is becoming more costly. As a consequence, the unweighted in-vehicle waiting time is decreased from 4740 pax-min for $\mu = 0.1$ to 1119 pax-min, a 76% reduction, for $\mu = 0.5$ at a cost of a further 1000 pax-min increase in in-platform waiting time, a 3% increase (see Table 4.23).

Model Parameters	Total In-Platform Waiting Time (pax-min)	Unweighted In-Vehicle Time (pax-min)
No Hold	68794 (10177)	0
$\mu = 0$	34816 (701)	22361
$\mu = 0.1$	35418 (839)	4740
$\mu = 0.5$	36390 (1102)	1119

Table 4.23: In-platform waiting time and in-vehicle time (with left behind pax time in parenthesis)

Except for train 1, holding occurs only at the next station arrived at or at the terminal[12]. This emphasizes the two observations that we made earlier: *i)* Early holding actions are preferred as they can yield benefits down the line while avoiding the costs of holds at a later station and, *ii)* Terminal holding and use of the layover time are effective as there are no associated holding costs.

Tables 4.24 and 4.25 summarize the above findings by showing the cumulative holding times by stations and by trains respectively. They clearly show, as μ increases, each train is held less and that there is less holding at non-terminal stations while terminal holding increases.

[12]Train 1 is held at Davis since its hold at the preceding station –Alewife– is limited by the maximum deviation from schedule constraint

	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	Total
Infinite Capacity, $\mu = 0$	10.0	10.0	2.7	2.3	7.0	8.0	0.0	8.0	0.0	7.9	0.0	8.0	0.1	7.8	0.0	72
Finite Capacity, $\mu = 0$	10.0	10.0	2.6	2.4	3.6	11.1	0.6	9.8	4.2	3.2	0.0	6.6	0.0	7.9	0.0	72
Finite Capacity, $\mu = 0.1$	10.2	5.8	0.0	11.8	3.0	5.8	0.0	1.8	0.0	0.0	0.9	0.0	0.0	0.0	0.0	39.3
Finite Capacity, μ =0.5	7.9	1.7	0.0	18.2	0.5	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	30.9

Table 4.24: Cumulative (active) holding times at stations (min): Harvard Northbound disruption

	-6	-5	-4	-3	-2	-1	0	1	2	T_1	T_2	1_R	2_R	3_R	4_R	5_R	Total
Infinite Capacity, $\mu = 0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.4	12.8	11.2	9.6	8.0	6.4	4.8	3.2	1.6	72
Finite Capacity, $\mu = 0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.4	12.8	11.2	9.6	8.0	6.4	4.8	3.2	1.6	72
Finite Capacity, $\mu = 0.1$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.9	9.1	6.9	5.2	3.9	1.8	0.5	0.0	0.0	39.3
Finite Capacity, μ =0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	10.5	8.2	5.9	3.7	2.6	0.0	0.0	0.0	0.0	30.9

Table 4.25: Cumulative (active) holding times of trains (min): Harvard Northbound disruption

In-Platform Waiting Time From Table 4.20, the headway sequences at stations again show a high level of variability, which suggests that in-platform waiting time at stations is far from optimal. Nevertheless, Table 4.27 again shows that in-platform waiting time is close to its optimal value in spite of the uneven headway sequences. Table 4.27 also shows that total passenger time savings are reduced to 46% for $\mu = 0.5$ against 48% for $\mu = 0.1$, but these savings are still both significant and comparable.

Station	Finite Capacity No Hold	Infinite Capacity $\mu = 0$	Finite Capacity $\mu = 0$	Finite Capacity $\mu = 0.1$	Finite Capacity $\mu = 0.5$
POR	294	197	197	197	202
DAV	250	143	143	146	165
ALW	9278	5509	5509	5540	5549
DAV	10632	6096	6146	6167	6276
POR	7455	4248	4311	4374	4412
HAR	9250	5272	5364	5427	5475
CEN	6972	3984	4061	4076	4215
KEN	1434	819	825	838	867
MGH	968	557	557	578	597
PKS	5588	3217	3218	3307	3446
DTX	4986	2900	2900	3011	3131
STA	988	575	575	597	621
BRW	149	88	88	92	95
AND	373	220	220	230	238
Total	58617	33825	34115	34579	35288

Table 4.26: Average in-plaform waiting time (pax-min) at stations

	Finite Cap. No Hold	Infinite Cap. $\mu = 0$	Finite Cap. $\mu = 0$	Finite Cap. $\mu = 0.1$	Finite Capacity $\mu = 0.5$
Total In-Platform Waiting Time (Left-Behind-Pax Time)	68794 (10177)	33825 (0)	34816 (701)	35418 (839)	36390 (1102)
Unweighted In-Vehicle Delay Time	0	21141	22361	4740	1119
Weigthed Total Time ($\mu = 0$)	68794	33825 (51%)	34816 (49%)		
Weighted Total Time ($\mu = 0.1$)	68794			35892 (48%)	
Weighted Total Time ($\mu = 0.5$)	68794				36949 (46%)

Table 4.27: Benefits of control strategies

c) The Porter Southbound disruption Case

Tables 4.28 through 4.31 present the application results of the Porter Southbound disruption case with $\mu = 0.5$. We note that the holding times of train -3 and train -2 at Alewife are respectively 17 minutes and 15 minutes. These times actually include the remaining layover time of train -2 (four minutes) so that these trains' actual holds (due to queuing) are 11 minutes for train -3 and 13 minutes for train -2. In the remainder of this section, we focus on the main findings from this application, in light of the previous results.

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	17.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train -2	15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train -1		13.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train 0			15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train 1					7.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.9	0.0	0.0	0.0	0.0
Train 2							0.0	2.3	0.0	0.0	0.0	0.0	0.0	11.4	0.0	0.0	0.0	0.0
Train 3									0.0	0.7	0.0	0.0	0.0	11.3	0.0	0.0	0.0	0.0
Train 4											0.0	0.0	0.0	10.3	0.0	0.0	0.0	0.0
Train T_1														8.7	0.0	0.0	0.0	0.0
Train T_1														3.4	0.0	0.0	0.0	0.0
Train 1_R															0.0	0.0	0.0	0.0
Train 2_R																0.0	0.0	0.0
Train 3_R																	0.0	0.0

Table 4.28: Holding times (min): Porter Square Southbound disruption; $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train 0			19.0	19.0	11.6	11.6	11.6	11.6	11.6	11.6	11.6	11.6	11.6	5.7	5.7	5.7	5.7	5.7
Train 1					11.4	11.4	11.4	9.1	9.1	9.1	9.1	9.1	9.1	5.7	5.7	5.7	5.7	5.7
Train 2							4.0	6.3	6.3	5.6	5.6	5.6	5.6	5.7	5.7	5.7	5.7	5.7
Train 3									4.0	4.7	4.7	4.7	4.7	5.7	5.7	5.7	5.7	5.7
Train 4											4.0	4.0	4.0	5.6	5.6	5.6	5.6	5.6
Train T_1														5.4	5.4	5.4	5.4	5.4
Train T_2														5.4	5.4	5.4	5.4	5.4
Train 1_R															4.0	4.0	4.0	4.0
Train 2_R																4.0	4.0	4.0
Train 3_R																	4.0	4.0

Table 4.29: Preceding departing headway: Porter Square Southbound disruption; $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	0	77	165	224	404	718	622	573	443	289	117	108	103	103	295	310	317	270
Train -2	0	580	663	716	960	960	827	760	574	364	146	134	127	127	295	310	317	270
Train -1		153	905	956	960	960	827	760	574	364	146	134	127	127	295	310	317	270
Train 0			270	838	960	960	881	844	836	689	304	283	278	278	838	880	899	767
Train 1					475	781	732	711	703	578	255	237	233	233	833	875	893	762
Train 2							468	441	445	372	162	150	147	147	833	875	893	762
Train 3									399	301	132	122	120	120	833	875	893	762
Train 4											132	122	119	119	830	873	891	760
Train T_1														0	798	839	856	731
Train T_2														0	798	839	856	731
Train 1_R															633	662	674	569
Train 2_R																657	669	565
Train 3_R																	671	567

Table 4.30: Entering loads: Porter Square Southbound disruption; $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train 0	0	0	0	470	267	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -1	0	0	0	406	274	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	137	280	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
													Additional Wait = 3668 pax-min					

Table 4.31: Passengers left behind: Porter Square Southbound disruption; $\mu = 0.5$, capacity = 960 passengers/train

From these tables, the structure of the optimal holding strategy follows the observations made for the Harvard Northbound disruption case:

- Trains are preferentially held at the earliest station arrived at, in order to benefit from the impacts of early control actions. This hold might be deferred to a later station if fewer people are negatively affected by the hold (and the benefits of the hold at stations down the line are comparable). For instance, Train 3 is not held at Downtown Crossing but at South Station since 132 on-board passengers would be delayed in the latter case against 301 passengers in the former case (see Table 4.30 and Table 4.30).

- Terminal holding and use of layover time can also be effective as they modify headways at stations down the whole line with no associated in-vehicle delay time costs.

In the case of the Harvard Northbound disruption, we noted that the headway distribution was irregular but that in-platform waiting time savings were close to its optimal value. Here, we observe from Table 4.29 that the headway sequences of trains 1, 2, 3, 4, T_1 and T_2 are essentially even at stations located after the JFK terminal[13]. *Thus, terminal holding is used in order to even out train headway sequences.*

Moreover, we determined that the passenger time is decreased "only" by 19% , as compared to 46% for the Harvard Northbound case. This suggests that there might be greater potential for short-turning in this disruption case.

d) Conclusion

We have shown that, when accounting for the costs of holding, the optimal holding strategy generally shows simple characteristics:

- Trains are preferentially held at the earliest station arrived at, in order to benefit from the impacts of early control intervention

[13]Such a pattern also occurs in the case of Harvard Northbound, but it is less obvious as it involves fewer trains: four are available for holding at the Alewife terminal.

- Terminal holding and use of layover time are also effective as they can even out headways at stations down the line without incurring in-vehicle waiting time

These characteristics can lead to optimal holding strategies that are easy to implement since only a few holding actions are required. Nonetheless, one might wonder whether such simple optimal holding strategies would be less efficient in terms of total passenger time savings compared to the one obtained without considering in-vehicle delay time ($\mu = 0$). From comparison of the achieved gains in both in-platform waiting time and total passenger time (see Table 4.27), we noted that the holding pattern is *no less efficient* that the "complicated" holding strategies shown in Tables 4.4 and 4.6. For instance, the gains in total passenger time are significant (46 % decrease for $\mu = 0.5$ against 48 % for $\mu = 0$ for the Harvard Northbound disruption), notwithstanding the striking difference in the structure of the optimal holding solution. This is in line with the findings of Barnett [5], who highlights the simplicity of the optimal strategies derived analytically, and notes that:

> This is interesting because there might, one might conjecture, be a reflexive fear that the use of detailed cost functions in optimization problems entails optimal strategies of such complexity as to be unimplementable. Our results suggest at least some situations where such fears do not seem warranted.

The structure and the efficiency of these optimal control strategies are of major importance since we showed that optimal holding strategies are sufficiently simple to be implemented by dispatchers and can yield large benefits[14]. This certainly supports the potential use of this holding model either in a computer based decision support system or in a manual guidance process.

4.3.5 Execution Time

We used version 12.0 of XPRESS-MP with a branch-and-cut strategy on an 800 MHz Pentium processor to solve the disruption scenarios described above. The execution

[14] This conclusion however contradicts the results from the model presented by Shen [30].

times are shown in Table 4.32. We also present in this table the effectiveness of the two-step solution procedure described in Chapter 4. For each scenario, we show the number of binary variables left after Step 1[15] of the solution procedure along with the solution time of each step. These times do not include the time needed to generate the model, which is independent of the model formulation.

Disruption Case	μ	# of $\nu_{i,m}$	# of $\nu_{i,m}$ Non-Fixed after Step 1	Solution Time without 2-Step Procedure (sec)	Solution Time of Step 1 (sec)	Solution Time of Step 2 (sec)
Harvard Square 20-minute	0.0	203	13	14	2	4
	0.1	203	13	56	1	3
	0.5	203	13	14	2	3
Porter Square 10-minute	0.5	116	17	19	1	3

Table 4.32: Execution Times

We note that, in all cases, the number of binary variables, which is the bottleneck of the solution procedure, was considerably reduced so that less than 20 binary variables (associated with the train capacity) remain at Step 2 of the solution procedure. The resulting solution times are significantly smaller: less than 6 seconds is necessary to achieve optimality with the two-step solution procedure, while 56 seconds were necessary to solve the Harvard case with $\mu = 0.1$ without the two-step solution procedure. For the other cases, the decrease is less pronounced but is still significant (it is reduced by at least by a factor of two).

For comparison purposes, O'Dell [26] solved the same scenario using CPLEX (Version 3.0) on a Sun SPARC 20 workstation with a solution time of 25 seconds. One could argue that this time and the one obtained by our model are both suitable for solving the deterministic holding problem in real-time. Nevertheless, this would not apply to the use of the models in a stochastic formulation of the problem. As accounting for uncertainty usually increases the size of the model, it is likely that the

[15]The solver was used here to solve the linear system of constraints. This is done by specifying no objective value and recording the first (and unique) feasible solution found.

model developed in this thesis would be better for such a purpose.

4.4 Conclusion

In this chapter, we presented and analyzed the implementation results of the model developed in Chapter 3. The results obtained confirmed that minimizing in-platform time is equivalent to achieving a regular headway pattern, under the system operational and physical constraints. In particular, the train capacity constraint is likely to be binding, which limits the holding actions. The patterns of the holding actions and the headway sequence were also explained.

In addition, the results showed that minimizing only in-platform time leads to holding strategies that are complicated and probably unimplementable. In constrast, including in-vehicle delay in the objective function yields simple optimal control strategies that can be implemented by dispatchers. The optimal control actions consist of holding a few trains at a few of the next stations they arrive at and/or at the terminal stations. Moreover, large gains in passenger waiting time can be achieved with these simple control actions and compared to the ones obtained with more complex holding strategies. The solution times obtained are also compatible with real-time operation controls and with a potential use of the model in a stochastic formulation of the train control problem.

Chapter 5

The Short-Turning Control Strategy

In the two previous chapters, we presented and analyzed a train holding model and derived general guidelines and results that can be used by dispatchers in developing effective holding strategies.

Although the benefits of holding in general seem large in terms of waiting time savings, as we discussed in Chapter 1, there are limitations to the holding strategy. Indeed, in the case of longer disruptions, train capacity limits the possibility of holding trains ahead of the blockage to achieve even headways, and of course spreading a longer delay over the trains ahead results in longer headways. This results in increased waiting times and possible congestion concerns at stations ahead of the blockage. In this case, an alternative to the holding strategy is to short-turn trains to compensate for the loss of service in the peak demand direction.

In this chapter, we develop the discussion in Chapter 1 to identify the main characteristics of the short-turning strategy. We will also show how the short-turning problem can be viewed as an extension of the holding sub-problem. In particular, our discussion will focus on several points that can be useful for the dispatchers to define appropriate short-turning actions:

- The conditions under which short-turning is an effective (complementary) alternative control operation to holding

- The critical factors in making a short-turning decision. In particular, we will highlight the importance of the type of short-turn (ahead of, or behind the blockage), the set of trains to be short-turned and the sequence of trains in the after-short-turn direction that can produce significant passenger time savings
- How, after these critical factors have been identified and analyzed in the short-turn decision-making process, the holding model might be used to determine the complementary holding actions that can provide the best service quality. These holding decisions are made, given the trains to be short-turned and the sequence of trains in the after-short-turn direction

We will first present a general analysis of the short-turn control strategy that summarizes and extends the discussion provided in Section 1.3.2. It also explains the critical factors in making a short-turning decision. For this purpose, we must differentiate two types of short-turn: short-turning ahead the blockage and short-turning behind the blockage. We will then show how the holding model might be modified to evaluate the benefits of each type of short-turning action. Control issues associated with each type of short-turn will be discussed and modeling solutions will be provided.

5.1 General Analysis of the Short-Turn Strategy

5.1.1 Types and Goals of the Short-Turning Strategy

Types of Short-Turning Strategy

According to Wilson et al. [36],

> [. . .] short-turning is the decision to turn a train before it reaches its terminus with the aim of reducing headway variance in the reverse direction by filling in a large headway gap.

Although this goal is essentially true for trains that are short-turned from the reverse direction to the disruption to fill the gap that is developing in front of the blockage (short-turning ahead of the blockage), this definition of the short-turning strategy can be further extended and refined.

First, there exist two types of short-turn control strategies as seen in Section 1.3.2. Indeed, most heavy rail transit lines serve a CBD on which heavy passengers flows are focused during the peak periods, and the CBD generally consists of only a few stations located in the middle of the line. Hence, three different disruption scenarios can be identified:

1. The blockage is located at, or close to, the CBD
2. The blockage is located before the CBD
3. The blockage is located after the CBD

The potential for short-turning and the issues associated with each of these three disruption scenarios are different. Here we briefly present these differences with more discussion of the issues specific to each corresponding short-turn strategy in later sections:

The blockage is located at, or close, to the CBD In this case, short-turning a train to serve the CBD would require a train to be short-turned at a cross-over track located near the CBD. As train loads are usually high in these areas, short-turning a train would result in many passengers being dumped and worsening congestion levels. Thus, for this type of disruption, short-turning is usually not an appropriate response.

The blockage is located before the CBD In this case, trains in the reverse direction have already served the CBD. If the blockage is sufficiently far from the CBD, train loads are likely to be low and trains with a low passenger load can be short-turned into the gap that is developing in front of the blockage. Consequently, the crossover track to use is the one closest to the disruption, downline from the disruption.

The ideal case would be a crossover track located just beyond the blockage as depicted in Fig. 5-1: trains in the off-peak direction would then be redirected behind the trains ahead of the blockage, thus virtually maintaining a close-to-normal service in front of the blockage. In this case, only a few holds would be necessary for the short-turned trains and the trains ahead to even out the headway sequence.

In the reverse direction, the short-turn results in a loss of service and the creation of a gap, due to the removal of the short-turn trains (see Fig. 5-1). We will provide a more detailed discussion of this issue in Section 5.2.1.

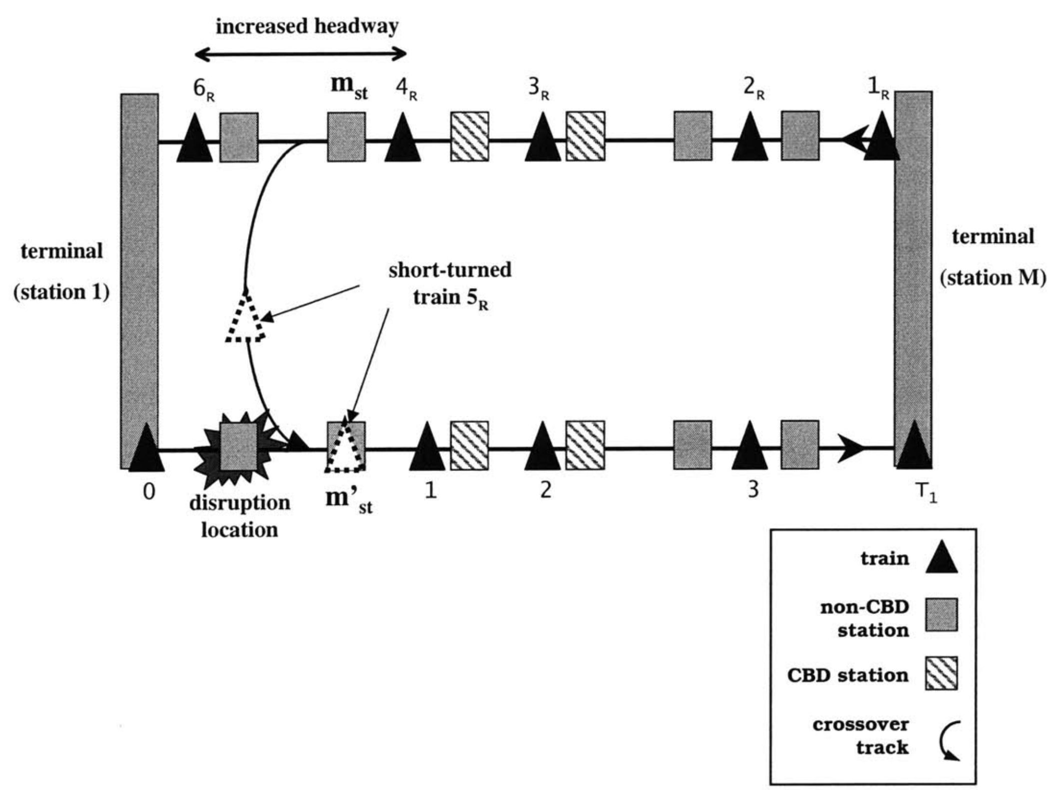

Figure 5-1: Short-turning ahead of the blockage

The blockage is located after the CBD In this case, the peak demand is in the reverse direction. If the disruption location is far enough beyond the CBD, trains behind the blockage will generally have low passenger loads (see Fig. 5-2). These trains could then be short-turned at an appropriate crossover track. This short-turn location must be located between the CBD and the disruption location, and be as close as possible to the disruption location, so that more stations are served by the short-turned trains in the reverse direction.

Short-turning here also leads to a loss of service behind the blockage. A more detailed discussion of this issue in the case of short-turning behind will be provided in a later section (Section 5.3).

In the case of short-turning behind the blockage, there are generally stations and trains located between the disruption location and the crossover track in the reverse direction (unless the blockage occurs at the terminal). Trains are thus short-turned into a sequence of trains at locations which must be carefully determined. If the disruption and the crossover track are close enough to the terminal so that all trains ahead of the blockage can be past the crossover track in the reverse direction before the short-turn operation is completed, the sequence of trains is straightforward (see Fig. 5-2). Trains are short-turned behind these trains[1]. Otherwise, the selected train sequence must balance:

- The negative impacts of *potential* holds on trains in the reverse direction: these holds could be necessary in order to safely insert the short-turned trains in the desired train sequence
- The impacts of short-turning trains on the level of service provided behind the blockage
- The benefits from the additional train service in the reverse direction and the complementary holds to even out headways

We will discuss the above tradeoffs in more detail in Section 5.3 to give guidance on how an appropriate train sequence may be identified to achieve significant overall benefits.

[1]This case of short-turning is equivalent to short-turning ahead of the blockage since trains are short-turned into the headway gap developing in front of the blocked train

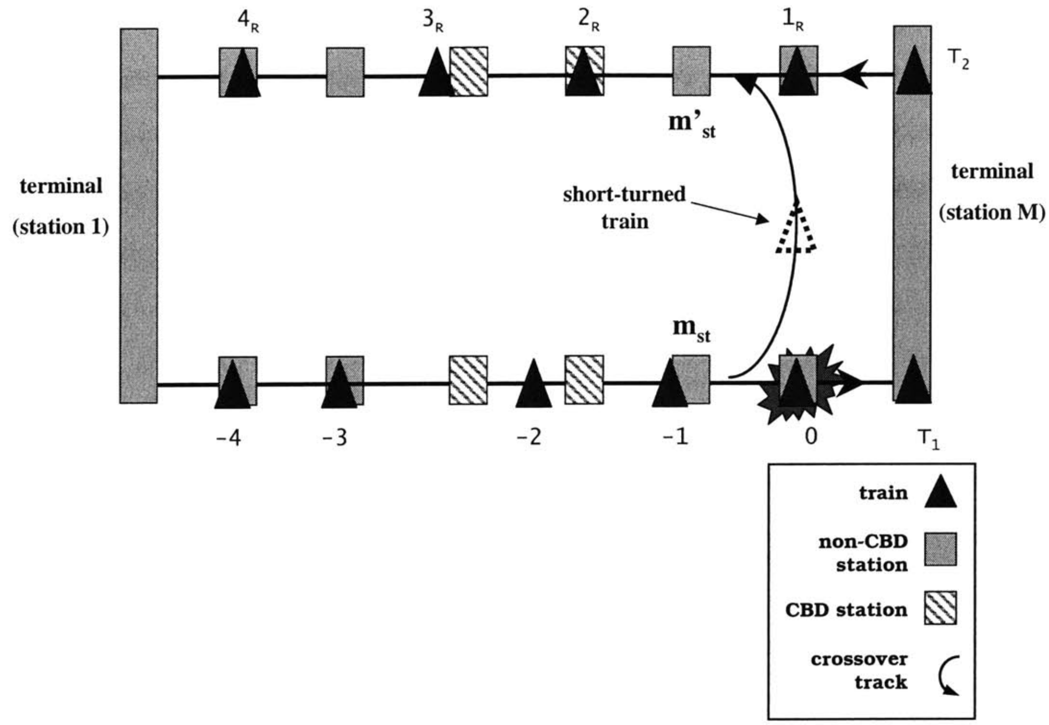

Figure 5-2: Short-turning behind the blockage

General Goals of a Short-Turning Strategy

In all cases, one of the major consequences of a blockage is the development of a service gap in front of the blocked train (the other being the accumulation of passengers both behind and ahead of the blockage). This results in increased headway variance, and greater average headways at stations located beyond the blockage.

Short-turning, if considered, must then provide additional train capacity to serve the CBD (see Section 1.3.2) *and also reduce the headway means and variances at stations*, according to Chapter 4. Adding a short-turned train allows average headways to be smaller while additional holding actions can even out the headway sequences at stations. In addition, from the results of Chapter 4, smaller holds would be required at stations served by the new trains to obtain even headway sequences.

Nevertheless, the negative impacts and the exact overall benefits of a short-turning action are different for each type of short-turn. As identified by Wilson et al. in [36], there are actually generally four groups of passengers that are affected by a short-turn decision:

- Skipped segment boarders–passengers who, if the train had not been short-turned, would have boarded it at stations outside the short-turn loop, in both direction
- Skipped segment alighters–those passengers who are dumped by a short-turned train and must await a following train in order to reach their destination
- Short-turn point boarders–those passengers waiting at the station before the crossover track and would have boarded a short-turned train had it continued
- Reverse direction passengers–those traveling to the CBD who board one of the short-turned trains

The last group benefits from a short-turn decision while the three first groups are negatively affected. Yet, depending on the type of short-turn, the benefits and levels of inconvenience experienced by each of these groups are different.

Hence, we will separately analyze in more detail each of the short-turn types and derive in each case a means to determine an efficient short-turn strategy, based on simple logical considerations and the use of the holding model presented earlier in this thesis. In each case, we will provide guidance in selecting the trains to be short-turned and the sequence of trains in the after-short-turn direction by discussing how these decisions might affect the efficiency of the short-turning strategy. The discussion provided is intended to allow a dispatcher or an automated DSS to select a small number of short-turn strategies to assess. We will also show in each of the short-turn types how the holding model developed in Chapter 3 can be used to determine the complementary holding actions (ahead of the blockage) and evaluate the efficiency of a given short-turn strategy.

We first analyze the simplest case of short-turning–ahead of the blockage– and then extend our analysis to the short-turn behind strategy.

5.2 Short-Turning Ahead the Blockage

5.2.1 Analysis

Impacts of Short-Turning Ahead

We investigate in more details here the impacts mentioned in Section 5.1.1 of a short-turn ahead decision on train operations and the level of service provided on various parts of the line. We will also assess the potential for control actions on these line segments.

At stations outside the short-turn loop in the reverse direction, train service is removed, which results in headway gaps and uneven headway sequences if no control action is taken. Yet, we note that there are a small number of these stations and that low passenger flows are located there since the short-turn occurs near the terminal (passenger flows are focused on the CBD during peak periods). Hence, there are little benefits from holding trains at these stations: the uneven headway sequence would lead to a waiting time increase for the skipped segment boarders that is likely negligible in comparison to the time savings achieved in the peak direction[2].

A similar argument–low passengers flows– holds for the negative impacts incurred by the skipped segment alighters and the short-turn point boarders. For each short-turn train, few passengers travel beyond the short-turn point and are forced to wait for another train. Additionally, as underlined in Section 1.3.2, there are a limited number of trains that can be short-turned, due to the duration of crossover operations[3] so we can conclude that the overall negative impacts of a short-turn option incurred by

[2]The model results presented by Shen [30] provides support for such a statement.

[3]For instance, in the case of the Porter Southbound disruption presented in Chapter 4, based on a six-minute short-turn operation duration and the normal inter-station running times, only three trains can be short-turned at the cross-over track located between Central and Harvard stations.

skipped segment alighters and short-turn point boarders are small, in comparison to the waiting time savings achieved in the disruption direction.

We thus conclude that there is no need for controlling the non-short-turned trains in the direction reverse to the disruption.

In the disruption direction, trains are short-turned into the gap, behind the trains located immediately ahead of the blockage (see Fig. 5-1). This additional train service reduces the gap developing in front of the blockage, and thus the average headway at stations downline from the disruption. Moreover, we know that passenger waiting time at stations is also a function of the headway sequence at stations. This implies that complementary holds might further increase the benefits of the additional train service downline from the disruption.

5.2.2 Solving the Corresponding Holding Problem

From the above description, it appears that, once a short-turn option is chosen, minimizing passenger waiting time in the system is equivalent to finding the optimal holding actions on the new train sequence in front of the disruption. We show below how trains in the system can be represented in the holding model developed in this thesis to determine the complementary holds.

We first note that trains in the reverse direction are not subject to active control actions. Indeed, the analysis of the short-turn ahead decision has shown that trains in the reverse direction need not to be held to respond to train service removal. Moreover, model results from the Porter Southbound disruption along with results presented in Shen[30] show that trains in the reverse direction are not controlled to respond to the headway gap developing in front of the blockage: for the delay durations considered, terminal holding downline from the disruption and the use of layover times are sufficient to buffer against the delay and minimize the *weighted* passenger time at stations in the reverse direction.

Therefore, only trains located in the disruption direction, at the terminal downline from the disruption and trains behind the blockage (to pick up passengers left behind) need to be represented in the system model. These trains are of course reindexed according to the new train order in the disruption direction[4].

Trains initially at stations and the terminal ahead of the blockage are represented as before, that is using their current position $m(i)$, load $L_0(i)$ and headway H_i at the beginning of the disruption (refer again to Fig. 5-1).

The headways of the short-turned trains are modified only at stations beyond the crossover track in the after-short-turn direction (i.e., for $m \geq m'_{st}$ in Fig. 5-1). Thus, we can represent the trains as sitting at station m'_{st} with no passenger load. Their initial preceding headway at this station are calculated using the estimated train departure time from this station –when no holding is applied–, based on nominal inter-station running times and the time required to short-turn a train at the specified crossover.

Trains behind the blockage are represented with the same parameters as in the holding model, except for train 0's headway H_0. Indeed, at stations $m \geq m'_{st}$, this quantity is changed since the preceding train is a new short-turned train. This implies that in all constraints involving H_0 and associated with stations $m \geq m'_{st}$, H_0 must be replaced by H'_0, the difference between the estimated departure times of train 0 and the preceding (short-turned) train 1, assuming free running conditions. This quantity might be negative since a non- delayed train 0 can arrive at platform m'_{st} before its preceding short-turned train.

The simple abovementioned changes to the train representation in the holding model can then be used to determine the complementary holds on the new train sequence in the after-short-turn direction and at the terminal downline from the disruption.

---

[4]For instance, in Fig. 5-1, the crossover track is located just ahead of the blockage, and all trains ahead of the blockage are located downline from this track. If we were to short-turn train 5_R using this track, the new order of trains ahead of the blockage would simply be $5_R, 1, 2, 3$.

5.2.3 Assessing the Short-Turn Strategy

As in the case of the holding strategy, the impacts of a short-turn strategy can be evaluated by calculating the waiting times and in-vehicle delay times associated with the sets of trains and stations impacted by both the short-turn decision and the complementary holds.

Passenger times associated with trains located downline from the disruption are evaluated at control stations m from $m(i)$ through $2M-3$. Passenger times associated with trains behind the blockage are evaluated at stations $m \geq m'_{st}$ since the short-turn strategy only affects the loads of these trains and the headway of train 0 at these stations.

Moreover, given the analysis of Section 5.2.1, one must evaluate the waiting time increase associated with skipped segment boarders and alighters, and short-turn point boarders. For a given short-turn strategy, this is derived directly from the nominal trains' headway and loads, and passenger arrival rates at stations.

5.2.4 Conclusion

In this section, it was shown that short-turning ahead is a rather simple decision to analyze. For the delay durations considered in this study, only a handful of trains can be short-turned during the duration of the blockage and these trains are simply short-turned into the gap developing in front of the blockage. This leads to only a very few possible short-turn options to consider.

It was also shown that for a given short-turn option, the associated holding problem can be easily solved –using the previously developed holding model– and the impacts of the short-turn strategy assessed.

5.3 Short-Turning Behind the Blockage

The short-turn behind decision is described in more detail here. We will highlight the main differences of these impacts as compared with a short-turn ahead decision.

In particular, it will be shown that, under some favorable conditions, short-turning behind the blockage is similar to a short-turning ahead decision but that otherwise, the short-turn decision is a much more difficult decision to make.

In the case of a short-turn behind strategy, we first note that skipped segment alighters and short-turn point boarders incur the same detrimental effects of the short-turn decision, that is increased in-platform waiting time. Nevertheless, removing trains from behind the blockage has specific consequences that we describe below.

First, the skipped segment boarders are affected by the train service removal only if they would have boarded a short-turn train at a station located between the crossover track and the blockage. At stations located downline from the blockage, passengers would board the first blocked train (train 0), assuming there is no train capacity issue at these stations (passenger arrival rates are low at these stations since the blockage is located near the terminal).

Second, and more importantly, train service removal can free platforms behind the blockage and limit the propagation of the queue of trains developing behind the blockage. If the disruption is sufficiently long, this queue could propagate to the CBD area and hinder travel to the CBD. Thus, depending on the delay duration, removing trains from behind the blockage can yield benefits (decreased in-vehicle delay time) in the disruption direction. This beneficial consequence of short-turning in the before-short-turn direction was not relevant in the case of short-turning ahead since the end of the line was located between the short-turn location and the blockage: the terminal provided an additional platform for trains to queue behind the blockage and trains might be pulled out of service to a yard at the terminal.

In the after short-turn direction, the new train sequence must achieve overall benefits from the additional train service. Nevertheless, this task is made more difficult in this case of short-turning because there is no natural gap where trains can be short-turned into. We refer here to Fig. 5-2. Trains in the reverse direction are operating with a normal service headway of four minutes, which means that either train 1_R or train T_1 might have to be held to create a gap into which train -1 could be short-turned.

Hence, it appears that the choice of the train sequence must balance:

- The cost of holding trains travelling in the other direction to the CBD
- The waiting time benefits from the additional trains in the peak direction
- The negative effects of holding short-turned trains behind the blockage to achieve the desired train sequence, as trains can queue up behind the blockage

We note that these tradeoffs are difficult to assess in general. Moreover, we recognize that, even for a given set of short-turn trains and a predetermined train sequence, the more complicated train sequence generally achieved in the after-short-turn direction does not lend itself to a simple use of the holding model to determine the optimal complementary holds to exert. One reason is that holding trains ahead of the blockage (in both directions) might affect the train sequence that can be achieved, as timing is a critical factor for more complicated train sequences. Another reason is that several trains might now be preceded by a short-turn train, which make short-turned trains difficult to represent in the holding model. Nevertheless, we describe below a case where such difficulties do not arise.

In the special case the crossover is very close to the terminal, the reverse trains can travel beyond the crossover location in the after-short-turn direction before the first short-turn can be completed (as in the instance shown in Fig. 5-2). In this case, it is most "natural" to create the gap behind the first reverse train by holding trains located between the blockage and the terminal downline from the disruption. This is because there is no cost associated with terminal holding and because trains downline from the disruption carry low passenger loads.

The above case of short-turning is similar to the short-turn ahead case since there is one gap into which trains are short-turned. For this particular type of short-turn, the holding model can be used in the same fashion as for the short-turn ahead strategy to determine the complementary holds to exert on trains ahead of the blockage. The main difference here is that the change in the initial headway H_i now applies only to the next train to be dispatched in the other direction (the one which is held to create a gap). In addition, this train's cumulative holding time at the terminal must be greater than the value necessary to create the desired gap, which can be specified straightforwardly in the holding model by adding a simple inequality.

5.4 Conclusion

In this section, we presented a general analysis of the short-turn control strategy that summarizes and extends the discussion provided in Section 1.3.2. We differentiated two types of short-turning (short-turning ahead of the blockage and short-turning behind the blockage) and we discussed the choice of the short-turned trains and the train sequence in the after-short-turn direction in each case.

It was shown that the short-turn ahead strategy is generally the simplest to assess and that the holding model developed in this thesis can be used in this case to determine the complementary holds that optimize the benefits of any given short-turn ahead decision.

The short-turn behind strategy was also discussed and it was shown that this type of short-turn is more difficult to assess as the desired train sequence is difficult to determine. It involves many tradeoffs that need to be made simultaneously and the operational constraints of this type of short-turn do not allow a simple use of the holding model to determine the complementary holds to exert on trains ahead of the blockage. Nevertheless, we presented one special case of short-turning behind that is similar to the short-turn ahead case and that can be analyzed in the same fashion.

It must be concluded from this chapter that, although each given short-turn strategy can be seen as a holding problem, a more integrated approach to the short-turning strategy is probably needed. Specifically, a model based on modified headways (as for the holding model developed in this thesis) could be developed and the difficult problem of train reordering must be addressed for this purpose. It is likely that such a model would make use of additional integer variables and thus increase the solution times. To prune branches of the solution tree and reduce solution times, methods based on simple logical considerations similar to the ones developed for the holding model could be appropriate.

Chapter 6

Summary and Conclusions

In this thesis, a deterministic holding problem for real-time disruption recovery has been developed for addressing service delays in a loop transit line. The model was applied to two disruption scenarios with different disruption locations and durations on a simplified version (non-branch) of the MBTA Red Line. Sensitivity of the optimal holding strategy was conducted with respect to two major model parameters: the finite train capacity and the value of the weight of in-vehicle delay time against in-platform waiting time. We also discussed the use of short-turning in certain disruption scenarios and proposed means to assess the efficiency of some short-turn strategies with the holding model. In this chapter, the findings are summarized and suggestions are offered for future research.

6.1 Summary and Conclusions

6.1.1 A Simple Control Strategy

In Chapter 2, we analyzed a simple train control strategy for a simplified subway system, consisting of trains operating on a unidirectional subway line, on which a disruption of a known duration occurs. The control strategy considered was the "Hold First" strategy, wherein controlled trains ahead of the blockage are held at the first station they arrive at after the disruption occurs. The model did not consider capacity constraints for trains behind the blockage and the impacts of passengers left behind by these trains.

Other assumptions were made in order to obtain a more tractable formulation of this control strategy. First, passenger arrival rates, alighting fractions and train dwell-times were assumed to be constant and station-specific. In addition, train movements between stations were supposed to be deterministic. Finally, full trains ahead of the blockage were not allowed to be held.

The "Hold First" problem was then formulated as a minimization program, subject to some physical and operational constraints. These linear constraints calculated train loads at stations and constrained full trains ahead of the blockage not to be held. They also specified that trains must be operated at a minimum safe headway. The choice of the objective function to minimize was discussed and this function was chosen as the weighted sum of total in-platform waiting time at stations and the in-vehicle delay time due to the blockage and the control actions. The resulting cost function had a quadratic form and the "Hold First" problem was thus equivalent to a constrained quadratic program with linear constraints.

This constrained quadratic program with linear constraints was then rewritten in a standard matrix form, which provided an interesting interpretation of its coefficients. First, it showed that when there are more passengers waiting for a given train than for its successor, holding the former is likely to increase the value of the cost function (and reduce it otherwise). In addition, the matrix coefficients showed that the interaction between two consecutive trains' holds contributes to a waiting time decrease for the following train at stations beyond its predecessor's control station.

A model application was solved with *MATLAB* in negligible solution time.

6.1.2 The General Train Holding Problem

In Chapter 3, we developed a more realistic (albeit deterministic) holding model for a subway loop-line ("Hold All"). Specifically, holding trains at any station or at multiple stations (in both directions of the line) was considered as well as passengers left behind by fully loaded trains.

The formulation of the general holding problem extended the objective function of the "Hold First" model to include trains located behind the blockage, at the terminal downline from the disruption and in the reverse direction. The objective function also accounted for the additional waiting time incurred by passengers left behind by full trains. Passenger times associated with each train were only evaluated from the train's current location through the end station in the reverse direction.

The problem formulation also addressed operations at the terminal downline from the disruption. Specifically, linear constraints were added to the mathematical programming formulation to model layover times and the minimum turn-around time of trains at terminal stations. Also, the dispatching schedule at terminal was also taken into account by limiting the amount of time a train's dispatching time could deviate from the schedule. The formulation obtained for the extended holding model was a mixed integer quadratic program with linear constraints. Integrality was introduced by the use of binary variables that indicated whether a train leaves passengers behind at a station.

In order to solve this type of mathematical program in times compatible with real-time implementation, a two-step solution procedure was proposed. The first step of the solution procedure finds a feasible solution to the "No Hold" control strategy and locates the stations where the train capacity issue does not arise in the absence of any holding. The second step solves the holding problem by constraining the train capacity issue not to arise at these stations, since a better solution is sought when control actions are considered. The described procedure dramatically reduced the number of free binary variables when few trains and/or few stations are affected by the train capacity issue without holding.

6.1.3 Model Application

In Chapter 4, results from the general holding model application were presented. Two problem instances on the MBTA Red Line were treated, both involving a disruption on the line during the morning peak period. The analysis of the results focused on *i)* the structure and the benefits achieved by the optimal holding strategy, *ii)* the impact of train capacity on the optimal holding strategy, *iii)* the sensitivity of the holding strategy to the cost associated with in-vehicle delay due to holding and *iv)* the viability of the resulting holding strategies for use by dispatchers.

Minimizing In-Platform Waiting Time

The general holding model was first applied to one of the disruption instances with the assumption of infinite train capacity and without considering the impact of holding on on-board passengers. The optimal holding strategy obtained yielded nearly perfectly even headways at each station across all controlled trains, which is consistent with an analytical result (Welding [29]), which states that average passenger waiting time at a given station is minimized when the variance of headways between trains is minimized. Moreover, the results from this application showed interesting properties of the optimal holding strategy, that were demonstrated mathematically or explained intuitively:

- No train is considered for (active) holding at a station located between two stations $m(i)$ and $m(i+1)$, where $m(i)$ is the first station a train i arrived at after the disruption starts
- The value of the constant headway sequence decreases at stations further down the line
- At a given station, a train's holding time is smaller than its preceding train's holding time
- For a given train traveling in a given direction, its holding time (at holding stations) is monotically decreasing

We also showed how the maximal deviation from schedule constraint could limit the evenness of headway sequences at stations.

It was concluded from this application that, under some "ideal" conditions (no consideration of the holding cost, no maximal deviation from schedule constraints, and infinite train capacities), minimizing passenger waiting time was equivalent to achieving perfectly even headway sequences at stations. Yet, the holding strategy required to achieve perfectly even headway sequences was complicated, as trains were held at multiple stations along their trip, which might be difficult to implement in real situations. The passenger waiting time was reduced in this case by 51%.

Solving the same problem with *finite train capacity* yielded a much different holding pattern with uneven headway sequences at stations. The results suggested that the headway distribution must in general have a high level of regularity to be optimal, but that this optimization goal was *constrained* by the train capacity. At stations where this constraint was binding, achieving perfectly even headway sequences did not necessarily lead to minimal waiting times since additional wait experienced by passengers left behind must be accounted for. Rather, different regular headway distributions were achieved for different groups of trains, each of these experiencing different levels of congestion. It was observed that minimizing only in-platform waiting time with finite train capacity also led to complex holding patterns. The passenger waiting time was reduced in this case by 49%.

Minimizing Total Passenger Time

In-vehicle delay was then included in the objective function and the same disruption scenario was solved for two non-zero values for the relative weight μ of in-vehicle delay against in-platform waiting time ($\mu = 0.1$ and 0.5).

From these model applications, it appeared that the optimal holding pattern becomes very simple when in-vehicle delay is accounted for. Less than ten trains were held at less than ten stations in both applications, as the holding costs incurred by on-board passengers were actually large, even for small values of μ. Moreover, it was noted that:

- Trains are preferentially held at the earliest station arrived at, in order to benefit from the impacts of early control interventions
- Terminal holding and use of layover time are also effective as they can even out headways at stations down the line without incurring in-vehicle waiting time

costs

The optimal holding strategies also resulted in uneven headway sequences at stations. Yet, comparison of in-platform time at stations with the case of infinite capacity and $\mu = 0$ –for which even headway sequences were observed– showed that the irregular headway pattern resulted in waiting times that were close to their minimal value.

Finally, comparison of the total passenger times showed significant savings (48% and 46% for $\mu = 0.1$ and $\mu = 0.5$ respectively) that are close to the ones obtained for an infinite train capacity and $\mu = 0$. This supported the idea that holding actions need not be complicated to be optimal and yield significant time savings.

An additional disruption case was also solved with a finite capacity and $\mu = 0.5$. Application results essentially confirmed the findings above. Nonetheless, passenger waiting time was reduced by "only" 19%, suggesting that the benefits of holding were limited in this disruption case and that short-turning might prove more beneficial.

Execution Time

Model applications were solved using the two-step procedure with Version 12.0 of XPRESS-MP. In all cases, the number of binary variables, which is the bottleneck of the solution procedure, was considerably reduced so that less than 20 binary variables remained at Step 2 of the solution procedure. The solution times were reduced by at least by a factor of two by using the two-step solution procedure and all were under 10 seconds (as compared to a solution time of 25 seconds obtained by O'Dell [26] for one of our disruption scenario).

6.1.4 The Short-Turn Strategy

In Chapter 5, we presented a general analysis of the short-turn control strategy that summarized and extended the discussion provided in Section 1.3.2. Two types of short-turning were defined (short-turning ahead of the blockage and short-turning behind the blockage) and we discuss the choice of the short-turned trains and the train sequence in the after-short-turn direction in each case.

It was shown that the short-turn ahead strategy is generally the simplest to assess and that the holding model developed in this thesis can be used in this case to determine the complementary holds that optimize the benefits of any given short-turn ahead decision.

The short-turn behind strategy was also discussed and it was shown that this type of short-turn is more difficult to assess as the desired train sequence is difficult to determine. It involves many tradeoffs that need to be made simultaneously and the operational constraints of this type of short-turn do not allow simple application of the holding model to determine the complementary holds to exert on trains ahead of the blockage. Nevertheless, we presented one special case of short-turning behind that is similar to short-turn ahead and that can be analyzed in the same fashion.

6.2 Recommendations for Future Research

The holding model developed in this thesis has provided us with insights into the train control problem: it has highlighted the impacts of train capacity and the cost of holding on the structure and the efficiency of optimal holding strategies. Moreover, the model shows a simple formulation and solution times were significantly reduced in comparison with those presented in Shen [30] and O'Dell [26], which could allow for real-time implementation of the model. Furthermore, this formulation of the holding model can be used for routine operations control.

Nevertheless, there is still much work to be done in developing real-time decision support systems. This thesis suggests the following areas of further research into this topic:

- First and foremost, the short-turning strategy needs to be addressed with the use of a mathematical programming formulation. Although it was shown that short-turning ahead could be probably addressed using some heuristical procedure combined with the holding model, we highlighted the complexity of the short-turning behind decision. It is suggested that model be based on modified headways, extending the holding model developed in this thesis. The difficult problem of train reordering must be addressed and means to solve in-real time the corresponding MIP should be investigated.

- Second, the model presented here is a deterministic model. The most questionable assumption made here is that the disruption duration is known with certainty. One way to relax this assumption would be to use a *stochastic programming* formulation of the train control problem, where uncertainty of the delay duration could be described by some probability distribution, based on data analysis.

  Nevertheless, when some of the data is random, then solutions and the optimal objective value to the optimization problem are themselves random. One logical way to pose the problem is to require that we make one control decision now and minimize the expected passenger time of the consequences of that decision. This is called the recourse model. We refer to Birge [7] for more detailed explanation of the recourse model as well as for a complete presentation of the theory of stochastic programming –we provide below some elements that are relevant to our topic.

  More importantly, solving a recourse problem is generally much more difficult than solving the deterministic version as evaluating expected costs implies high-dimensionality numerical integration in the solutions to mathematical programs. Yet, it is shown that when the random data is discretely distributed –which may be a plausible simplification for the delay duration parameter, the problem can be written as a large deterministic problem. The expectations can be written as finite sums, and each constraint can be duplicated for each realization of the random data. The resulting equivalent deterministic problem can be solved using any general purpose optimization package.

  The holding model presented in this thesis could be easily adapted to such a formulation and be implementable in real-time. Although the size of the model would increase, we showed that the two-step solution procedure can discard most of the binary variables that create the bottleneck in solving the corresponding MIP. Hence, it is expected that the stochastic formulation of the holding problem will have solution times that are compatible with real-time implementation of the model.

As for short-turning, an efficient deterministic formulation and a solution procedure still need to be developed before tackling a stochastic version of the problem. Nevertheless, it is believed that the approach used for the holding model would apply here as well.

Appendix A

"Hold All" Model Application Results. Harvard Square Northbound Twenty-Minute Disruption

Station Name / Station Acronym	JFK / JFK	Andrews / AND	Broadway / BRW	South Station / STA	DTX / DTX	Park Street / PKS	Charles MGH / MGH	Kendall / KEN	Central Square / CEN	Harvard Square / HAR	Porter Square / POR	Davis Square / DAV	Queuing Location / QUE
Passenger arrival rate (pax/min)	147.6	10.5	6.3	24.3	19.6	18.1	4.7	1.3	2.6	4.3	1.0	0.8	0.0
Alightings by train	0	11	15	198	272	170	56	96	45	87	13	14	0
Departing loads by train	633	657	671	572	408	286	253	162	130	57	47	37	37
Alighting fraction	0	0.02	0.02	0.30	0.48	0.42	0.20	0.38	0.28	0.67	0.23	0.29	0

Station Name / Station Acronym	Alewife / ALW	Davis Square / DAV	Porter Square / POR	Harvard Square / HAR	Central Square / CEN	Kendall / KEN	Charles MGH / MGH	Park Street / PKS	Downtown Crossing / DTX	South Station / STA	Broadway / BRW	Andrew / AND
Passenger arrival rate (pax/min)	38.7	44.3	30.1	37.3	27.2	5.6	3.7	21.2	18.3	3.6	0.5	1.3
Alightings by train	0	2	4	53	25	72	42	134	171	190	12	8
Departing loads by train	153	270	382	475	487	468	442	399	306	132	124	120
Alighting fraction	0	0.01	0.01	0.14	0.05	0.15	0.09	0.30	0.43	0.62	0.09	0.07

Table A.1: Station-specific parameters

Station	JFK	AND	BRW	STA	DTX	PKS	MGH	KEN	CEN	HAR	POR	DAV
Train -6	*											
Train -5		*										
Train -4			*									
Train -3				*								
Train -2							*					
Train -1									*			
Train 0										Blockage		
Train 1											*	
Train 2												*

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	JFK
Terminal Train T_1	*												
Terminal Train T_2	*												
Reverse Train 1_R			*										
Reverse Train 2_R					*								
Reverse Train 3_R							*						
Reverse Train 4_R									*				
Reverse Train 5_R											*		

Table A.2: Train locations

	Entering Load L_0	Headway H_0	Remaining Layover Time
Train -6	0	4	2
Train -5	633	4	2
Train -4	657	4	2
Train -3	671	4	2
Train -2	286	4	2
Train -1	162	4	2
Train 0	130	4	2
Train 1	57	4	6
Train 2	47	4	6
Train T_1	38	4	6
Train T_2	38	4	2
Train 1_R	270	4	na
Train 2_R	475	4	na
Train 3_R	468	4	na
Train 4_R	399	4	na
Train 5_R	132	4	na

Table A.3: Train loads, headways and layover times

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -3	0.0	14.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -2		0.0	16.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -1				18.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 0					20.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1						0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2							0.0	0.0	6.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_1									6.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_2									2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1_R											0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2_R													0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 3_R															0.0	0.0	0.0	0.0	0.0	0.0
Train 4_R																	0.0	0.0	0.0	0.0
Train 5_R																			0.0	0.0

Table A.4: Holding times (min): No hold, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	16.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -3	4.0	18.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -2		4.0	20.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -1				22.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train 0					24.0	24.0	24.0	24.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0
Train 1						4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 2							4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train T_1									4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train T_2									4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 1_R											4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 2_R													4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 3_R															4.0	4.0	4.0	4.0	4.0	4.0
Train 4_R																	4.0	4.0	4.0	4.0
Train 5_R																			4.0	4.0

Table A.5: Preceding departing headway: No hold, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2321	9	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -3	145	756	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -2		37	267	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -1				621	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train 0					1229	278	230	0	7732	8860	6012	7460	5447	1120	733	4233	3666	727	107	267
Train 1						8	6	0	309	354	240	298	218	45	29	169	147	29	4	11
Train 2							6	0	309	354	240	298	218	45	29	169	147	29	4	11
Train T_1									309	354	240	298	218	45	29	169	147	29	4	11
Train T_2									309	354	240	298	218	45	29	169	147	29	4	11
Train 1_R											240	298	218	45	29	169	147	29	4	11
Train 2_R													218	45	29	169	147	29	4	11
Train 3_R															29	169	147	29	4	11
Train 4_R																	147	29	4	11
Train 5_R																			4	11

Table A.6: In-platform waiting time (pax-min): No hold, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	345	490	401	252	186	70	56	41	41	77	165	224	267	472	413	383	310	214	88	82
Train -3	309	252	286	180	135	53	43	32	32	77	165	224	590	960	827	760	574	364	146	134
Train -2		199	178	137	104	43	35	26	26	77	165	763	960	960	827	760	574	364	146	134
Train -1				78	113	46	37	28	28	77	857	960	960	960	827	760	680	424	169	154
Train 0					23	110	108	96	96	773	960	960	960	960	928	918	960	914	420	393
Train 1						23	22	19	19	155	330	447	534	616	546	512	443	326	138	128
Train 2							14	13	13	155	330	447	534	616	546	512	443	326	138	128
Train T_1									0	155	330	447	534	616	546	512	443	326	138	128
Train T_2									0	155	330	447	534	616	546	512	443	326	138	128
Train 1_R											331	448	534	617	547	512	443	326	138	128
Train 2_R													548	630	557	522	450	330	140	129
Train 3_R															566	530	455	333	141	130
Train 4_R																	489	352	148	137
Train 5_R																			234	215

Table A.7: Train loads entering at stations: No hold, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	0	0	0	0	0	0	0	0	0	691	592	612	497	0	0	106	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	540	552	503	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	323	510	0	0	0	0	0	0	0
Train -3	0	0	0	0	0	0	0	0	0	0	0	0	164	0	0	0	0	0	0	0
Train -4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table A.8: Passengers left behind: No hold, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -3	0.0	14.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -2		0.0	16.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -1				18.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 0					20.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1						10.0	3.3	0.4	2.3	2.8	0.5	0.0	0.4	0.0	0.3	0.0	0.2	0.0	0.2	0.0
Train 2							6.7	2.3	4.5	2.1	1.1	0.0	0.8	0.0	0.6	0.0	0.4	0.0	0.3	0.0
Train T_1									11.0	1.4	1.6	0.0	1.1	0.0	0.8	0.0	0.7	0.0	0.5	0.0
Train T_2									4.5	0.7	2.1	0.0	1.5	0.0	1.1	0.0	0.9	0.0	0.7	0.0
Train 1_R											2.7	0.0	1.9	0.0	1.4	0.0	1.1	0.0	0.9	0.0
Train 2_R													2.3	0.0	1.7	0.0	1.3	0.0	1.0	0.0
Train 3_R															2.0	0.0	1.6	0.0	1.2	0.0
Train 4_R																	1.8	0.0	1.4	0.0
Train 5_R																			1.6	0.0

Table A.9: Holding times (min): $\mu = 0$, infinite capacity

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	16.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -3	4.0	18.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -2		4.0	20.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -1				22.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train 0					24.0	14.0	10.7	10.3	10.0	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 1						14.0	10.7	8.7	6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 2							10.7	13.0	6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train T_1									6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train T_2									6.5	7.2	6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 1_R											6.7	6.7	6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 2_R													6.3	6.3	6.0	6.0	5.8	5.8	5.6	5.6
Train 3_R															6.0	6.0	5.8	5.8	5.6	5.6
Train 4_R																	5.8	5.8	5.6	5.6
Train 5_R																			5.6	5.6

Table A.10: Preceding departing headway: $\mu = 0$, infinite capacity

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2321	9	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -3	145	756	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -2		37	267	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -1				621	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train 0					1229	95	46	0	1933	1148	668	829	538	110	66	381	306	61	8	21
Train 1						95	46	0	817	1148	668	829	538	111	66	381	306	61	8	21
Train 2							46	0	817	1148	668	829	538	111	66	381	306	61	8	21
Train T_1									817	1148	668	829	538	111	66	381	306	61	8	21
Train T_2									817	1148	668	829	538	111	66	381	306	61	8	21
Train 1_R											668	829	538	111	66	381	306	61	8	21
Train 2_R													538	111	66	381	306	61	8	21
Train 3_R															66	381	306	61	8	21
Train 4_R																	306	61	8	21
Train 5_R																			8	21

Table A.11: In-platform waiting time (pax-min): $\mu = 0$, infinite capacity

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -3	0.0	14.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -2		0.0	16.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -1				18.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 0					20.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1						10.0	3.3	0.3	2.3	1.5	0.6	0.0	0.3	1.7	0.0	0.0	0.2	0.0	0.2	0.0
Train 2							6.7	2.3	4.6	1.1	1.4	0.1	0.8	1.2	0.0	0.0	0.3	0.0	0.3	0.0
Train T_1									11.0	0.7	2.2	0.1	1.3	0.8	0.0	0.0	0.5	0.0	0.5	0.0
Train T_2									4.5	0.4	3.0	0.2	1.8	0.4	0.0	0.0	0.6	0.0	0.7	0.0
Train 1_R											3.8	0.2	2.3	0.0	0.0	0.0	0.8	0.0	0.9	0.0
Train 2_R													3.2	0.0	1.1	0.0	1.1	0.0	1.0	0.0
Train 3_R															2.1	0.0	1.4	0.0	1.2	0.0
Train 4_R																	1.8	0.0	1.4	0.0
Train 5_R																			1.6	0.0

Table A.12: Holding times (min): $\mu = 0$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	16.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -3	4.0	18.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -2		4.0	20.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -1				22.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train 0					24.0	14.0	10.7	10.3	10.0	8.5	7.9	7.9	7.6	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train 1						14.0	10.7	8.7	6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train 2							10.7	12.9	6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train T_1									6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train T_2									6.5	6.9	6.1	6.0	5.5	5.9	5.9	5.9	5.8	5.8	5.6	5.6
Train 1_R											7.8	8.0	7.2	7.2	6.1	6.1	5.8	5.8	5.6	5.6
Train 2_R													7.2	7.2	6.1	6.1	5.8	5.8	5.6	5.6
Train 3_R															6.1	6.1	5.8	5.8	5.6	5.6
Train 4_R																	5.8	5.8	5.6	5.6
Train 5_R																			5.6	5.6

Table A.13: Preceding departing headway: $\mu = 0$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2321	9	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -3	145	756	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -2		37	267	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -1				621	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train 0					1229	95	46	0	1933	1619	939	1165	786	99	64	372	306	61	8	21
Train 1						95	46	0	817	1043	555	677	414	99	64	372	306	61	8	21
Train 2							46	0	817	1043	555	677	414	98	64	372	306	61	8	21
Train T_1									817	1043	555	677	414	98	64	372	306	61	8	21
Train T_2									817	1043	555	677	414	98	64	372	306	61	8	21
Train 1_R											910	1194	701	144	69	396	306	61	8	21
Train 2_R													701	144	69	396	306	61	8	21
Train 3_R															69	396	306	61	8	21
Train 4_R																	306	61	8	21
Train 5_R																			8	21

Table A.14: In-platform waiting time (pax-min): $\mu = 0$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	345	490	401	252	186	70	56	41	41	77	165	224	267	308	273	256	221	163	69	64
Train -3	309	252	286	180	135	53	43	32	32	77	165	224	267	308	273	256	221	163	69	64
Train -2		199	178	137	104	43	35	26	26	77	165	224	267	308	273	256	221	163	69	64
Train -1				78	113	46	37	28	28	77	165	255	454	645	559	516	404	267	109	100
Train 0					23	110	98	78	78	387	761	960	960	960	849	795	682	495	209	193
Train 1						23	31	31	31	251	553	730	852	960	849	795	682	495	209	193
Train 2							14	18	18	251	553	730	852	960	849	795	682	494	209	193
Train T_1									0	251	553	730	852	960	849	795	682	494	209	193
Train T_2									0	251	553	730	852	960	849	795	682	494	209	193
Train 1_R											331	562	781	938	837	784	678	493	208	192
Train 2_R													548	716	649	613	558	424	182	169
Train 3_R															566	537	506	394	171	158
Train 4_R																	489	385	167	155
Train 5_R																			234	216

Table A.15: Train loads entering at stations: $\mu = 0$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	0	0	0	0	0	0	0	0	0	0	31	160	159	0	0	0	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table A.16: Passengers left behind: $\mu = 0$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2420	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	2831	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2		0	1768	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -1				1016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train 0					152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train 1						178	74	10	0	361	352	0	248	1361	1	1	59	1	32	1
Train 2							66	42	0	271	779	35	659	1019	2	1	118	1	65	2
Train T_1									0	180	1207	70	1071	679	2	1	178	1	98	3
Train T_2									0	90	1635	105	1482	339	1	1	238	1	132	3
Train 1_R											1235	107	1736	0	0	0	296	1	165	3
Train 2_R													1648	1	623	1	351	1	174	3
Train 3_R															1090	1	416	1	191	2
Train 4_R																	497	0	215	2
Train 5_R																			340	2

Table A.17: In-vehicle waiting time (extra ride-time): $\mu = 0$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -3	0.0	14.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -2		0.0	16.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -1				18.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 0					20.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1						10.2	0.8	0.0	4.9	1.5	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0
Train 2							4.9	0.0	9.4	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_1									12.1	0.5	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_2									5.4	0.2	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1_R											3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2_R													1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 3_R															0.0	0.5	0.0	0.0	0.0	0.0
Train 4_R																	0.0	0.0	0.0	0.0
Train 5_R																			0.0	0.0

Table A.18: Holding times (min): $\mu = 0.1$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	16.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -3	4.0	18.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -2		4.0	20.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -1				22.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train 0					24.0	13.8	12.9	12.9	10.0	8.5	8.5	8.5	8.5	8.5	8.5	8.1	8.1	8.1	8.1	8.1
Train 1						14.2	10.1	10.1	5.7	6.4	6.4	6.4	6.4	6.4	6.4	6.8	6.8	6.8	6.8	6.8
Train 2							8.9	8.9	6.2	6.5	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train T_1									6.8	7.0	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7
Train T_2									7.4	7.6	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3
Train 1_R											7.9	7.9	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train 2_R													5.8	5.8	5.8	5.3	5.3	5.3	5.3	5.3
Train 3_R															4.0	4.5	4.5	4.5	4.5	4.5
Train 4_R																	4.0	4.0	4.0	4.0
Train 5_R																			4.0	4.0

Table A.19: Preceding departing headway: $\mu = 0.1$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2321	9	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -3	145	756	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -2		37	267	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -1				621	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train 0					1229	92	67	0	1933	1617	1097	1361	994	204	134	691	598	119	17	44
Train 1						98	41	0	623	900	610	757	553	114	74	495	428	85	12	31
Train 2							32	0	741	923	570	707	516	106	69	401	347	69	10	25
Train T_1									887	1094	491	609	445	91	60	346	299	59	9	22
Train T_2									1046	1280	418	519	379	78	51	294	255	51	7	19
Train 1_R											947	1175	516	106	69	401	347	69	10	25
Train 2_R													455	94	61	298	258	51	7	19
Train 3_R															29	212	184	36	5	13
Train 4_R																	147	29	4	11
Train 5_R																			4	11

Table A.20: In-platform waiting time (pax-min): $\mu = 0.1$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	345	490	401	252	186	70	56	41	41	77	165	224	267	308	273	256	221	163	69	64
Train -3	309	252	286	180	135	53	43	32	32	77	165	224	267	308	273	256	221	163	69	64
Train -2		199	178	137	104	43	35	26	26	77	165	224	267	308	273	256	221	163	69	64
Train -1				78	113	46	37	28	28	77	165	274	495	709	614	566	438	287	116	107
Train 0					23	110	98	80	80	387	761	960	960	960	864	817	743	572	247	229
Train 1						23	31	30	30	220	500	686	828	960	852	798	704	526	225	208
Train 2							14	17	17	239	523	703	834	960	850	797	688	505	214	198
Train T_1									0	262	571	737	847	960	848	793	676	490	207	191
Train T_2									0	284	618	771	859	960	846	789	664	475	200	184
Train 1_R											331	566	783	912	809	759	662	490	209	193
Train 2_R													548	678	609	575	515	391	168	155
Train 3_R															566	530	466	348	148	137
Train 4_R																	489	352	148	137
Train 5_R																			234	215

Table A.21: Train loads entering at stations: $\mu = 0.1$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	0	0	0	0	0	0	0	0	0	0	50	184	185	0	0	0	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table A.22: Passengers left behind: $\mu = 0.1$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2420	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	2831	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2		0	1768	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -1				1016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train 0					152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train 1						181	19	0	0	316	0	0	0	0	0	258	0	0	0	0
Train 2							49	0	0	180	0	0	0	0	0	0	0	0	0	0
Train T_1									0	130	167	0	0	0	0	0	0	0	0	0
Train T_2									0	69	982	0	0	0	0	0	0	0	0	0
Train 1_R											1286	0	0	0	0	0	0	0	0	0
Train 2_R													926	0	0	0	0	0	0	0
Train 3_R															0	177	0	0	0	0
Train 4_R																	0	0	0	0
Train 5_R																			0	0

Table A.23: In-vehicle waiting time (extra ride-time): $\mu = 0.1$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -3	0.0	14.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -2		0.0	16.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train -1				18.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 0					20.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1						7.9	0.0	0.0	8.1	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2							1.7	0.0	12.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_1									11.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train T_2									5.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 1_R											2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 2_R													0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Train 3_R															0.0	0.0	0.0	0.0	0.0	0.0
Train 4_R																	0.0	0.0	0.0	0.0
Train 5_R																			0.0	0.0

Table A.24: Holding times (min): $\mu = 0.5$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	16.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -3	4.0	18.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -2		4.0	20.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -1				22.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train 0					24.0	16.1	16.1	16.1	10.0	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5	9.5
Train 1						11.9	10.2	10.2	5.8	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3
Train 2							5.7	5.7	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train T_1									6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2	6.2
Train T_2									7.7	7.7	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1	5.1
Train 1_R											6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6
Train 2_R													4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 3_R															4.0	4.0	4.0	4.0	4.0	4.0
Train 4_R																	4.0	4.0	4.0	4.0
Train 5_R																			4.0	4.0

Table A.25: Preceding departing headway: $\mu = 0.5$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2321	9	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -3	145	756	3	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -2		37	267	5	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train -1				621	9	2	2	0	77	89	60	75	54	11	7	42	37	7	1	3
Train 0					1229	126	104	0	1933	1996	1354	1681	1227	252	165	954	826	164	24	60
Train 1						68	41	0	657	889	603	749	547	112	74	425	368	73	11	27
Train 2							13	0	751	861	584	725	529	109	71	411	356	71	10	26
Train T_1									751	861	584	725	529	109	71	411	356	71	10	26
Train T_2									1148	1315	392	486	355	73	48	276	239	47	7	17
Train 1_R											655	812	593	122	80	461	399	79	12	29
Train 2_R													218	45	29	169	147	29	4	11
Train 3_R															29	169	147	29	4	11
Train 4_R																	147	29	4	11
Train 5_R																			4	11

Table A.26: In-platform waiting time (pax-min): $\mu = 0.5$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	345	490	401	252	186	70	56	41	41	77	165	224	267	308	273	256	221	163	69	64
Train -3	309	252	286	180	135	53	43	32	32	77	165	224	267	308	273	256	221	163	69	64
Train -2		199	178	137	104	43	35	26	26	77	165	224	267	308	273	256	221	163	69	64
Train -1				78	113	46	37	28	28	77	165	344	590	826	713	656	502	323	130	119
Train 0					23	110	100	84	84	387	803	960	960	960	869	826	779	618	269	250
Train 1						23	29	29	29	225	504	689	829	960	851	798	693	511	217	201
Train 2							14	15	15	241	515	697	832	960	851	797	690	508	215	199
Train T_1									0	241	515	697	832	960	851	797	690	508	215	199
Train T_2									0	298	636	783	864	960	845	787	659	469	197	182
Train 1_R											331	526	699	843	754	710	637	484	208	193
Train 2_R													548	630	557	522	450	330	140	129
Train 3_R															566	530	455	333	141	130
Train 4_R																	489	352	148	137
Train 5_R																			234	215

Table A.27: Train loads entering at stations: $\mu = 0.5$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train 0	0	0	0	0	0	0	0	0	0	0	121	220	210	0	0	0	0	0	0	0
Train -1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table A.28: Passengers left behind: $\mu = 0.5$, capacity = 960 passengers/train

	PKS	MGH	KEN	CEN	HAR	POR	DAV	QUE	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND
Train -4	2420	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	2831	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2		0	1768	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -1				1016	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train 0					152	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Train 1						140	0	0	0	113	0	0	0	0	0	0	0	0	0	0
Train 2							17	0	0	0	0	0	0	0	0	0	0	0	0	0
Train T_1									0	0	0	0	0	0	0	0	0	0	0	0
Train T_2									0	0	0	0	0	0	0	0	0	0	0	0
Train 1_R											849	0	0	0	0	0	0	0	0	0
Train 2_R													0	0	0	0	0	0	0	0
Train 3_R															0	0	0	0	0	0
Train 4_R																	0	0	0	0
Train 5_R																			0	0

Table A.29: In-vehicle waiting time (extra ride-time): $\mu = 0.5$, capacity $= 960$ passengers/train

Appendix B

"Hold All" Model Application Results. Porter Square Southbound Fifteen-Minute Disruption

Station Name / Station Acronym	Alewife ALW	Davis Square DAV	Porter Square POR	Harvard Square HAR	Central Square CEN	Kendall KEN	Charles MGH MGH	Park Street PKS	Downtown Crossing DTX	South Station STA	Broadway BRW	Andrew AND
Passenger arrival rate (pax/min)	38.7	44.3	30.1	37.3	27.2	5.6	3.7	21.2	18.3	3.6	0.5	1.3
Alightings by train	0	2	4	53	25	72	42	134	171	190	12	8
Departing loads by train	153	270	382	475	487	468	442	399	306	132	124	120
Alighting fraction	1.00	0.01	0.01	0.14	0.05	0.15	0.09	0.30	0.43	0.62	0.09	0.07

Station Name / Station Acronym	JFK JFK	Andrews AND	Broadway BRW	South Station STA	DTX DTX	Park Street PKS	Charles MGH MGH	Kendall KEN	Central Square CEN	Harvard Square HAR	Porter Square POR	Davis Square DAV
Passenger arrival rate (pax/min)	147.6	10.5	6.3	24.3	19.6	18.1	4.7	1.3	2.6	4.3	1.0	0.8
Alightings by train	0	11	15	198	272	170	56	96	45	87	13	14
Departing loads by train	633	657	671	572	408	286	253	162	130	57	47	37
Alighting fraction	0	0.02	0.02	0.30	0.48	0.42	0.20	0.38	0.28	0.67	0.23	0.29

Table B.1: Station-specific parameters.

Note: Train 8_R is modeled as train -3 sitting at the Alewife terminal with a current headway of 0 minute and a remaining layover time of 2 minutse at Alewife. These values are indicated in the tables within parenthesis.

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK
Train - 2	*													
Train - 1		*												
Train 0			Blockage											
Train 1					*									
Train 2							*							
Train 3									*					
Train 4											*			
Terminal Train T_1														*
Terminal Train T_2														*

Station	AND	BRW	STA	DTX	PKS	MGH	KEN	CEN	HAR	POR	DAV
Reverse Train 1_R	*										
Reverse Train 2_R		*									
Reverse Train 3_R			*								
Reverse Train 4_R						*					
Reverse Train 5_R								*			
Reverse Train 6_R									*		
Reverse Train 7_R										*	
Reverse Train 8_R (-3)											*

Table B.2: Train locations

	Entering Load L_0	Headway H_0	Remaining Layover Time
Train - 2	0	0	2
Train - 1	153	4	2
Train 0	270	4	2
Train 1	475	4	6
Train 2	468	4	6
Train 3	399	4	6
Train 4	132	4	6
Train T_1	0	4	6
Train T_2	590	4	2
Train 1_R	633	4	na
Train 2_R	657	4	na
Train 3_R	671	4	na
Train 4_R	286	4	na
Train 5_R	162	4	na
Train 6_R	130	4	na
Train 7_R	57	4	na
Train -3	47(0)	4(0)	2

Table B.3: Train loads, headways and layover times

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	17.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train -2	15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train -1		13.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train 0			15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train 1					0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0
Train 2							0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0
Train 3									0.0	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0
Train 4											0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0
Train T_1														6.0	0.0	0.0	0.0	0.0
Train T_2														2.0	0.0	0.0	0.0	0.0
Train 1_R															0.0	0.0	0.0	0.0
Train 2_R																0.0	0.0	0.0
Train 3_R																	0.0	0.0

Table B.4: Holding times (min): No hold, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -2	15.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -1		17.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train 0			19.0	19.0	19.0	19.0	19.0	19.0	19.0	19.0	19.0	19.0	19.0	15.0	15.0	15.0	15.0	15.0
Train 1					4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 2					4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 3							4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train 4											4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
Train T_1														4.0	4.0	4.0	4.0	4.0
Train T_2														4.0	4.0	4.0	4.0	4.0
Train 1_R															4.0	4.0	4.0	4.0
Train 2_R																4.0	4.0	4.0
Train 3_R																	4.0	4.0

Table B.5: Preceding departing headway: No hold, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	77	89	60	75	54	11	7	42	37	7	1	3	0	295	21	12	48	39
Train -2	4349	89	60	75	54	11	7	42	37	7	1	3	0	295	21	13	48	39
Train -1		6401	60	75	54	11	7	42	37	7	1	3	0	295	21	13	48	39
Train 0			5426	6733	4916	1011	662	3821	3309	656	96	241	0	16605	1181	709	2730	2205
Train 1					218	45	29	169	147	29	4	11	0	1181	84	50	194	157
Train 2							29	169	147	29	4	11	0	1181	84	50	194	157
Train 3									146	29	4	11	0	4723	336	202	776	627
Train 4											4	11	0.0	7380	525	315	1213	980
Train T_1														1181	84	50	194	156
Train T_2														1181	84	50	194	157
Train 1_R															84	50	194	157
Train 2_R																50	194	157
Train 3_R																	194	157

Table B.6: In-platform waiting time: No hold, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	0	77	165	404	404	920	793	729	553	352	141	129	123	123	295	378	445	360
Train -2	0	580	663	960	960	960	827	760	574	364	146	134	127	127	884	960	960	721
Train -1		153	905	960	960	960	827	760	653	409	163	149	141	141	960	960	960	797
Train 0			270	960	960	960	922	909	960	895	409	382	381	381	960	960	960	960
Train 1					475	560	499	468	413	308	132	122	119	119	590	621	633	540
Train 2							468	441	393	297	128	118	115	115	590	621	633	540
Train 3									399	301	129	119	116	116	590	621	633	540
Train 4											132	122	119	119	590	621	633	540
Terminal Train 1														0	590	621	633	540
Terminal Train 2														0	590	621	633	540
Reverse Train 1															633	662	674	569
Reverse Train 2																657	669	565
Reverse Train 3																	671	567

Table B.7: Train loads entering at stations: No hold, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train 0	0	0	0	470	469	0	0	79	0	0	0	0	0	1254	138	75	76	0
Train -1	0	0	0	406	476	0	0	0	0	0	0	0	0	589	140	69	0	0
Train -2	0	0	0	137	482	0	0	0	0	0	0	0	0	0	68	62	0	0
Train -3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table B.8: Passengers left behind: No hold, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	17.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train -2	15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train -1		13.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train 0			15.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0
Train 1					7.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.9	0.0	0.0	0.0	0.0
Train 2							0.0	2.3	0.0	0.0	0.0	0.0	0.0	11.4	0.0	0.0	0.0	0.0
Train 3									0.0	0.7	0.0	0.0	0.0	11.3	0.0	0.0	0.0	0.0
Train 4											0.0	0.0	0.0	10.3	0.0	0.0	0.0	0.0
Train T_1														8.7	0.0	0.0	0.0	0.0
Train T_1														3.4	0.0	0.0	0.0	0.0
Train 1_R															0.0	0.0	0.0	0.0
Train 2_R																0.0	0.0	0.0
Train 3_R																	0.0	0.0

Table B.9: Holding times (min): $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -2	15.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train -1		17.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Train 0			19.0	19.0	11.6	11.6	11.6	11.6	11.6	11.6	11.6	11.6	11.6	5.7	5.7	5.7	5.7	5.7
Train 1					11.4	11.4	11.4	9.1	9.1	9.1	9.1	9.1	9.1	5.7	5.7	5.7	5.7	5.7
Train 2							4.0	6.3	6.3	5.6	5.6	5.6	5.6	5.7	5.7	5.7	5.7	5.7
Train 3									4.0	4.7	4.7	4.7	4.7	5.7	5.7	5.7	5.7	5.7
Train 4											4.0	4.0	4.0	5.6	5.6	5.6	5.6	5.6
Train T_1														5.4	5.4	5.4	5.4	5.4
Train T_2														5.4	5.4	5.4	5.4	5.4
Train 1_R															4.0	4.0	4.0	4.0
Train 2_R																4.0	4.0	4.0
Train 3_R																	4.0	4.0

Table B.10: Preceding departing headway: $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	77.3	88.6	60.1	74.6	54.5	11.2	7.3	42.3	36.7	7.3	1.1	2.7	0.0	295.2	21.0	12.6	48.5	39.2
Train -2	4349.2	88.6	60.1	74.6	54.5	11.2	7.3	42.3	36.7	7.3	1.1	2.7	0.0	295.2	21.0	12.6	48.5	39.2
Train -1		6401.3	60.1	74.6	54.5	11.2	7.3	42.3	36.7	7.3	1.1	2.7	0.0	295.2	21.0	12.6	48.5	39.2
Train 0			5425.8	6732.6	1826.4	375.6	245.9	1419.5	1229.3	243.7	35.8	89.4	0.0	2376.9	169.1	101.5	390.8	315.6
Train 1					1775.4	365.1	239.0	880.1	762.2	151.1	22.2	55.4	0.0	2376.9	169.1	101.5	390.8	315.6
Train 2							29.3	420.0	363.7	56.5	8.3	20.7	0.0	2376.9	169.1	101.5	390.8	315.6
Train 3									146.6	40.5	5.9	14.9	0.0	2376.9	169.1	101.5	390.8	315.6
Train 4											4.3	10.7	0.0	2296.4	163.4	98.0	377.5	304.9
Train T_1														2120.7	150.9	90.5	348.7	281.6
Train T_2														2120.7	150.9	90.5	348.7	281.6
Train 1_R															84.0	50.4	194.1	156.8
Train 2_R																50.4	194.1	156.8
Train 3_R																	194.1	156.8

Table B.11: In-platform waiting time: $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train -3	0	77	165	224	404	718	622	573	443	289	117	108	103	103	295	310	317	270
Train -2	0	580	663	716	960	960	827	760	574	364	146	134	127	127	295	310	317	270
Train -1		153	905	956	960	960	827	760	574	364	146	134	127	127	295	310	317	270
Train 0			270	838	960	960	881	844	836	689	304	283	278	278	838	880	899	767
Train 1					475	781	732	711	703	578	255	237	233	233	833	875	893	762
Train 2							468	441	445	372	162	150	147	147	833	875	893	762
Train 3									399	301	132	122	120	120	833	875	893	762
Train 4											132	122	119	119	830	873	891	760
Train T_1														0	798	839	856	731
Train T_2														0	798	839	856	731
Train 1_R															633	662	674	569
Train 2_R																657	669	565
Train 3_R																	671	567

Table B.12: Train loads entering at stations: $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train 0	0	0	0	470	267	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -1	0	0	0	406	274	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -2	0	0	0	137	280	0	0	0	0	0	0	0	0	0	0	0	0	0
Train -3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table B.13: Passengers left behind: $\mu = 0.5$, capacity = 960 passengers/train

Station	ALW	DAV	POR	HAR	CEN	KEN	MGH	PKS	DTX	STA	BRW	AND	QUE	JFK	AND	BRW	STA	DTX
Train 0			3996.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Train 1					3340.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Train 2								0.00	705.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Train 3									0.00	82.72	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Train 4											0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Train T_1														0.00	0.00	0.00	0.00	0.00
Train T_2														0.00	0.00	0.00	0.00	0.00
Train 1_R															0.00	0.00	0.00	0.00
Train 2_R																0.00	0.00	0.00
Train 3_R																	0.00	0.00

Table B.14: In-vehicle waiting time (extra ride-time): $\mu = 0.5$, capacity = 960 passengers/train

Bibliography

[1] M. Abkowitz and I. Engelstein. Optimal Control of Headway Variation on Transit Routes. *Journal of Advanced Transportation*, 20(1):73–88, 1986.

[2] Dash Associates. *XPRESS-MP's User Guide/Reference Manual*, 1999.

[3] C. Audet, P. Hansen, B. Jaumard, and G. Savard. A branch and cut algorithm for non-convex quadratically constrained quadratic programming. *Mathematical Programming*, 87(1):131–152, 2000.

[4] A. Barnett. On Controlling Randomness in Transit Operations. *Transportation Science*, 8(2):102–116, 1974.

[5] A. Barnett. Control Strategies for Transport Systems with Nonlinear Waiting Costs. *Transportation Science*, 12(2):119–136, 1978.

[6] C. Barnhart, E. Johnson, G. Nemhauser, M. Savelsbergh, and P.H. Vance. Branch-and-Price: Column Generation for Huge Integer Programs. *Operations Research*, 46:316–329, 1998.

[7] J.R. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Springer Series in Operations Research. Springer-Verlag, 1997.

[8] A. Caprara and M. Fischetti. Branch-and-cut algorithms. In M. Dell'Amico, F. Maffioli, and S. Martello, editors, *Annotated bibliographies in combinatorial optimization*, pages 45–63. John Wiley & Sons, Chichester, 1997.

[9] V. Chandru and M.R. Rao. Integer Programming. In M.J. Atallah, editor, *Algorithms and Theory of Computation Handbook.* CRC Press, 1999.

[10] V. Chandru and M.R. Rao. Linear Programming. In M.J. Atallah, editor, *Algorithms and Theory of Computation Handbook.* CRC Press, 1999.

[11] X. Eberlein. *Real-Time Control Strategies in Transit Operations: Models and Analysis.* PhD thesis, Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, 1995.

[12] X. Eberlein, N.H.M. Wilson, and D. Bernstein. Modeling Real-Time Control Strategies in Public Transit Operations. In N.H.M. Wilson, editor, *Lecture Note in Economics and Mathematical Systems*, 471, pages 325–346. Springer-Verlag, Berlin, Heidelberg, 1999.

[13] C. A. Floudas and V. Visweswaran. Quadratic optimization. In *Handbook of Global Optimization*, pages 217–269. Kluwer Acad. Publ., Dordrecht, 1995.

[14] P. Furth. A Headway Control Strategy for Recovering From Transit Vehicle Delays. In *Transportation Congress: Civil Engineers–Key to the World Infrastructure*, volume 2, pages 2032–2039. Transportation Research Board, 1995.

[15] Z. Gu, G. L. Nemhauser, and M. W. P. Savelsbergh. Lifted cover inequalities for 0–1 integer programs: computation. *INFORMS Journal on Computing*, 10:427–438, 1998.

[16] D.E. Heimburger, A.Y. Herzenberg, and N.H.M. Wilson. Using Simple Simulation Models in The Operational Analysis of Real Transit Lines: A Case Study of the MBTA Red Line. *Transportation Research Record*, 1677:21–30, 1999.

[17] K.L. Hoffman. Combinatorial Optimization: Current Successes and Directions for the Future. *Journal of Computational and Applied Mathematics*, 124:341–360, 2000.

[18] M. Jünger, G. Reinelt, and S. Thienel. Practical Problem Solving with Cutting Plane Algorithms in Combinatorial Optimization. In W. Cook, L. Lovasz, and P. Seymour, editors, *Combinatorial optimization. Papers from the DIMACS*

special year. Papers from workshops held at DIMACS at Rutgers University, volume 20, pages 111–152. 1995.

[19] V. Lefebvre and P. Feautrier. Automatic storage management for parallel programs. *Parallel Computing*, 24(3–4):649–671, 1998.

[20] T.-M.. Lin. Dwell-Time Relationships for Urban Rail Systems. Master's thesis, Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, 1990.

[21] T.-M. Lin and N.H.M. Wilson. Dwell-Time Relationships for Light Rail Systems. *Transportation Research Record*, 1361:287–295, 1992.

[22] A. Lucena and J. Beasley. Branch and cut algorithms. In *Advances in Linear and Integer Programming*, Oxford Lecture Series in Mathematics and its Applications. 1996.

[23] R. Macchi. Expressing Vehicles on the MBTA Green Line. Master's thesis, Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, 1989.

[24] G. Nemhauser, M. Savelsbergh, and G. Sigismondi. MINTO, a Mixed INTeger Optimizer. *Operations Research Letters*, 15:47–58, 1994.

[25] G. Nemhauser and L. Wolsey. *Integer and Combinatorial Optimization*. Wiley Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, 1988.

[26] S. O'Dell. Optimal Control Strategies for a Rail Transit Line. Master's thesis, Massachusetts Institute of Technology, Operations Research Center, 1997.

[27] S. O'Dell and N.H.M. Wilson. Optimal Real-Time Control Strategies for Rail Transit Operations during Disruption. In N.H.M. Wilson, editor, *Lecture Note in Economics and Mathematical Systems*, volume 471, pages 299–323. Springer-Verlag, Berlin, Heidelberg, 1999.

[28] E. Osuna and G. Newell. Control Strategies for an Idealized Public Tranportation System. *Transportation Science*, 6:52–72, 1972.

[29] Welding P. The Instability of Close Interval Service. *Operations Research Quarterly*, 8(3):133–148, 1957.

[30] S. Shen. Integrated Real-Time Disruption Recovery Strategies: A Model for Rail Transit Systems. Master's thesis, Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, jan 2000.

[31] W. Song. A Dispatching Control Model for Rail Transit Line. Master's thesis, Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, 1998.

[32] Central Transportation Planning Staff. *1997 Passengers Counts: MBTA Rapid Rapid Transit.* Boston, MA, 1997.

[33] S. Thienel. *ABACUS: A Branch-And-CUt System.* PhD thesis, Universitat zu Koln, Germany, Department of Applied Mathematics, 1995.

[34] M. Turnquist. Real-Time Control for Improving Transit Level-of-Service. In *First International Conference on Applications of Advanced Technologies in Transportation Engineering*, volume 5, pages 217–222, San Diego, California, feb 1989.

[35] M. Turnquist and S. Blume. Evaluating Potential Effectiveness of Headway Control Strategies for Transit Systems. *Transportation Research Record*, 746:25–29, 1980.

[36] N.H.M. Wilson, R. Macchi, R. Fellows, and A. Deckoff. Improving Service on the MBTA Green Line through Better Operations Control. *Transportation Research Record*, 1361:296–304, 1992.

[37] L. Wolsey. *Integer Programming.* Wiley Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, 1998.

[38] S. Wright. On the convergence of the Newton/log-barrier method. Technical Report Preprint ANL/MCS-P681-0897, Argonne National Laboratory, Argonne, Illinois 60439, 1997.