Optimal Control Strategies for a Rail Transit Line

by

Susan W. O'Dell

B.S., Cornell University (1986)
M.S., University of California, Berkeley (1987)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Science in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1997

Signature of Author . . . .
Department of Electrical Engineering and Computer Science
23 May 1997

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

Accepted by . . . .
Thomas L. Magnanti
George Eastman Professor of Management Science
Professor of Electrical Engineering and Computer Science
Codirector, Operations Research Center

Optimal Control Strategies for a Rail Transit Line

by

Susan W. O'Dell

Submitted to the Department of Electrical Engineering and Computer Science
on 23 May 1997, in partial fulfillment of the
requirements for the degree of
Master of Science in Operations Research

Abstract

Rail transit systems are frequently subject to disruptions which temporarily block traffic, leading to increased passenger waiting times and overcrowding of trains. Even a short-term blockage (10-15 minutes) can affect the level of service throughout the line for long after the blockage has been cleared. Transit agencies employ various control strategies such as short-turning and holding of trains to mitigate the effects of these disruptions. Unfortunately, it is extremely difficult for a controller to determine, in real time, the best actions to take from a system-wide perspective.

In this thesis, we focus on the development of a real-time decision support system for rail transit operations. We study the use of holding and short-turning strategies to minimize passenger waiting time following a disruption. We consider blockages of relatively short duration, for which an alternate operations plan is unwarranted. First, we analyze holding and short-turning strategies using a simplified system model for which we obtain closed-form results; these results provide insight into the behavior of a more realistic system. Next, we develop a more realistic, but deterministic, generalized model of a rail system. We present linear and mixed integer programming formulations for several holding and short-turning strategies.

We apply the generalized model to four problem instances based on the MBTA Red Line. We demonstrate that passenger waiting time can be significantly reduced (on the order of 15-50%) by applying the set of short-turning and holding controls which are output by the model. We also show that the majority of benefits can be realized by applying only limited control actions on a small set of trains.

We were able to determine optimal control strategies using our generalized system formulations, typically in less than 30 seconds. We believe that this thesis represents a significant step towards the development of a real-time decision support system for rail transit systems .

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

Acknowledgements

I would like to thank Nigel Wilson for his guidance and encouragement, and the Operations staff of the Massachusetts Bay Transit Authority for answering questions and providing data. I would also like to thank Doug Heimburger for developing a Red Line simulation tool, Guglielmo Guastalla for helping with software, and my friends at the ORC for being there. Finally, I would like to thank John, for everything.

Contents

Chapter 1

Introduction

1.1 Motivation and Statement of Problem

When a rail transit system is operating close to capacity, a relatively minor service disruption can lead to serious degradation in system performance if appropriate control actions are not taken immediately. Examples of disruption sources include a faulty door which won't close, a malfunctioning signal, or a medical emergency on board. It is unlikely that any of these incidents would delay a train at the point of occurrence for more than 10-15 minutes; however, operations at stations both ahead of and behind the disruption location may be negatively affected long after the initial source of the disruption has been remedied. The dependence of a train's dwell time on its preceding headway causes the long headways ahead of the blockage to lengthen further, while the short headways following the blockage are further shortened. Thus, the problem is amplified in an uncontrolled, or poorly controlled, system.

It is extremely difficult, if not impossible, for even an experienced controller to assess the situation and quickly determine the best actions from a system-wide perspective. Fortunately, recent advances in automatic vehicle identification, control, and information technologies make it feasible to design computer-based real-time decision support systems for rail transit operations. This thesis focuses on the development of such a system for control of a rail transit line during and following minor service disruptions.

Our work is partly motivated by the development of a new Operations Control System for the Massachusetts Bay Transit Authority (MBTA) rail system[10]. We use the MBTA Red

Line as a case study; however, the methods we develop can be applied to any rail line for which train location information is available centrally. Furthermore, our formulations can be easily modified for use in routine operations control.

We study train operations during, and immediately following, a disruption which temporarily blocks traffic in one or both directions. We address the incidents of relatively short duration (i.e. 10 to 20 minutes) which occur frequently in many transit systems. We do not consider longer incidents which would necessitate a systematic re-routing of trains according to an alternate operations plan. Specifically, we do not consider disruptions for which single-track, shuttle, or substitute bus operations would be warranted.

We investigate two of the most commonly used control strategies: holding and short-turning. Our primary objective is the minimization of passenger waiting time. Although the use of holding to minimize waiting time may seem counterintuitive, it has significant benefits. Consider a train with a preceding headway of 2 minutes; people boarding the train have an average (expected) waiting time of 1 minute. Suppose that this train has a following headway of 6 minutes. Then, we expect three times as many people to board the following train, with an average waiting time of 3 minutes. Thus, the average wait for the two trains is $(1/4)(1) + (3/4)(3) = 2.5$ minutes. However, if we hold the first train for 2 minutes, then each train has a preceding headway of 4 minutes, and the average waiting time for the two trains is only 2 minutes.

Since we are analyzing real-time control strategies, operating costs such as wages and benefits, electricity, and equipment depreciation are virtually fixed. Therefore, we selected the minimization of passenger waiting time as our objective, since passenger satisfaction greatly affects future ridership.

We first analyze the holding and short-turning strategies using a simple, idealized system model for which we obtain closed-form results, in order to gain a better understanding of the problem. We then eliminate many simplifying assumptions and develop mathematical programming formulations for a more realistic, generalized model. We test our model on several problem instances using data from the MBTA Red Line.

1.2 Review of Existing Work

There has been little prior work towards developing real-time decision support systems for control of rail transit operations, particularly with regard to service disruptions. Some of the earliest work in this area was done by Barnett [2], [3]. In his 1974 paper, Barnett considered the use of holding to control headway variance, and presented an algorithm for constructing an optimal dispatching strategy from a single control point on a transit line. His objective was the minimization of both passenger and on-board time. Coincidentally, Barnett applied his algorithm to operations data from the MBTA Red Line. Since real-time vehicle location information was unavailable at the time this work was done, vehicle delays were approximated by a two-point discrete probability distribution. Barnett's application of the concepts of random incidence and the effect of variance on mean waiting times inspired our work on the simplified system in Chapter 2.

In the 1980's, Araya et al [1] developed a method for generating optimal schedules for online train traffic in disturbed situations. Their objective was to minimize the sum of schedule deviations, to which end they used speed variation, holding, and order changes (overtaking). The problem was formulated as a 0-1 mixed integer program; an expert system provided an initial solution for the branch-and-bound. This work was the starting point for the generalized system formulations we develop in Chapter 3. The project was continued in the late 1980's by Komaya and Fukuda [8]. At this time, however, the researchers believed that mathematical programming techniques could not be applied to problems on real railway networks, because of the computational demands of solving what is "essentially a large scale combinatorial problem" with " a lot of complicated criteria". Instead, they attempted to duplicate a skilled dispatcher's problem solving process by combining simulation and "if/then" rules. Since the 1980's, there have been significant improvements in MIP solution techniques and, most dramatically, the cost, speed, and capacity of computers.

The existing work most closely related to ours was done by Eberlein in her 1995 Ph.D. dissertation [7]. Eberlein studied the use of dead-heading, expressing, and holding strategies to minimize passenger waiting time during routine operations. She formulated models for two types of systems: a simplified, abstract one for which properties and solutions could be studied analytically, and a more general one for which the simplifying assumptions were removed. In addition to using Eberlein's two system approach, we adopted her idea of considering a limited impact set of trains and stations.

Despite similarities, our work departs from Eberlein's in several important ways. We consider problems in which the vehicle order changes. We study control strategies during service disruptions, for which the capacity of trains is an important consideration. Therefore, we include capacity constraints, which are largely ignored by Eberlein, and the effect of crowding. Finally, Eberlein's quadratic objective function and (assumed) non-linear constraints resulted in non-linear integer programs for which she developed heuristics. We formulate the capacitated holding and short-turning problems in a linear form which can be solved efficiently using a commercial optimization package.

This brief review has summarized the prior work which most influenced our model development and problem formulations.

1.3 Thesis Contents

In Chapter 2, we analyze holding and short-turning strategies using a simple, idealized model of a rail system. We compare the passenger waiting time and on-board time for the different strategies. In Chapter 3, we develop a more realistic, generalized system model. We give linear and mixed integer programming formulations for several holding and short-turning problems, for a transit system which may include branches. In Chapter 4, we apply the generalized system model to the MBTA Red Line. We test our formulations on four problem instances, and compare the effectiveness of various control strategies. We discuss several implementation issues. Finally, in Chapter 5 we summarize our findings and offer suggestions for future research.

Chapter 2

Analysis of a Simplified System

2.1 Description

2.1.1 Assumptions

In order to gain a better understanding of various control strategies, we first analyze the strategies using a simple, idealized model of a rail system. This simplified system, System S, consists of a set of trains operating on a one-way loop line at even headways prior to a disruption which temporarily blocks traffic. The following assumptions are used in System S:

- The passenger arrival rate is constant and equal at each station.
- The dwell time at each station is constant.
- The passenger load on each train is constant. Thus, the number of passengers who board at each station is equal to the number of passengers who alight at each station, when trains are operating at even headways.
- The train capacity is not a binding constraint. This assumption is reasonable for some systems during non-peak hours.

We will consider two bounding cases for System S:

Case 1: We have "sufficiently long" layovers at the terminal station (between round trips), so that uneven headways resulting from a disruption may be evened out when the trains reach the terminal.

Case 2: There are no layovers at the terminal station, so that the headway distribution following the disruption remains in effect until the operating plan changes.

2.1.2 Notation

The following notation is used in the analysis of System S. Several of the variables are illustrated in Figure 2-1. The "blocked train" refers to a disabled train, or the train immediately behind a track or station blockage.

H	=	standard operating headway
N	=	total number of stations in loop line
N_A	=	number of stations ahead of the blockage (from the blockage to the terminal)
L	=	passenger load on each train
A	=	passenger arrival rate at each station
M	=	number of round trips per train until operating plan changes
T	=	total number of trains in the system
F	=	expected number of stations between trains, when trains are operating at standard headways ($F = N/T$)
K	=	blockage duration factor (duration of blockage = KH)
ΔW	=	increase in passenger waiting time due to blockage
ΔO	=	increase in passenger on-board time due to blockage
n_a	=	number of trains held ahead of the blocked train
n_b	=	number of trains held behind the blocked train
N_{AH}	=	number of stations between the blockage location and the first held train
N'	=	$N_A + 1$ for the case where we have sufficient layovers between round trips
	=	NM for the case where we have no layovers between round trips

2.2 The System S Holding Problem

Consider a short-term blockage on one track, which leads to a delay of K times the standard operating headway H at that location. The magnitude of the resulting increases in passenger waiting time and passenger on-board time will depend on the control strategy used. An analysis of four different holding strategies is given below.

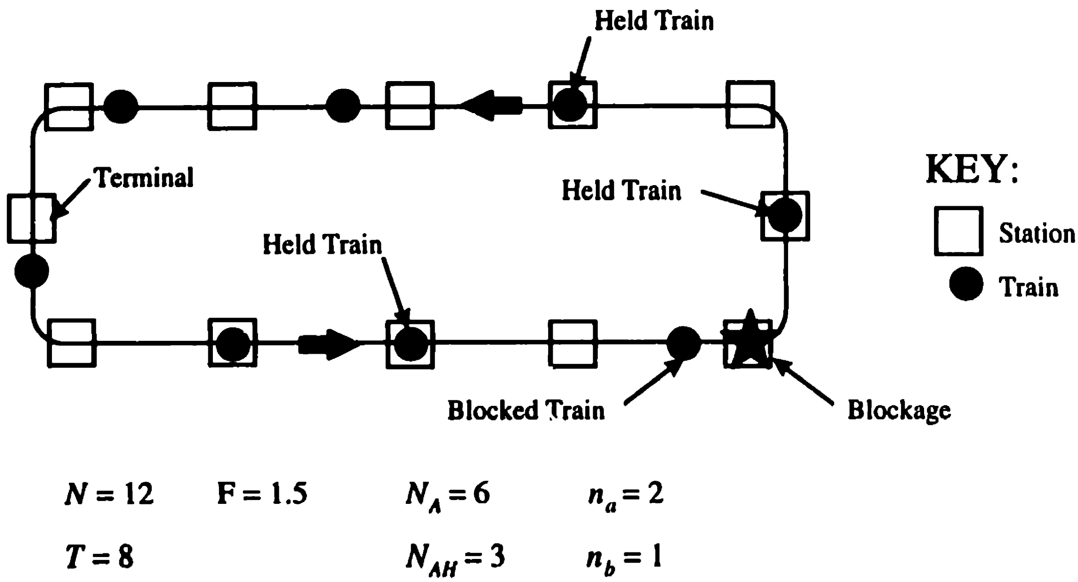


Figure 2-1: System S Notation

2.2.1 Strategy I

In this strategy, each train in the system is held for a time equal to KH at the next station the train reaches after the blockage occurs. The distance between trains remains constant, but at each station there is one headway of $(H + KH)$ followed by normal headways of H between consecutive train departures. For either Case 1 (with layovers), or Case 2 (without layovers), the increases in passenger waiting time and on-board time are:

$$\Delta W_{Strat.I} = NAH^2\frac{K(K+1)}{2} \quad (2.1)$$

$$\Delta O_{Strat.I} = TKHL \quad (2.2)$$

2.2.2 Strategy II

In this strategy, n_a trains are held in front of, and n_b trains are held behind, the blocked train. The holding time for each train is determined so that consecutive headways at stations are as even as possible for a given n_a and n_b. Each train is held at the next station it reaches after the disruption occurs. For example, suppose that $n_a = n_b = 3$, so that there are $(n_a + 1 + n_b) = 7$ held or blocked trains. The blocked (fourth) train is "held" for KH, the third and fifth trains are held for $3KH/4$, the second and sixth trains are held for $2KH/4$, and the first and seventh trains are held for $KH/4$. No other trains are held or blocked. The consecutive departure headways at stations *beyond* the current locations of the 7 trains are $...H, H, (H+KH/4), (H+KH/4), (H+KH/4), (H+KH/4), (H-KH/4), (H-KH/4), (H-KH/4), (H-KH/4), H, H....$ The headways are different at the stations located *within* the set of 7 trains at the time the blockage occurs; these headways depend on the relative location of the stations and the trains. The consecutive departure headways at each station are given in the table below, for the case with $n_a = n_b = 3$:

Location of Station:	Consecutive Departure Headways:							
Before Train 1	$H+\frac{KH}{4}$	$H+\frac{KH}{4}$	$H+\frac{KH}{4}$	$H+\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$
Between Tr 1 & 2	H	$H+\frac{KH}{2}$	$H+\frac{KH}{4}$	$H+\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$
Between Tr 2 & 3	H	H	$H+\frac{3KH}{4}$	$H+\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$
Between Tr 3 & 4	H	H	H	$H+KH$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$
Between Tr 4 & 5	H	H	H	H	$H+\frac{3KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$
Between Tr 5 & 6	H	H	H	H	H	$H+\frac{KH}{2}$	$H-\frac{KH}{4}$	$H-\frac{KH}{4}$
Between Tr 6 & 7	H	H	H	H	H	H	$H+\frac{KH}{4}$	$H-\frac{KH}{4}$
After Train 7	H	H	H	H	H	H	H	H
Headways for Blocked Train ↑								

Passenger Waiting Time

If Strategy II Holding is used, the increase in passenger waiting time resulting from a KH blockage can be written as the sum of three terms:

$$\Delta W_{Strat.II} = X + Y + Z \tag{2.3}$$

where

X = the increased waiting time at the stations behind the blockage,

Y = the increased waiting time at the $(N' - N_{AH})$ stations which are ahead of the blockage and initially beyond the first held train, and

Z = the increased waiting time at the N_{AH} stations which are at or ahead of the blockage location, but initially behind the first held train.

By studying the headway patterns at each station, we can find the X, Y, and Z terms by subtracting the waiting times in an undisturbed situation from the waiting times resulting from a KH blockage. Thus,

$$
\begin{aligned}
X &= AF\left(\frac{1}{2}\times\sum_{i=1}^{n_b}\left(H+\frac{iKH}{n_b+1}\right)^2+\frac{1}{2}\times\sum_{j=1}^{n_b} j\times\left(H-\frac{KH}{n_b+1}\right)^2-\frac{1}{2}\times H^2\sum_{k=2}^{n_b+1}k\right)\\
&= \frac{AFK^2H^2}{2}\frac{\left(n_b^2+2n_b\right)}{3\left(n_b+1\right)} \qquad (2.4)\\
Y &= A\left(\frac{(n_b+1)}{2}\left(H-\frac{KH}{n_b+1}\right)^2+\frac{(n_a+1)}{2}\left(H+\frac{KH}{n_a+1}\right)^2-\frac{(n_a+n_b+2)}{2}H^2\right)\\
&\qquad \times\left(N'-N_{AH}\right)\\
&= \frac{AK^2H^2}{2}\left(\frac{1}{n_a+1}+\frac{1}{n_b+1}\right)\times\left(N'-N_{AH}\right) \qquad (2.5)
\end{aligned}
$$

The waiting time at each station in the set N_{AH} (term Z) is equal to the waiting time at a station initially beyond the set of held trains *plus* waiting time caused by the additional headway variance at the stations in N_{AH}. The first of these components, summed over the set N_{AH}, is simply equal to $Y\times(N_{AH})/(N'-N_{AH})$. Studying the headways patterns in the table above and noting the symmetry, we find that the second component is equal to X , with n_a in place of n_b. Summing the "X" and "Y" components, we have

$$Z=\frac{AFK^2H^2}{2}\frac{\left(n_a^2+2n_a\right)}{3\left(n_a+1\right)}+\frac{AK^2H^2}{2}\left(\frac{1}{n_a+1}+\frac{1}{n_b+1}\right)\times N_{AH}. \qquad (2.6)$$

Finally, we sum the X, Y and Z terms to get an equation for the increase in passenger waiting time if Strategy II is employed:

$$\Delta W_{Strat.II}=\frac{AK^2H^2}{2}\left(\frac{\frac{F}{3}\left(n_a^2+2n_a\right)+N'}{n_a+1}+\frac{\frac{F}{3}\left(n_b^2+2n_b\right)+N'}{n_b+1}\right). \qquad (2.7)$$

In order to find the values of n_a and n_b which minimize the waiting time, we take the derivative of ΔW with respect to each of these variables, set the derivatives equal to 0, and solve the resulting quadratic equations. We are left with $n_a = n_b = -1 \pm \sqrt{1-2+3N'/F}$. Since we know that n_a and n_b must be positive, we are left with a single solution:

$$n_a^* = n_b^* = n^* = \sqrt{\frac{3N'}{F} - 1} - 1 \tag{2.8}$$

Note that the value n^* which minimizes ΔW is independent of KH , and thus is independent of duration of the blockage.

As an example, suppose that for the case with no layovers, $N_A = 9$ and $F = 1.5$. Then, $n^* = 3.80$. Since n must be an integer (number of trains), we compare ΔW for $\lfloor n_* \rfloor = 3$ with ΔW for $\lceil n^* \rceil = 4$, and select the optimal value of $n = 4$.

We can substitute the equation for n^* into Equation 2.7 to get the optimal (minimum) increase in passenger waiting time when Strategy II Holding is used for a KH blockage:

$$\Delta W^{OPT}_{Strat.II} = \frac{2AK^2H^2\left(N' + \frac{F}{3}\right)}{\sqrt{\frac{3N'}{F} - 1}} \tag{2.9}$$

Note that there exists a minimum feasible value for the number of trains held behind the blockage. This value depends on the ratio of the minimum safe headway to the standard operating headway H. If the number of trains held from behind is less than the minimum value $n_{b_{\min}}$, then $(n_{b_{\min}} - n_b)$ trains will run up from behind and will be blocked by the n_b trains, so that in effect $n_{b_{\min}}$ trains are "held". Values for $n_{b_{\min}}$ are given in the table below:

MINIMUM SAFE HEADWAY	0	$\frac{H}{2}$	$\frac{2H}{3}$	$\frac{3H}{4}$	$\frac{4H}{5}$	$\frac{(C-1)H}{C}$
MINIMUM FEASIBLE n_b	0	$2K-1$	$3K-1$	$4K-1$	$5K-1$	$CK-1$

Similarly, there exists a maximum feasible value for the number of trains held ahead of the blockage. For Case 1, where we have layovers between round trips, we know that n_a must be no greater than N_A/F, the number of trains between the blockage and the terminal. Therefore, the following inequality must hold in order for Equation 2.9 to be valid:

$$\begin{aligned} n_a^* &\leq \frac{N_A}{F} \\ \sqrt{\frac{3N'}{F} - 1} - 1 &\leq \frac{N_A}{F} \end{aligned} \tag{2.10}$$

After rearranging terms and varying parameter F over a range of possible values, we find that this inequality is satisfied for $N_A \geq 2$, in the worst case. Thus, Equation 2.9 is typically valid for Case 1 (with layovers).

For Case 2, where we do not have layovers, we know that n_a must be less than $T/2 - 1$. In order for Equation 2.9 to be valid, the following inequality must hold:

$$\begin{aligned} n_a^* &\leq \frac{T}{2} - 1 \\ \sqrt{\frac{3NM}{F} - 1} - 1 &\leq \frac{T}{2} - 1 \end{aligned}$$

Solving for M in terms of F and T, and recalling that $F = N/T$, we get the following equation:

$$M \leq \frac{T}{12} + \frac{1}{3T} \tag{2.11}$$

This equation is satisfied only when M is small and T is large (i.e. $M < 1.7$ for $T = 20$). Thus, for most situations without layovers, Equation 2.9 is not valid. Instead, we can set $n_a = n_b = T/2 - 1 = (N/2F) - 1$, and substitute into Equation 2.7 to get an equation which typically holds for Case 2 when Strategy II is used:

$$\Delta W_{Strat.II}^{OPT} = AK^2H^2 \left(2FM + \frac{N}{6} - \frac{2F^2}{3N} \right) \tag{2.12}$$

Passenger On-Board Time

After studying the headway patterns which result from Strategy II holding, we can write an equation for the increase in passenger on-board time as a result of a KH blockage:

$$\begin{aligned} \Delta O_{Strat.II} &= \left(1 + \frac{1}{n_a + 1} \sum_{i=1}^{n_a} i + \frac{1}{n_b + 1} \sum_{i=1}^{n_b} i \right) \times KHL \\ &= \left(1 + \frac{n_a + n_b}{2} \right) \times KHL \end{aligned} \tag{2.13}$$

Note that ΔO increases as n_a and n_b increase. As a result, the value of $n = n_a = n_b$ which minimizes the total of waiting time *plus* on-board time will be smaller than the value n^* found for waiting time alone. Furthermore, this value now depends on the duration of the blockage,

KH.

2.2.3 Strategy III

Strategy III is the same as Strategy II in that n_a trains are held in front of, and n_b trains are held behind, the blocked train so that the headways at stations are evened out to the greatest extent possible. However, in Strategy III the trains behind the blockage are not held until they run up as close to the blockage as possible. Each train in front of the blockage is held at the next station it reaches after the blockage occurs, as is done for Strategy II. The increase in on-board time will be equal to that for Strategy II (see Equation 2.13). The passenger waiting time increase, however, will be less than for Strategy II, by an amount less than or equal to X (see Equation 2.4). We can now bound the waiting time increase for Strategy III:

$$\frac{AK^2H^2}{2}\left(\frac{\frac{F}{3}\left(n_a^2+2n_a\right)+N'}{n_a+1}+\frac{N'}{n_b+1}\right) \leq \Delta W_{Strat.III} \leq \frac{AK^2H^2}{2} \times \left(\frac{\frac{F}{3}\left(n_a^2+2n_a\right)+N'}{n_a+1}+\frac{\frac{F}{3}\left(n_b^2+2n_b\right)+N'}{n_b+1}\right) \tag{2.14}$$

2.2.4 Strategy IV

Strategy IV is the "do nothing" strategy; no trains are held. The trains behind the blockage run up as close to the blockage as possible, and proceed as soon as the blockage is cleared. This case is equivalent to Strategy III with $n_a = 0$ and $n_b = n_{b_{\min}}$. Therefore, this strategy can be no more effective than Strategy III, for which we can choose to hold an optimal number of trains.

2.2.5 Comparison of Alternative Holding Strategies

As indicated above, Strategy IV ("do nothing") is less effective than Strategy III, and need not be considered in an analysis which determines the optimal holding strategy. Although Strategy III is slightly more effective than Strategy II, the two are approximately equivalent. Therefore, for computational ease, we will compare Strategies I and II in order to determine the optimal holding strategy.

Case 1: With Layovers

Since $(1+(n_a+n_b)/2) < T$, it is obvious from Equations 2.2 and 2.13 that $\Delta O_{Strat.II} < \Delta O_{Strat.I}$: the on-board time increase is less for Strategy II than for Strategy I. The waiting time for Strategy II will be less than for Strategy I for the case with layovers if the following inequality holds (See Equations 2.1 and 2.9):

$$\begin{aligned} \Delta W_{Strat.II} &< \Delta W_{Strat.I} \\ \frac{2AK^2H^2\left(N' + \frac{F}{3}\right)}{\sqrt{\frac{3N'}{F} - 1}} &< NAH^2\frac{K(K+1)}{2} \\ \frac{4\left(N' + \frac{F}{3}\right)}{\sqrt{\frac{3N'}{F} - 1}}\left(\frac{K}{K+1}\right) &< N \end{aligned} \tag{2.15}$$

In Table 2.1, we vary parameters F, N' , and K over a range of values in order to determine the conditions for which Inequality 2.15 is satisfied. Since $N \geq (N' = N_A + 1)$, Inequality 2.15 is typically satisfied. The waiting time under Strategy II is less than for Strategy I for most situations where we have layovers.

We conclude that Strategy II (or III) is preferable to Strategy I when there are layovers between round trips.

Case 2: Without Layovers

In a system without terminal layovers, the total passenger time (waiting plus on-board) is less for Strategy II than for Strategy I if the following inequality holds (See Equations 2.1, 2.2, 2.12 and 2.13):

$$\begin{aligned} \Delta W_{Strat.II} + \Delta O_{Strat.II} &< \Delta W_{Strat.I} + \Delta O_{Strat.I} \\ AK^2H^2\left(2FM + \frac{N}{6} - \frac{2F^2}{3N}\right) + \left(1 + \frac{n_a + n_b}{2}\right)KHL &< NAH^2\frac{K(K+1)}{2} + \\ & \quad TKHL \end{aligned} \tag{2.16}$$

Rearranging terms, and recalling that $T = N/F$, we find:

F	N'	K	$\frac{4(N'+F/3)K}{(3N'/F-1)^{0.5}(K+1)}$
1/3	2	1/2	0.7
1/3	2	4	1.6
1/3	10	1/2	1.4
1/3	10	4	3.4
1/3	100	1/2	4.5
1/3	100	4	11
1	2	1/2	1.4
1	2	4	3.3
1	10	1/2	2.6
1	10	4	6.1
1	100	1/2	7.7
1	100	4	19
5	6	1/2	6.3
5	6	4	15
5	10	1/2	7
5	10	4	17
5	100	1/2	18
5	100	4	42

Table 2.1: Case 1

$$M < \frac{T}{4}\left(\frac{K+1}{K} - \frac{1}{3}\right) + \frac{T}{4}\left(\frac{L}{FAKH}\right) + \frac{1}{3T} \tag{2.17}$$

The first term on the right side of this inequality is due to the difference in waiting time, the second term is due to the increase in on-board time, and the third term is relatively insignificant. In the table below, we vary parameters T and K over a range of reasonable values in order to determine the conditions for which Inequality 2.17 is satisfied. Neglecting the on-board time, we have:

T	3	3	10	10	20	20	40	40
K	0.5	4	0.5	4	0.5	4	0.5	4
$\frac{T}{4}\left(\frac{K+1}{K} - \frac{1}{3}\right) + \frac{1}{3T}$	2.1	0.8	6.7	2.3	13.4	4.6	26.7	9.2

For the on-board time, we find that $L/(FAKH) = O(L/AH \times 1/F \times 1/K) = O(5 \times 0.7 \times 0.5) = O(2)$. Strategy II becomes preferable to Strategy I as the length of the blockage or number of remaining round trips decrease, or as the number of trains in the system increases.

2.3 The System S Short-Turning Problem

2.3.1 Short-Turning Strategies

Consider a one-track blockage which leads to a delay of K times the standard operating headway H at that location. One intuitive control strategy is to short-turn K trains. The passenger waiting and on-board time increases are analyzed below.

First, assume that K trains are immediately short-turned from *behind* the blockage (see Figure 2-2 below). The trains outside the short-turning loop (defined by a set of N_1 stations)

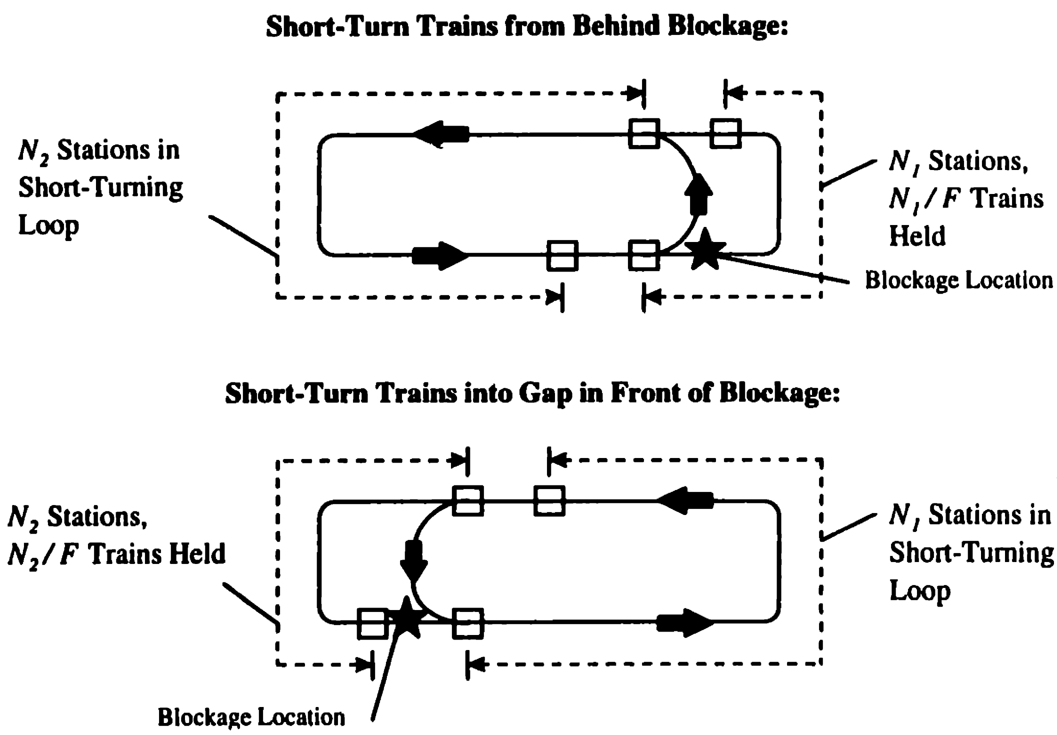


Figure 2-2: Short-Turning Options

are held for the duration of the blockage. For the set of N_1 stations, the on-board and waiting time increases are:

$$\Delta O_{N_1} = \frac{N_1}{F} KHL \tag{2.18}$$

$$\Delta W_{N_1} = N_1 A H^2 \frac{K(K+1)}{2} \tag{2.19}$$

Within the set of N_2 stations (within the short-turning loop), there is no on-board time increase. However, an increase in waiting time is experienced by people who are forced to alight from the short-turning trains. Given the assumption that A and L are constant, $(L - AH)$ people are forced to alight, and must wait until the blockage is cleared. The length of their wait depends on when they arrive at the short-turning location. Thus,

$$\begin{aligned} \Delta W_{N_2} &= (L - AH)(KH + (K-1)H + (K-2)H + \ldots + H) \\ &= (L - AH) H \sum_{i=1}^{K} i \\ &= (LH - AH^2) \frac{K(K+1)}{2} \end{aligned} \tag{2.20}$$

This result is valid when $K \leq N_2/F$, which will hold for the short-term disruptions we are considering. Summing our results for the two sets, we find:

$$\Delta O_{ST} = \frac{N_1}{F} KHL \tag{2.21}$$

$$\Delta W_{ST} = \left((N_1 - 1) A H^2 + LH\right) \frac{K(K+1)}{2} \tag{2.22}$$

If we next assume that the K trains are short-turned *ahead* of the blockage (see Figure 2-2), and N_2 trains are held, it is simple to show that:

$$\Delta O_{ST} = \frac{N_2}{F} KHL \tag{2.23}$$

$$\Delta W_{ST} = \left((N_2 - 1) A H^2 + LH\right) \frac{K(K+1)}{2} \tag{2.24}$$

Therefore, to minimize passenger time, we select the minimum of N_1 and N_2 and short-turn from behind or ahead of the blockage accordingly.

2.3.2 Comparison of Short-Turning and Holding

Short-Turning Versus Strategy I Holding

In systems without terminal layovers (Case 2) and in which there are a large number of round trips, a small set of trains, and/or a short blockage duration, we showed above that Strategy I is the best holding strategy. We now compare the effect of Strategy I Holding to the short-turning strategy described in Section 2.3.1.

The total passenger time (waiting plus on-board) is less for the short-turning strategy than for Strategy I Holding if the following inequality is satisfied (See Equations 2.1, 2.2, 2.22, and 2.21):

$$\begin{aligned}
\Delta O_{ST} + \Delta W_{ST} &< \Delta O_{Strat.I} + \Delta W_{Strat.I} \\
\frac{N_1}{F} KHL + ((N_1 - 1)\, AH^2 + LH)\, \frac{K\,(K+1)}{2} &< TKHL + NAH^2 \frac{K\,(K+1)}{2}
\end{aligned}$$

After substituting $N_1 + N_2$ for N and $N_1/F + N_2/F$ for T, we rearrange terms to get:

$$\begin{aligned}
LH\frac{K\,(K+1)}{2} &< \frac{N_2}{F} KHL + (N_2+1)\, AH^2 \frac{K\,(K+1)}{2} \\
\frac{(K+1)}{2} L &< \frac{N_2}{F} L + (N_2+1)\, AH \frac{(K+1)}{2} \qquad (2.25)
\end{aligned}$$

Given the assumption that $K < N_2/F$ (on which Equation 2.22 is based), $(K+1)/2 < N_2/F$ and the left side of Inequality 2.25 is less than the right side. Thus, the waiting time is less if we short-turn. Note that the same conclusion is reached if the assumption that $K < N_2/F$ is relaxed, although the derivation is more complicated.

Short-Turning Versus Strategy II Holding

In systems where we have sufficiently long terminal layovers (Case 1), or in systems with no layovers but a small number of round trips, a large set of trains, and/or a long blockage duration, we showed above that Strategy II is preferable to Strategy I. We now compare Strategy II

Holding with the short-turning strategy described in Section 2.3.1. Assume that for Strategy II, we hold an optimal number of trains (i.e. $n = \sqrt{3N'/F - 1} - 1$). By comparing Equations 2.13 and 2.21, we find that the on-board time is less for the short-turning strategy if the inequality below is satisfied:

$$\begin{aligned} \Delta O_{ST} &< \Delta O_{Strat.II} \\ \frac{N_1}{F} KHL &< (1+n) \times KHL \\ \frac{N_1}{F} &< \sqrt{\frac{3N'}{F} - 1} \qquad (2.26) \end{aligned}$$

It is difficult to draw conclusions from Inequality 2.26.

By comparing Equations 2.9 and 2.22, we find that the passenger waiting time is less for the short-turning strategy if the following inequality is satisfied:

$$\begin{aligned} \Delta W_{ST} &< \Delta W_{Strat.II} \\ \left((N_1 - 1) AH^2 + LH\right) \frac{K(K+1)}{2} &< \frac{2AK^2H^2\left(N' + \frac{F}{3}\right)}{\sqrt{\frac{3N'}{F} - 1}} \\ N_1 &< \frac{4\left(N' + \frac{F}{3}\right)}{\sqrt{\frac{3N'}{F} - 1}} \left(\frac{K}{K+1}\right) + 1 - \frac{L}{AH} \qquad (2.27) \end{aligned}$$

Comparing this result to Inequality 2.15 and the accompanying table in Section 2.2.5, we can see how the first term on the right hand side varies over a range of parameter values. Realistic values for L/AH are on the order of 5 to 15. Thus, the waiting time will be less for short-turning only if N_1 is quite small.

2.4 Conclusions Drawn from System S

The following points summarize our key results for System S:

- In Strategy II, trains are held in front of and behind the blockage by varying amounts so as to even out headways at stations to the greatest extent possible. The optimal number of trains to hold is independent of the duration of the blockage.

- There is a minimum feasible number of trains which are held behind a blockage; this number depends on the ratio of the minimum safe headway to the standard operating headway H.

- The "do nothing" strategy is always less effective than at least one of the active control strategies we examined.

- For Case 1 with layovers between trips, both the on-board time and waiting time are typically less for Strategy II (hold $n_a + n_b$ trains) than for Strategy I (hold all trains).

- For a system without layovers, the total passenger time (waiting plus on-board) becomes less for Strategy II than for Strategy I as the length of the blockage or number of remaining round trips decrease, or as the number of trains in the system increases.

- When faced with a single track blockage, it is possible to short-turn trains from behind a blockage, or to short-turn trains from the opposite track into the gap ahead of the blockage. In order to minimize waiting and on-board time, short-turn whichever trains will skip fewer stations as a result.

- We described systems for which Strategy I results in less passenger time than Strategy II. In such systems, passenger time can be reduced by short-turning. For systems in which Strategy II is preferable to Strategy I, short-turning is preferable only if N_1 is quite small.

Chapter 3

Generalized System Model

3.1 Description

The Generalized System Model ("System G") closely approximates the behavior of an actual rail transit system. System G is significantly more realistic and thus more complicated than System S; we did not find closed form results for optimal control strategies in System G. Instead, System G is the basis for a series of mathematical programming formulations which can be used to find optimal control strategies, and which can be solved in real time as part of a decision support system for transit operations. Although the formulations were developed for use during minor service disruptions, they could be easily adapted for use in routine operations control.

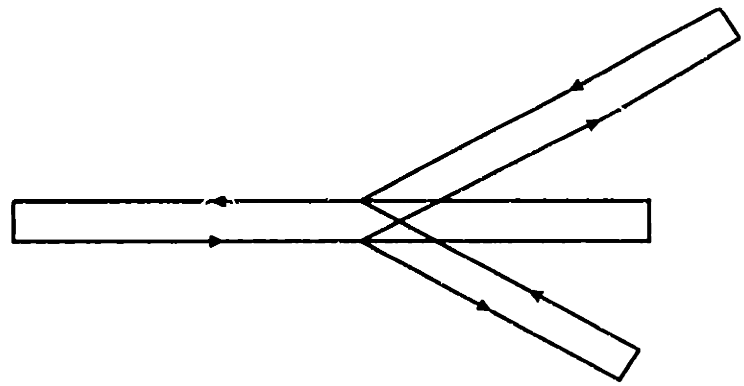
3.1.1 Features and Assumptions

The System G Model incorporates the following features:

- The transit system need not be a simple loop line; it may contain branches (See Figure 3-1). However, the addition of branches adds complexity and increases the size of the problem. The model application in Chapter 4 is based on the MBTA Red Line, which includes two branches.

- The passenger arrival rate and alighting fraction, train running times and minimum safe headways between trains are all station-specific parameters.

Figure 3-1: Transit Line with Branching Structure



- Dwell time of trains at stations is modeled as a linear function of the number of passengers boarding and alighting, as well as the crowding conditions on the train. The dwell time function can be station specific.
- The order of trains need not remain constant. Trains may enter or leave service, or may change order as in the case of short-turning strategies.
- Many types of disruptions can be considered, including a temporary closure or speed restriction on a track section, a delayed train, or a disabled train which must be taken off the line.
- The capacity of the train is considered; during peak hours, passengers may be left behind on the platform.

The System G Model is subject to the following assumptions and limitations:

- The number of passengers boarding and alighting and the load on the trains at the time the disruption occurs are estimated from passenger flow data collected previously and stored. This data may vary by time of day and day of the week.
- Stochastic elements, such as actual running and dwell times, are approximated by their expected values in order to make the problem more tractable.

- The dwell time function, which may be non-linear, has been approximated by a linear function.

- The problem formulations and control strategies consider only a limited "impact set" of trains and stations (see discussion of the impact set in Section 3.1.3 below). As a result, the effects of the disruption and control strategies at stations beyond the impact set are approximated by a function of the headway distribution at the last station in the impact set.

- Perhaps the greatest limitation of the System G Model is the treatment of the running time and minimum headways between stations. The movement of trains through an actual rail transit system can be quite complex, depending on the signal system and control lines. In order to make the problem more tractable, we do not allow a train to depart from a station until it can travel to the next station at the maximum permissible (free-running) speed, without stopping in between. Inter-station stopping or deceleration due to the position of the preceding train is prohibited. Since many transit systems, including the MBTA Red Line, do not operate this way, the model uses holding at stations in order to meet this requirement. If the controls output by the model are applied, then the trains will run at maximum speed until they leave the impact set of stations.

There are several advantages to an operating strategy which prohibits inter-station stopping. Some passengers are uncomfortable when "trapped" in a dark tunnel between stations. If a disruption is expected to cause significant delays, passengers can choose to alight if a train is stopped at a station rather than in a tunnel. Furthermore, additional people can board a train which is held at a station, assuming it is not already full. Although a free-running operating strategy may decrease the capacity of a line because of the increase in safe headways, this decrease may not be significant.

3.1.2 Data Requirements

The following data is required to find optimal control strategies using the System G model formulations:

- Track configuration including branches and crossover locations.

- Train schedule information, including when and where trains enter and leave service, routes, and any specific schedule constraints.
- Passenger arrival rate and alighting fraction at each station, by time of day and day of week.
- Minimum safe headway between trains at each station.
- Minimum (free) running time between stations, including acceleration and deceleration time.
- Dwell-time function parameters, which may be station specific.
- Passenger capacity of train.
- Real-time information on train locations, from which the last station departure time for each train may be calculated. This information can also be used to estimate the current load on each train.
- Location and estimated duration of blockage, and corresponding restrictions on train movements.
- Short-turning parameters, such as the time to cross a train between the tracks, time to operate the switch, time to notify an operator of a short-turn, and minimum safe headways around the short-turning train.

3.1.3 Impact Set

The System G problem formulations and control strategies consider only a limited "impact set" of trains and stations in front of and behind the blockage at the time of the disruption. The station set is determined by the train set: the first (last) station in the impact set is the station from which the first (last) train departs next, immediately following the disruption. The number of trains in the set is governed by several factors. First, the problem must be solved in real time. As the number of variables and constraints is increased, the problem size and solution time increase. Second, the number of trains to be considered should be no larger than can feasibly be controlled, given the existing signal and communication systems. Finally, it is unrealistic to project train movements far into the future, due to the stochastic nature of any transit system. Each of these factors presents an upper bound on the size of the impact set. It is intuitively obvious, however, that larger impact sets result (up to some point) in more effective control strategies, and a decrease in passenger waiting time. It is desirable, therefore, to find a set size which balances these conflicting requirements.

Note that there is also a minimum number of trains behind the blockage which must be held or will be blocked by their predecessors.. If the impact set does not include a sufficient number of trains behind the blockage, then the deleterious effects of the blockage will not be entirely captured by the model. This minimum number depends on the duration of the blockage, and the difference between the scheduled operating headways and the minimum safe headways.

The selection of an appropriate impact set is explored in Section 4.3.1 below.

3.1.4 Variable, Parameter, and Set Definitions

The following variables are used in the System G formulations:

$d_{i,k}$	=	departure time of train i from station k
$l_{i,k}$	=	load on train i departing station k
$p_{i,k}$	=	people left behind on platform by train i at station k
i'	=	predecessor of train i at a particular station (see definition of Set $IatK$ below)
i'_b	=	predecessor of train i on the branched portion of the line
$h_{i,k}$	=	preceding headway of train i at station k, $h_{i,k} = d_{i,k} - d_{i',k}$
$\nabla h_{i,k}$	=	absolute value of the difference between headways of train i and its predecessor train
$z_{i,k}$	=	variable used to approximate quadratic objective function (see Section 3.1.5)
$z'_{i,k}$	=	variable used to approximate quadratic objective function
$v_{i,k}$	=	$\begin{cases} 1, & \text{if train } i \text{ is loaded to capacity when it departs station } k \\ 0, & \text{otherwise} \end{cases}$
$w_{i,k}$	=	$\begin{cases} 1, & \text{if train } i \text{ is held at station } k \\ 0, & \text{otherwise} \end{cases}$

The following parameters are used in the System G formulations:

A_k = passenger arrival rate at station k

A_k^o = arrival rate at station k, for passengers with a destination on the trunk portion of the line

A_k^b = arrival rate at station k, for passengers with a destination on branch b

Q_k = passenger alighting fraction at station k

R_k = free running time from station $(k-1)$ to station k, including acceleration and deceleration time

H_k = minimum safe headway at station k

C_j = dwell time function parameter

$Sch_{i,k}$ = scheduled departure time of train i from station k

M = "sufficiently large" coefficient, used in integer programming constraints

U_j = weighting factor in the objective function

$L_{\max}$ = passenger capacity of train

T_{BL} = earliest time at which the blocked or disabled train can leave station

k_{if} = the first station reached by train i after the disruption occurs

i_{BL} = blocked or disabled train

k_l = last station in impact set

The following sets are used in the System G formulations:

I is the set of trains in the impact set.

I_b is the set of trains from branch b in the impact set.

K is the set of stations in the impact set.

K_o is the set of stations on the trunk portion of the line which are beyond k_l but which are included in the objective function.

K_b is the set of stations in branch b which are beyond k_l, but are included in the objective function.

$IatK$ is the train/station impact set which contains one element corresponding to each decision variable $d_{i,k}$. An element (i, i', k) consists of train i, station k,and the predecessor train i' of train i at station k .

$IatK'$ is equivalent to $IatK$ except that it excludes any element (i, i', k) such that $(i', i'', k) \notin IatK$. Thus, for each element (i, i', k) of $IatK'$, both $d_{i,k}$ and $d_{i',k}$ are decision variables.

Additional variables, parameters and sets which apply specifically to short-turning formulations are described in Section 3.3.1 below.

3.1.5 Objective Function

During short-term service disruptions, our chosen objective in determining optimal control strategies is to minimize passenger waiting time. An alternate objective would be to minimize total passenger time (waiting plus on-board time); a discussion of this objective is given in Section 4.3.3 below.

The passenger waiting time at stations within the impact set, assuming no passengers are left behind by an overloaded train, is given by:

$$\frac{1}{2}\sum_{IatK}\left(A_k^o\left(d_{i,k}-d_{i',k}\right)^2+A_k^b\left(d_{i,k}-d_{i_b',k}\right)^2\right)$$

The first term is the waiting time for passengers whose destination is either on the trunk portion of the line, or on the same branch from which they originate. These passengers may board any train. The second term is the waiting time for passengers who board at a station on the trunk portion of the line, but whose destinations are on one of the branches. Note that the sub/superscript b in this term corresponds to the destination branch b for train i.

For stations beyond the impact set, we can approximate the passenger waiting time by using the headways at the last station in the impact set. The set of stations included in the objective function is bounded by several considerations. First, if there is "sufficient" layover time at the dispatching terminal, then uneven headways will be corrected there, and stations beyond the terminal should not be included in the objective function. Second, it is unrealistic to include stations far beyond k_l, because the headway distribution at k_l will not be a good approximation of the headways at those stations. For stations in the trunk portion of the line, we have:

$$\frac{1}{2}\sum_{K_o}\left(A_k^o\sum_{I}\left(d_{i,k_l}-d_{i',k_l}\right)^2+\sum_{b}A_k^b\left(\sum_{I_b}\left(d_{i,k_l}-d_{i_b',k_l}\right)^2\right)\right)$$

and for stations in the branched portion of the line, we have:

$$\frac{1}{2}\sum_{b}\left(\sum_{K_b} A_k \left(\sum_{I_b}\left(d_{i,k_l} - d_{i'_b,k_l}\right)^2\right)\right)$$

The waiting time terms can be multiplied by weighting factors U_j based on the distance of the stations from k_l; as the distance from k_l increases, the approximation of the headway distribution becomes less accurate and the weighting factor should decrease.

There is additional waiting time for passengers who are left behind at a station when they are unable to board an overloaded train. This waiting time is equal to the number of passengers left behind by each train, multiplied by the headway of the applicable *following* train.

The waiting time terms, for passengers who are able to board the first train departing for their destination, are quadratic functions. Since the formulations developed below for System G are large mixed integer programs, and are intended for real-time use, we developed a piecewise linear approximation for these functions. This approximation is of the form

$$h_{i,k}^2 = \left(d_{i,k} - d_{i',k}\right)^2 \approx z_{i,k} = \max_{n=1..N}\left[f_n \times h_{i,k} - g_n\right] \tag{3.1}$$

where N is the number of piecewise segments used in approximating the quadratic function.

There is an inherent problem with the piecewise linear approximation. For a given mean headway, the passenger waiting time is minimized when the headway variance is minimized, i.e. all headways are equal, if possible. With the piecewise linear approximation, however, the passenger waiting time is minimized for an infinite number of headway combinations chosen along one piecewise segment. Although the passenger waiting time will be close to optimal, erratic headway patterns may result. To remedy the situation, we added a term to the objective function which penalizes headway variance:

$$PENALTY = U_{pen}\sum_{IatK'} A_k \left|h_{i,k} - h_{i',k}\right| = U_{pen}\sum_{IatK'} A_k \nabla h_{i,k} \tag{3.2}$$

Even a small weighting factor U will achieve the desired result of eliminating unnecessary headway variance.

The additional waiting time for passengers left behind an overloaded train is also described by a nonlinear function. Unfortunately, the function is the inseparable product of two different variables, $p_{i,k}$ and $h_{i+1,k}$, and is not readily approximated by a linear function. Therefore, in our System G formulations we simply minimize the number of passengers left behind, multiplied by a weighting factor. This factor depends on whether the passengers are left behind by a train which is in front of the blocked train, where headways are longer than scheduled, or behind the blocked train, where they are shorter. The solutions which result from this simplified objective function should be close to optimal for the minor disruptions we are considering.

After we incorporate the piecewise linear approximation and penalty terms, and eliminate the constant 1/2, the objective function for the System G Holding and Short-Turning Problems is:

$$\min \sum_{IatK} \left(A_k^o z_{i,k} + A_k^b z'_{i,k} \right) + U_o \sum_{K_o} \left(A_k^o \sum_{I} z_{i,k_l} + \sum_{b} A_k^b \sum_{I_b} z'_{i,k_l} \right)$$
$$+ \sum_{b} U_b \sum_{K_b} A_k \sum_{I_b} z'_{i,k_l} + U_{pen} \sum_{IatK'} A_k \nabla h_{i,k} + U_{ahd} \sum_{\substack{IatK \\ ahd}} p_{i,k} + U_{bhd} \sum_{\substack{IatK \\ bhd}} p_{i,k} \quad (3.3)$$

where

$$\begin{aligned} z_{i,k} &= \max_{n=1..N} \left[f_n \times h_{i,k} - g_n \right] \\ z'_{i,k} &= \max_{n=1..N} \left[f_n \left(d_{i,k} - d_{i'_b,k} \right) - g_n \right] \\ \nabla h_{i,k} &= \left| h_{i,k} - h_{i',k} \right| \end{aligned}$$

and "*ahd*" and "*bhd*" refer to the trains ahead of, and behind, the blocked train, respectively.

3.1.6 Dwell-Time Function

To successfully apply the System G Model, we need a credible dwell time function. Although the dwell time at each station could perhaps be approximated by its mean during routine operations, the dwell time variance will be larger during the service disruptions for which System G was developed.

In his 1990 thesis, Lin [9] demonstrated that a dwell time function which is linear in the number of passengers boarding and alighting can be quite powerful; he used such a model to explain 70% of the variation in dwell time for the MBTA Green Line, a light rail line, during routing operations. Lin also considered the effect of crowding by introducing various load variables into his model; this resulted in a small increase in explanatory power. However, observations of rail transit operations following service disruptions indicate that the effect of crowding is an important factor, and should be included in the System G Model.

Although the true dwell time function may be non-linear, we use a linear function in the System G problem formulations in order to make the problem more tractable. When the passenger load is well below the train capacity, crowding is not an issue and we use a dwell time function of the form:

$$dwelltime = C_o + C_1 A_k (d_{i,k} - d_{i',k}) + C_2 Q_k l_{i,k-1} + C_1 p_{i',k} \tag{3.4}$$

where C_o is the constant term, and C_1 and C_2 are the marginal boarding and alighting times that apply in uncrowded conditions. The number of passengers boarding the train equals $A_k(d_{i,k} - d_{i',k}) + p_{i',k}$; the number alighting equals $Q_k l_{i,k-1}$.

As the passenger loads approach the train capacity, we model the dwell time by a pair of functions. When train i will depart from station k with a load less than its capacity, we use Equation 3.4 for its dwell time. However, if train i will be loaded to capacity when it departs, we use:

$$dwelltime = C_3' + C_4 \left(L_{\max} - (1 - Q_k)\, l_{i,k-1} \right) + C_5' Q_k l_{i,k-1} \tag{3.5}$$

where C_3' is a constant term, and C_4 and C_5' are the marginal boarding and alighting times that apply in crowded conditions. In this case, the number of passengers who board is equal to the number of available spots in the train after the alighting process. We can simplify Equation 3.5, writing it as:

$$dwelltime = C_3 - C_4 l_{i,k-1} + C_5 Q_k l_{i,k-1} \tag{3.6}$$

where $C_3 = C_3' + C_4 L_{\max}$ and $C_5 = C_4 + C_5'$.

We could also add linear terms for the arriving or departing load ($C_6 l_{i,k-1}$, $C_6 l_{i,k}$), or the number of passengers left behind ($C_7 p_{i,k}$), to the dwell time functions.

3.2 Holding Formulations

We developed a series of mathematical programming formulations for the holding problem, in which one of several holding strategies is used to mitigate the effects of a short-term blockage. The three holding strategies we considered differ in the extent of required control actions, and potentially in their effectiveness. The first strategy is to hold each train at any of the stations in the impact set ("Hold All"). Although this strategy is the most effective, it is infeasible unless there exists an efficient mechanism for controlling train departure times. An example of such a mechanism would be a holding signal at each station which signals when a train should depart. The second, and simplest, strategy is to hold each train at the first station it reaches after the disruption occurs ("Hold at First"). The third strategy is to hold each train at only one optimally chosen station in the impact set ("Hold Once"). The three formulations differ in their complexity and in their solution time, as described in Chapter 4.

We developed formulations for both the Fixed Order and Variable Order problems. In the Fixed Order problem, the order in which trains from separate branches enter a junction is given. Any problem on a loop line is a Fixed Order problem. In the Variable Order problem, the order in which trains from separate branches enter a junction is a decision variable.

We distinguish between formulations in which the train capacity is constrained, and those in which it is not. The formulations without capacity constraints apply when the passenger load on the trains is well below the train capacity. This might be a realistic assumption for off-peak operations. During peak hours, however, capacity constraints are typically required.

3.2.1 Fixed Order Holding Problem Without Capacity Constraints

A formulation for the Fixed Order Holding Problem without capacity constraints (FOHP), for the "Hold All" strategy, is given below. Note that there are no terms in the objective function for passengers who are left behind by an overloaded train.

"Hold All"

$$\min \sum_{IatK} \left(A_k^o \left(d_{i,k} - d_{i',k}\right)^2 + A_k^b \left(d_{i,k} - d_{i'_b,k}\right)^2 \right)$$

$$+U_o \sum_{K_o} \left(A_k^o \sum_{I} \left(d_{i,k_l} - d_{i',k_l}\right)^2 + \sum_{b} A_k^b \left(\sum_{I_b} \left(d_{i,k_l} - d_{i'_b,k_l}\right)^2 \right) \right)$$

$$+\sum_{b} U_b \left(\sum_{K_b} A_k \left(\sum_{I_b} \left(d_{i,k_l} - d_{i'_b,k_l}\right)^2 \right) \right)$$

subject to:

Headway: $d_{i,k} \geq d_{i',k+1} + H_{k+1} - R_{k+1} \qquad \forall (i,i',k) \in IatK \,|\, k \neq k_l$

Load: $l_{i,k} = A_k^o \left(d_{i,k} - d_{i',k}\right) + A_k^b \left(d_{i,k} - d_{i'_b,k}\right) + (1 - Q_k)\, l_{i,k-1} \qquad \forall (i,i',k) \in IatK$

RunTime: $d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1 \left(A_k^o \left(d_{i,k} - d_{i',k}\right) + A_k^b \left(d_{i,k} - d_{i'_b,k}\right)\right) + C_2 Q_k l_{i,k-1} \qquad \forall (i,i',k) \in IatK$

Blockage: $d_{i_{BL},k_{BL}} \geq T_{BL}$

$d_{i,k}\,,\ l_{i,k} \geq 0 \qquad \forall i,k$

The Headway constraints ensure that the minimum safe headway between trains is maintained. These constraints prevent train i from departing station k until it can run to the next station $(k+1)$ without being blocked by the preceding train i'. Thus, inter-station stopping is prohibited. The RunTime constraints indicate that the earliest time of departure for train i from station k is equal to the departure time from the previous station, plus the running time to station i, plus the dwell time of train i at station k. The Load constraints calculate the number of passengers on board train i as it departs station k, as a function of the load entering the station and the number of passengers boarding and alighting. Finally, the Blockage constraint(s) restrict the movement of the train(s) directly affected by the disruption, based on the estimated duration of the blockage. Blockage constraints may be written for multiple trains at multiple stations.

The above formulation presents the most basic version of the problem. For transit system geometries which include branches, special constraints must be written to reflect the discontinuities in station numbering which must occur at junctions. There may be special headway constraints at intermediate terminals, or station-specific dwell-time functions in the RunTime constraints. There may also be schedule constraints which prevent trains from departing ahead of schedule, of the form $d_{i,k} \geq Sch_{i,k}$.

As written above, the Fixed Order Holding Problem has a quadratic objective function and must be solved as a quadratic programming problem. By replacing the objective function with its piecewise linear approximation, we can rewrite the formulation as a linear program which can be solved very quickly using a commercial solver. The objective function above is replaced by:

$$\min \sum_{IatK} \left(A_k^o z_{i,k} + A_k^b z'_{i,k}\right) + U_o \sum_{K_o} \left(A_k^o \sum_{I} z_{i,k_l} + \sum_{b} A_k^b \sum_{I_b} z'_{i,k_l}\right) + \sum_{b} U_b \sum_{K_b} A_k \sum_{I_b} z'_{i,k_l} + U_{pen} \sum_{IatK'} A_k \nabla h_{i,k}$$

In addition, we must add the following constraints:

Piecewise A: $z_{i,k} \geq f_n \times h_{i,k} - g_n \qquad \forall (i,i',k) \in IatK,\ \forall n \,|\, 1 \leq n \leq N$

Piecewise B: $z'_{i,k} \geq f_n \left(d_{i,k} - d_{i'_b,k}\right) - g_n \qquad \forall (i,i',k) \in IatK,\ \forall n \,|\, 1 \leq n \leq N$

Piecewise C: $h_{i,k} = d_{i,k} - d_{i',k} \qquad \forall (i,i',k) \in IatK$

Piecewise D: $\nabla h_{i,k} \geq h_{i,k} - h_{i',k} \qquad \forall (i,i',k) \in IatK'$

Piecewise E: $\nabla h_{i,k} \geq h_{i',k} - h_{i,k} \qquad \forall (i,i',k) \in IatK'$

$z_{i,k},\ z'_{i,k},\ h_{i,k},\ \nabla h_{i,k} \geq 0 \qquad \forall i,k$

Note that the absolute value and maximum functions in Equations 3.1 and 3.2 have been replaced by linear inequalities.

"Hold at First"

As discussed above, it is not always feasible to hold trains at multiple stations. Therefore, we consider a control strategy in which each train can be held at only the first station it reaches after the disruption occurs. At all other stations, the train must depart as soon as it is able, i.e. either the RunTime or the Headway constraint must be binding. Since our model assumes that trains do not depart from stations until they can travel at maximum speed to the next station, trains are held for sufficient duration at the first station to prevent inter-station stopping or deceleration due to blocking at any point within the impact set. Thus, the RunTime constraint must be the binding constraint. The "Hold at First" strategy will not be as effective as the "Hold All" strategy; we will compare their effectiveness for several problem instances in Chapter 4.

The "Hold at First" formulation is identical to the "Hold All" formulation presented above, with the exception of the RunTime constraints. The RunTime inequality constraints must be replaced by equality constraints at all stations except the first one each train reaches following the disruption, as given below:

$$\text{RunTimeG:} \quad d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1\left(A_k^o\left(d_{i,k} - d_{i',k}\right) + A_k^b\left(d_{i,k} - d_{i'_b,k}\right)\right) + C_2 Q_k l_{i,k-1} \quad \forall (i,i',k) \in IatK \mid k = k_{if}$$

$$\text{RunTimeEq:} \quad d_{i,k} = d_{i,k-1} + R_k + C_o + C_1\left(A_k^o\left(d_{i,k} - d_{i',k}\right) + A_k^b\left(d_{i,k} - d_{i'_b,k}\right)\right) + C_2 Q_k l_{i,k-1} \quad \forall (i,i',k) \in IatK \mid k \neq k_{if}$$

The "Hold at First" formulation, like the "Hold All" formulation, is a linear program for FOHP.

"Hold Once"

The "Hold Once" strategy entails holding each train at only one station, similar to the "Hold at First" strategy above. However, the holding station for this strategy need not be the first station, but instead can be any station in the impact set. The formulation for this strategy is the same as for "Hold All" or "Hold First", with the exception of the RunTime constraints and the addition of binary variable $w_{i,k}$, which equals 1 if train i is held at station k. The RunTime constraints for the "Hold Once" version of FOHP are as follows:

$$\text{RunTimeG:} \quad d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1\left(A_k^o\left(d_{i,k} - d_{i',k}\right) + A_k^b\left(d_{i,k} - d_{i'_b,k}\right)\right) + C_2 Q_k l_{i,k-1} \quad \forall (i,i',k) \in IatK$$

$$\text{RunTimeL:} \quad d_{i,k} \leq d_{i,k-1} + R_k + C_o + C_1\left(A_k^o\left(d_{i,k} - d_{i',k}\right) + A_k^b\left(d_{i,k} - d_{i'_b,k}\right)\right) + C_2 Q_k l_{i,k-1} + M_1 w_{i,k} \quad \forall (i,i',k) \in IatK$$

$$\textstyle\sum_k w_{i,k} \leq 1 \quad \forall i$$

$$w_{i,k} \in \{0,1\} \quad \forall i,k$$

M_1 is a sufficiently large coefficient that prevents the RunTimeL constraint from binding when $w_{i,k} = 1$. Thus, the RunTime constraints are satisfied with equality at all stations except the one holding station for each train. Note that the "Hold Once" formulation of FOHP is a mixed integer program.

3.2.2 Fixed Order Holding Problem With Capacity Constraints

The introduction of capacity constraints to the Fixed Order Holding Problems necessitates the use of binary variables. Therefore, all formulations of the Fixed Order Holding Problem with Capacity constraints (FOHPC) are mixed integer programs.

In the formulations given below for FOHPC, we will assume that the impact set does not include any stations on the trunk for which there is significant boarding of passengers with destinations on a branch, i.e. A_k^b is negligible for stations in the impact set. If we relax this assumption, then we replace the term $A_k(d_{i,k} - d_{i',k})$ in the constraints below with $A_k^o(d_{i,k} - d_{i',k}) + A_k^b(d_{i,k} - d_{i'_b,k})$. To be strictly correct, we should also consider that Q_k will vary between trains, depending on the ratio of $A_k^o(d_{i,k} - d_{i',k})$ to $A_k^b(d_{i,k} - d_{i'_b,k})$. However, this is a second order effect that can typically be neglected .As before, we will first develop the formulation for the "Hold All" strategy, and then extend it to the "Hold at First" and "Hold Once" strategies.

"Hold All"

$$\min \sum_{IatK} A_k z_{i,k} + U_o \sum_{K_o} \left(A_k^o \sum_{I} z_{i,k_l} + \sum_{b} A_k^b \sum_{I_b} z'_{i,k_l} \right) + \sum_{b} U_b \sum_{K_b} A_k \sum_{I_b} z'_{i,k_l} + U_{pen} \sum_{IatK'} \nabla h_{i,k} + U_{ahd} \sum_{\substack{IatK \\ ahd}} p_{i,k} + U_{bhd} \sum_{\substack{IatK \\ bhd}} p_{i,k}$$

subject to:

Piecewise A: $z_{i,k} \geq f_n \times h_{i,k} - g_n \quad \forall (i,i',k) \in IatK,\ \forall n \,|\, 1 \leq n \leq N$

Piecewise B: $z'_{i,k} \geq f_n \left(d_{i,k} - d_{i'_b,k} \right) - g_n \quad \forall (i,i',k) \in IatK,\ \forall n \,|\, 1 \leq n \leq N$

Piecewise C: $h_{i,k} = d_{i,k} - d_{i',k} \quad \forall (i,i',k) \in IatK$

Piecewise D: $\nabla h_{i,k} \geq h_{i,k} - h_{i',k} \quad \forall (i,i',k) \in IatK'$

Piecewise E: $\nabla h_{i,k} \geq h_{i',k} - h_{i,k} \qquad \forall(i, i', k) \in IatK'$

Load A: $l_{i,k} \leq A_k\left(d_{i,k} - d_{i',k}\right) + (1 - Q_k)\, l_{i,k-1} + p_{i',k} \qquad \forall(i, i', k) \in IatK$

Load B: $l_{i,k} \leq L_{\max} \qquad \forall(i, i', k) \in IatK$

Load C: $l_{i,k} \geq A_k\left(d_{i,k} - d_{i',k}\right) + (1 - Q_k)\, l_{i,k-1} + p_{i',k} - M_1 v_{i,k} \qquad \forall(i, i', k) \in IatK$

Load D: $l_{i,k} \geq L_{\max} \times v_{i,k} \qquad \forall(i, i', k) \in IatK$

Left Behind: $p_{i,k} = A_k\left(d_{i,k} - d_{i',k}\right) + (1 - Q_k)\, l_{i,k-1} + p_{i',k} - l_{i,k} \qquad \forall(i, i', k) \in IatK$

Headway: $d_{i,k} \geq d_{i',k+1} + H_{k+1} - R_{k+1} \qquad \forall(i, i', k) \in IatK \mid k \neq k_l$

RunTime A: $d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1 A_k(d_{i,k} - d_{i',k}) + C_2 Q_k l_{i,k-1} + C_1 p_{i',k} - M_2 v_{i,k}$
$$\forall(i, i', k) \in IatK$$

RunTime B: $d_{i,k} \geq d_{i,k-1} + R_k + C_3 - C_4 l_{i,k-1} + C_5 Q_k l_{i,k-1} - M_3\left(1 - v_{i,k}\right)$
$$\forall(i, i', k) \in IatK$$

Blockage: $d_{i_{BL},k_{BL}} \geq T_{BL}$

$$d_{i,k}\,,\ l_{i,k}\,,\ z_{i,k}\,,\ z'_{i,k}\,,\ h_{i,k}\,,\ \nabla h_{i,k}\,,\ p_{i,k} \geq 0 \qquad \forall i, k$$
$$v_{i,k} \in \{0, 1\} \qquad \forall i, k$$

M_1, M_2, and M_3 are "sufficiently large" coefficients which prevent Load C, RunTime A and RunTime B from binding when $v_{i,k}$ equals 0 or 1, as applicable.

The Load constraint has been replaced by four separate constraints which reflect the fact that the load is equal to the *minimum* of the load capacity and the number of passengers trying to board. When the train is full, $v_{i,k} = 1$, and the Load B and Load D constraints are binding. When the train is not yet full, $v_{i,k} = 0$, and the Load A and Load C constraints are binding. Load A and Load C include variable $p_{i',k}$, the number of passengers who were unable to board the previous train. Variable $p_{i,k}$ is calculated in the Left Behind constraints as the sum of the passengers who arrived at the station since the last train departed, the load on train i when it arrived, and the number left behind by the previous train, minus the fraction who alighted at station k and the load on train i as it departs.

There are two RunTime constraints in FOHPC. If train i is not full as it departs station k, then $v_{i,k} = 0$, and RunTime A is equivalent to the RunTime constraint in the uncapacitated problem (FOHP). When the train is full, however, $v_{i,k} = 1$, and constraint RunTime B is the binding constraint. RunTime B reflects the increased dwell time due to congestion effects, as was discussed in Section 3.1.6.

"Hold at First"

For the "Hold at First" strategy in FOHPC, the formulation is the same as for the "Hold All" strategy with the exception of the RunTime constraints, which must be modified as they were for FOHP. However, we can not use an equality for the RunTime constraints when $k \neq k_{i,f}$, because we do not know whether the RunTime A or RunTime B constraint will be binding. Therefore, we must use inequalities to ensure that *either* RunTime A or RunTime B is satisfied with equality when $k \neq k_{i,f}$. The RunTime constraints for FOHPC, "Hold at First", are:

RunTimeG A: $$d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1 A_k (d_{i,k} - d_{i',k}) + C_2 Q_k l_{i,k-1} + C_1 p_{i',k} - M_2 v_{i,k} \quad \forall (i, i', k) \in IatK$$

RunTimeG B: $$d_{i,k} \geq d_{i,k-1} + R_k + C_3 - C_4 l_{i,k-1} + C_5 Q_k l_{i,k-1} - M_3 (1 - v_{i,k}) \quad \forall (i, i', k) \in IatK$$

RunTimeEq A: $$d_{i,k} \leq d_{i,k-1} + R_k + C_o + C_1 A_k (d_{i,k} - d_{i',k}) + C_2 Q_k l_{i,k-1} + C_1 p_{i',k} + M_3 v_{i,k} \quad \forall (i, i', k) \in IatK \mid k \neq k_{if}$$

RunTimeEq B: $$d_{i,k} \leq d_{i,k-1} + R_k + C_3 - C_4 l_{i,k-1} + C_5 Q_k l_{i,k-1} + M_2 (1 - v_{i,k}) \quad \forall (i, i', k) \in IatK \mid k \neq k_{if}$$

"Hold Once"

For the "Hold Once" strategy in FOHPC, the RunTime constraints are:

RunTimeG A: $$d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1 A_k (d_{i,k} - d_{i',k}) + C_2 Q_k l_{i,k-1} + C_1 p_{i',k} - M_2 v_{i,k} \quad \forall (i, i', k) \in IatK$$

RunTimeG B: $$d_{i,k} \geq d_{i,k-1} + R_k + C_3 - C_4 l_{i,k-1} + C_5 Q_k l_{i,k-1} - M_3 (1 - v_{i,k}) \quad \forall (i, i', k) \in IatK$$

RunTimeEq A: $$d_{i,k} \leq d_{i,k-1} + R_k + C_o + C_1 A_k (d_{i,k} - d_{i',k}) + C_2 Q_k l_{i,k-1} + C_1 p_{i',k} + M_3 v_{i,k} + M_4 w_{i,k} \quad \forall (i, i', k) \in IatK$$

RunTimeEq B: $$d_{i,k} \leq d_{i,k-1} + R_k + C_3 - C_4 l_{i,k-1} + C_5 Q_k l_{i,k-1} + M_2 (1 - v_{i,k}) + M_4 w_{i,k} \quad \forall (i, i', k) \in IatK$$

M_4 is a sufficiently large coefficient that prevents the RunTimeEq constraints from binding when $w_{i,k} = 1$. Thus, either RunTime A or RunTime B must be satisfied with equality at all stations except the one holding station for each train.

3.2.3 Variable Order Holding Problem

The Variable Order Holding Problem (VOHP) applies only when a blockage occurs on one of the branches of the line. In VOHP, we allow trains to deviate from their scheduled order as they merge at a junction, i.e. we allow trains from the unblocked branch to "skip ahead" of one or more trains from the blocked branch. Assume without loss of generality that the trains are numbered according to their scheduled order beyond the junction (on the trunk line), so that trains with lower numbers are ahead of trains with higher numbers.

We introduce the following notation for the Variable Order Holding Problem:

K^T	is the set of stations in the impact set K which are on the trunk of the line
K^{BR}	is the set of stations in the impact set K which are on one of the branches
S_i	is the set all of possible trains which can precede train i at Stations $k \in K^T$
i_J	$= \max i \mid i$ is beyond the junction (within the trunk portion of the line) when the disruption occurs
i_l	is the last train in the impact set
$y_{j,i}$	$= \begin{cases} 1, & \text{if train } j \text{ precedes train } i \text{ through the junction and at stations in } K^T \\ 0, & \text{otherwise} \end{cases}$

In developing the VOHP formulation, we will assume that the transit system has two branches, and that the blockage occurs on branch $b = 1$. Then,

for $i \leq i_J$, $S_i = \{i - 1\}$,

for $i > i_J \mid i \in I_1$, $S_i = \{i'_b, (i-1), (i+1, i+2, ..., i_l) \cap I_2\}$,

and for $i > i_J \mid i \in I_2$, $S_i = \{i'_b, (i-1), (i_J, ..., i-3, i-2) \cap I_1\}$.

Note that we can decrease the size of S_i by limiting the number of trains which may skip ahead of, or slip behind, train i.

All versions of the Variable Order Holding Problem are formulated as mixed integer programs because of the binary variable $y_{j,i}$. We will assume that the impact set does not include any stations on the trunk for which there is significant boarding of passengers with destinations on a branch, i.e. A^b_k is negligible for stations in the impact set. For VOHP, the impact set includes the section of the line where the branches enter the trunk. This is typically far from the section where the trunk separates into branches again. A formulation for VOHP without capacity constraints, for the "Hold All" strategy, is given below.

$$\min \sum_{IatK} A_k z_{i,k} + U_o \sum_{K_o} \left(A_k^o \sum_{I} z_{i,k_l} + \sum_{b} A_k^b \sum_{I_b} z'_{i,k_l} \right) + \sum_{b} U_b \sum_{K_b} A_k \sum_{I_b} z'_{i,k_l} + U_{pen} \sum_{IatK'} \nabla h_{i,k}$$

subject to:

Pcwise A:	$z_{i,k} \geq f_n \times h_{i,k} - g_n \quad \forall (i,i',k) \in IatK,\ \forall n \mid 1 \leq n \leq N$
Pcwise B:	$z'_{i,k} \geq f_n \left(d_{i,k} - d_{i'_b,k} \right) - g_n \quad \forall (i,i',k) \in IatK,\ \forall n \mid 1 \leq n \leq N$
Pcwise C1:	$h_{i,k} = d_{i,k} - d_{i',k} \quad \forall (i,i',k) \in IatK \mid (k \in K^{BR})\ or\ (i \leq i_J\ and\ k \in K^T)$
Pcwise D1:	$\nabla h_{i,k} \geq h_{i,k} - h_{i',k} \quad \forall (i,i',k) \in IatK' \mid (k \in K^{BR})\ or\ (i \leq i_J\ and\ k \in K^T)$
Pcwise E1:	$\nabla h_{i,k} \geq h_{i',k} - h_{i,k} \quad \forall (i,i',k) \in IatK' \mid (k \in K^{BR})\ or\ (i \leq i_J\ and\ k \in K^T)$
Headway 1:	$d_{i,k} \geq d_{i',k+1} + H_{k+1} - R_{k+1}$ $\forall (i,i',k) \in IatK \mid (k \in K^{BR})\ or\ (i \leq i_J\ and\ k \in K^T \backslash k_l)$
Load 1:	$l_{i,k} = A_k \left(d_{i,k} - d_{i',k} \right) + (1 - Q_k)\, l_{i,k-1}$ $\forall (i,i',k) \in IatK \mid (k \in K^{BR})\ or\ (i \leq i_J\ and\ k \in K^T)$
RunTime 1:	$d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1 A_k \left(d_{i,k} - d_{i',k} \right) + C_2 Q_k l_{i,k-1}$ $\forall (i,i',k) \in IatK \mid (k \in K^{BR})\ or\ (i \leq i_J\ and\ k \in K^T)$
Pcwise C2:	$h_{i,k} \geq d_{i,k} - d_{j,k} + (y_{j,i} - 1) M_1 \quad \forall (i,i',k) \in IatK \mid (i > i_J\ ,\ k \in K^T), \forall j \in S_i$
Pcwise D2:	$\nabla h_{i,k} \geq h_{i,k} - h_{j,k} + (y_{j,i} - 1) M_2$ $\forall (i,i',k) \in IatK \mid (i > i_J\ ,\ k \in K^T), \forall j \in S_i$
Pcwise E2:	$\nabla h_{i,k} \geq h_{j,k} - h_{i,k} + (y_{j,i} - 1) M_2$ $\forall (i,i',k) \in IatK \mid (i > i_J\ ,\ k \in K^T), \forall j \in S_i$
Headway 2:	$d_{i,k} \geq d_{j,k+1} + H_{k+1} - R_{k+1} + (y_{j,i} - 1)\, M_3$ $\forall (i,i',k) \in IatK \mid (i > i_J\ ,\ k \in K^T \backslash k_l), \forall j \in S_i$
Load 2G:	$l_{i,k} \geq A_k \left(d_{i,k} - d_{j,k} \right) + (1 - Q_k)\, l_{i,k-1} + (y_{j,i} - 1)\, M_4$ $\forall (i,i',k) \in IatK \mid (i > i_J\ ,\ k \in K^T), \forall j \in S_i$
Load 2L:	$l_{i,k} \leq A_k \left(d_{i,k} - d_{j,k} \right) + (1 - Q_k)\, l_{i,k-1} - (y_{j,i} - 1)\, M_4$ $\forall (i,i',k) \in IatK \mid (i > i_J\ ,\ k \in K^T), \forall j \in S_i$
RunTime 2:	$d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1 A_k \left(d_{i,k} - d_{j,k} \right) + C_2 Q_k l_{i,k-1} + (y_{j,i} - 1)\, M_5$ $\forall (i,i',k) \in IatK \mid (i > i_J\ ,\ k \in K^T), \forall j \in S_i$

Blockage: $d_{i_{BL},k_{BL}} \geq T_{BL}$

OnePred: $\sum_{j \in S_i} y_{j,i} = 1 \quad \forall i \geq i_J$

OneSucc: $\sum_{i | j \in S_i} y_{j,i} = 1 \quad \forall j \in \cup_{i_J}^{i_l} S_i$

$d_{i,k}\,,\ l_{i,k},\ z_{i,k}\,,\ z'_{i,k}\,,\ h_{i,k}\,,\ \nabla h_{i,k} \geq 0 \quad \forall i,k$

$y_{j,i} \in \{0,1\}$

M_1 through M_5 are sufficiently large coefficients which prevent constraints Pcwise C2, Pcwise D2,..., RunTime 2 from binding unless $y_{j,i} = 1$. Thus, these constraints are only binding when train j precedes train i through the junction. Constraint Pcwise C2 is written as an inequality; however, it will be satisfied with equality when $y_{j,i} = 1$, because we are minimizing an objective function which varies directly with $h_{i,k}$. The Load C2 constraint has been written as a pair of inequalities which are equivalent to an equality when $y_{j,i} = 1$. Constraints OnePred and OneSucc ensure that each train has only one predecessor and one successor at the north end of the line.

Although the Variable Order Holding Problem formulation above was written for the "Hold All" strategy, VOHP may be easily modified for the "Hold at First" and "Hold Once" strategies as we did for FOHP.

3.3 Short-Turning Formulations

We developed mathematical programming formulations for the short-turning problem, in which short-turning is used in conjunction with holding to mitigate the effects of a minor disruption. To motivate these formulations, we first consider a simple loop line. Assume that we have a central business district (CBD) near the middle of the loop, and that peak traffic flow is in and out of the CBD in the morning and evening, respectively. When a one-track blockage occurs, there are three possibilities for the short-turning of trains (Refer again to Figure 2-2): (1) one or more trains may be short-turned from behind the blockage, (2) one or more trains may be short-turned from the opposite track into the gap in front of the blockage, or (3) trains are short-turned both from behind the blockage and into the gap in front of it. If a train which has not yet reached the CBD is short-turned during the morning peak hours, then it is likely that many people will be forced to alight, which is undesirable. If a train which has not yet reached the CBD is short-turned in the evening peak hours, then few people will benefit. Thus, the decision whether to short-turn from behind the blockage, or into the gap in front of the blockage, should be made so that the short-turned train(s) are already past the CBD.

The discussion above suggests several strategies for which we could develop formulations, including strategies in which none, one, or multiple trains are short-turned from behind or in front of the blockage. The final train order (after the short-turn) may be predetermined, or may be determined as part of the optimization. We present a formulation for the Short-Turning Problem with Predetermined Train Order below, and discuss how it can be extended to the problem with undetermined train order. The formulation is given for the strategy in which each train can be held at any of the stations in the impact set ("Hold All"). The formulation may be easily modified for the "Hold at First" and "Hold Once" strategies. Capacity constraints have been included.

The short-turning train may be several stations away from the cross-over location at the time the disruption occurs; we assume here that the train continues with normal passenger pick-ups and drop-offs until all passengers alight at the station immediately preceding the cross-over. Other assumptions may be modeled as well.

3.3.1 Variable, Parameter, and Set Definitions

The variable, parameter, and set definitions given in Section 3.1.4 apply to the System G short-turning formulations; additional notation is given below.

The following variables, illustrated in Figure 3-2, are used in the short-turning formulations:

i_{st}	=	train which is short-turned
k_{st}	=	station preceding the short-turn
$i_{st}-1$	=	train preceding i_{st} before short-turn
$i_{st}+1$	=	train following i_{st} before short-turn
i_p	=	train preceding i_{st} after short-turn
i_f	=	train following i_{st} after short-turn
k'_{st}	=	station across from k_{st} (platform opposite k_{st})

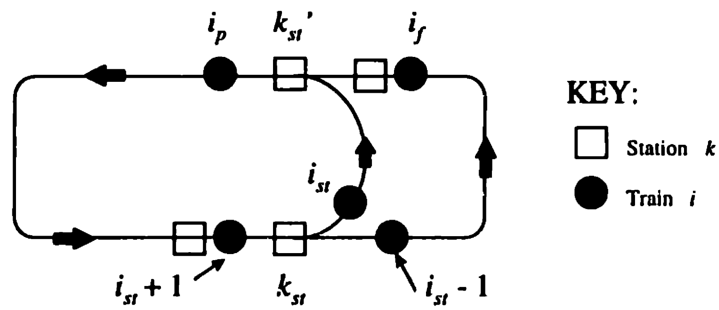


Figure 3-2: System G Short-Turning Variables

The following parameters are used in the short-turning formulations:

T_{sw} = time to throw the switch at the crossover. If train i does not cross over but train $(i+1)$ does, or vice versa, then T_{sw} is the minimum time between their departures at maximum speed from k_{st}.

$X_{k'_{st}}$ = running time to cross over (short-turn) and get into position at station k'_{st}. This time may be affected by T_{sw}.

R_{sw} = the running time from station $(k'_{st}-1)$ until past the switch, *minus* the running time from station k_{st} to the switch

T_{not} = time until the operator of train i_{st} can be notified of the short-turn

C_o^{st} = dwell-time constant for i_{st} at k_{st}, $C_o^{st} \geq C_o$

These parameters are more clearly explained by their context in the constraints given below.

3.3.2 Short-Turning Problem with Predetermined Train Order

The formulation for the Short-Turning Problem with Predetermined Train Order (STPP), using the "Hold All" strategy, is given below. In this formulation, the train/station impact set $IatK$, which was defined earlier in Section 3.1.4, includes element (i^*, i'^*, k^*) only if passengers can board train i^* at station k^*, i.e. there is no element for i_{st} at k_{st}. Note that the set $IatK$ reflects the *revised* ordering of trains after the short-turn.

$$\min \sum_{IatK} A_k z_{i,k} + U_o \sum_{K_o} \left(A_k^o \sum_{I} z_{i,k_l} + \sum_{b} A_k^b \sum_{I_b} z'_{i,k_l} \right) + \sum_{b} U_b \sum_{K_b} A_k \sum_{I_b} z'_{i,k_l} + U_{pen} \sum_{IatK'} \nabla h_{i,k} + U_{ahd} \sum_{\substack{IatK \\ ahd}} p_{i,k} + U_{bhd} \sum_{\substack{IatK \\ bhd}} p_{i,k}$$

subject to:

Piecewise A: $z_{i,k} \geq f_n \times h_{i,k} - g_n \quad \forall (i,i',k) \in IatK,\ \forall n \,|\, 1 \leq n \leq N$

Piecewise B: $z'_{i,k} \geq f_n \left(d_{i,k} - d_{i'_b,k} \right) - g_n \quad \forall (i,i',k) \in IatK,\ \forall n \,|\, 1 \leq n \leq N$

Piecewise C: $h_{i,k} = d_{i,k} - d_{i',k} \quad \forall (i,i',k) \in IatK$

Piecewise D: $\nabla h_{i,k} \geq h_{i,k} - h_{i',k} \quad \forall (i,i',k) \in IatK'$

Piecewise E: $\nabla h_{i,k} \geq h_{i',k} - h_{i,k} \quad \forall (i,i',k) \in IatK'$

Load A: $l_{i,k} \leq A_k \left(d_{i,k} - d_{i',k} \right) + (1 - Q_k)\, l_{i,k-1} + p_{i',k} \quad \forall (i,i',k) \in IatK$

Load B: $l_{i,k} \leq L_{\max} \quad \forall (i,i',k) \in IatK$

Load C: $l_{i,k} \geq A_k \left(d_{i,k} - d_{i',k} \right) + (1 - Q_k)\, l_{i,k-1} + p_{i',k} - M_1 v_{i,k} \quad \forall (i,i',k) \in IatK$

Load D: $l_{i,k} \geq L_{\max} \times v_{i,k} \quad \forall (i,i',k) \in IatK$

Left Behind: $p_{i,k} = A_k \left(d_{i,k} - d_{i',k} \right) + (1 - Q_k)\, l_{i,k-1} + p_{i',k} - l_{i,k} \quad \forall (i,i',k) \in IatK$

Headway: $d_{i,k} \geq d_{i',k+1} + H_{k+1} - R_{k+1} \quad \forall (i,i',k) \in IatK \backslash (i_f, i_p, k'_{st} - 1) \,|\, k \neq k_l$

RunTime A: $d_{i,k} \geq d_{i,k-1} + R_k + C_o + C_1 A_k (d_{i,k} - d_{i',k}) + C_2 Q_k l_{i,k-1} + C_1 p_{i',k} - M_2 v_{i,k}$
$\forall (i,i',k) \in IatK \setminus (i_{st},\, i_p,\, k'_{st})$

RunTime B: $d_{i,k} \geq d_{i,k-1} + R_k + C_3 - C_4 l_{i,k-1} + C_5 Q_k l_{i,k-1} - M_3 \left(1 - v_{i,k}\right)$
$\forall (i,i',k) \in IatK \setminus (i_{st},\, i_p,\, k'_{st})$

ST-1: $d_{i_f,k'_{st}-1} \geq d_{i_{st},k'_{st}} + H_{k'_{st}} - R_{k'_{st}}$

ST-2: $d_{i_{st},k'_{st}} \geq d_{i_{st},k_{st}} + X_{k'_{st}} + C_o + C_1 A_{k'_{st}} (d_{i_{st},k'_{st}} - d_{i_p,k'_{st}}) + C_1 p_{i_p,k'_{st}} - M_2 v_{i_{st},k'_{st}}$

ST-3: $d_{i_{st},k'_{st}} \geq d_{i_{st},k_{st}} + X_{k'_{st}} + C_3 - M_3 (1 - v_{i_{st},k'_{st}})$

ST-4: $d_{i_{st},k_{st}} \geq d_{i_{st}-1,k_{st}} + T_{sw}$

ST-5: $d_{i_{st}+1,k_{st}} \geq d_{i_{st},k_{st}} + T_{sw}$

ST-6: $d_{i_{st},k_{st}} \geq d_{i_p,k'_{st}-1} + R_{sw} + T_{sw}$

ST-7: $d_{i_{st},k_{st}} \geq d_{i_{st},k_{st}-1} + R_{k_{st}} + C_o^{st} + C_2 l_{i_{st},k_{st}-1}$

ST-8: $d_{i_{st},k_{st}} \geq T_{not} + C_o^{st} + C_2((1 - Q_{k_{st}})l_{i_{st},k_{st}-1} + A_{k_{st}}(T_{not} - d_{i_{st}-1,k_{st}}) + p_{i_{st}-1,k_{st}})$

ST-9: $d_{i_{st},k_{st}} \geq d_{i_p,k'_{st}} + H_{k'_{st}} - X_{k'_{st}}$

ST-10: $p_{i_{st}-1,k_{st}} = (1 - Q_{k_{st}})\, l_{i_{st},k_{st}-1}$

ST-11: $l_{i_{st},k'_{st}-1} = 0$

Blockage: $d_{i_{BL},k_{BL}} \geq T_{BL}$

$$d_{i,k}\,,\ l_{i,k}\,,\ z_{i,k}\,,\ z'_{i,k}\,,\ h_{i,k}\,,\ \nabla h_{i,k}\,,\ p_{i,k} \geq 0 \qquad \forall i,k$$

$$v_{i,k} \in \{0,1\} \qquad \forall i,k$$

This formulation is identical to the "Hold All" formulation for FOHPC given in Section 3.2.2, with the following exceptions:

- The Headway constraint for element (i_f, i_p, $k'_{st} - 1$) has been replaced by constraint ST-1, since train i_f will follow i_{st}, not train i_p, after station $k'_{st} - 1$.
- Constraints RunTime A and RunTime B for element (i_{st}, i_p, k'_{st}) have been replaced by ST-2 and ST-3, since train i_{st} reaches station k'_{st} via the cross-over.
- Constraints ST-4 through ST-11 have been added. ST-4 through ST-6 incorporate the time required to throw the switch at the cross-over, before and after the short-turn. ST-7 is a runtime constraint for the short-turning train which states that train i_{st} must arrive at k_{st} and unload all of its passengers before it can depart. ST-8 is another runtime constraint for train i_{st}; it states that train i_{st} must first be notified that it will be short-turned and must then unload all of its passengers, including any who boarded at k_{st} prior to the notification. ST-9 is a headway constraint for the short-turning train. In this form, ST-9 prevents i_{st} from departing station k_{st} until it can travel over the cross-over and into position at $k_{st'}$ without being blocked by train i_p. ST-9 could be rewritten to permit stopping on the crossover ($d_{i_{st},k'_{st}} \geq d_{i_p,k'_{st}} + H_{k'_{st}}$). Constraint ST-10 assigns the passengers left behind by train i_{st} at station k_{st} to train $i_{st} - 1$, so that they are considered correctly in the constraints for element ($i_{st} + 1$, $i_{st} - 1$, $k_{st} + 1$). Finally, ST-11 sets the passenger load on train i_{st} to 0 as it enters station k'_{st}. There may be additional constraints specific to the transit system being modeled; an example of such a constraint is given in the model application in Chapter 4.

3.3.3 Short-Turning Problem with Undetermined Train Order

Consider a situation in which we might short-turn a train from behind the blockage. In the Short-Turning Problem with Undetermined Train Order (STPU), the decision of whether, and in front of which train, we should short-turn is determined during the optimization. A similar problem occurs when we consider short-turning a train into the gap ahead of the blockage. For this problem, we must decide which (if any) train to short-turn.

For STPU, elements (i, i', k) of $IatK$ are generated assuming that no short-turning takes place; $IatK$ is the same set as for FOHP(C). We introduce the binary variable y_i, which equals 1 if we short-turn i_{st} in front of i $(y_i = 1$ for $i = i_f)$. We also introduce the set S, which is the set of all possible trains in front of which we can short-turn. $|S|$ is less than the number of trains which would pass k'_{st} before the disruption ends if no controls were applied. Then, the constraint $\sum_S y_i \leq 1$ requires that train i_{st} can be short-turned in front of no more than one train in S. The variable y_i is used as a switch in setting up the constraints, as is illustrated for the Load A set of constraints:

$$l_{i,k} \leq A_k \left(d_{i,k} - d_{i',k}\right) + \left(1 - Q_k\right) l_{i,k-1} + p_{i',k}$$
$$\forall (i, i', k) \in IatK \mid k < k_{st} \text{ for } i = i_{st} \text{ , and } k < k'_{st} \text{ for } i \in S$$

$$l_{i,k} \leq A_k \left(d_{i,k} - d_{i',k}\right) + \left(1 - Q_k\right) l_{i,k-1} + p_{i',k} + L_{\max} y_i$$
$$\forall (i, i', k) \in IatK \mid i \in S \text{ and } k \geq k'_{st}$$

$$l_{i,k} \leq A_k \left(d_{i,k} - d_{i_{st},k}\right) + \left(1 - Q_k\right) l_{i,k-1} + p_{i_{st},k} + L_{\max} \left(1 - y_i\right) \qquad \forall i \in S,\ k \geq k'_{st}$$

$$l_{i_{st},k} \leq A_k \left(d_{i_{st},k} - d_{i',k}\right) + \left(1 - Q_k\right) l_{i_{st},k-1} + p_{i',k} + L_{\max} \left(1 - y_i\right)$$
$$\forall (i, i', k) \in IatK \mid i \in S \text{ and } k \geq k'_{st}$$

$$l_{i_{st},k} \leq A_k \left(d_{i_{st},k} - d_{i_{st}-1,k}\right) + \left(1 - Q_k\right) l_{i_{st},k-1} + p_{i_{st}-1,k} + L_{\max} \textstyle\sum_S y_i \qquad \forall\, k \geq k_{st}$$

Similar blocks of constraints, using y_i as a switch, are required for the Load C, Left Behind, Headway, RunTime A, and RunTime B and Piecewise constraints. As a result, the formulation becomes rather long and complicated. More importantly, the number of constraints, problem complexity, and solution time increase. Note that we could simply compare the solutions of several versions of STPP, in lieu of solving STPU.

Chapter 4

Model Application

4.1 Problem Setting

4.1.1 The MBTA Red Line

We applied the System G formulations to several problems using data from the MBTA Red Line during morning peak hour operations. The MBTA Red Line is a heavy rail system with a branching structure (See Figure 4-1). Trains are typically dispatched onto the line from two terminal stations, Ashmont and Braintree, which are located at the southern end of the two branches. Train schedules allow for layovers of approximately ten minutes at Ashmont and Braintree under routine operations. In addition, there are layovers of approximately five minutes at Alewife Station, which is located at the northern end of the trunk portion of the line.

There are five stations on the Ashmont Branch, six stations on the Braintree branch, and twelve on the trunk. With the exception of the terminals, the northbound and southbound platforms are considered as separate stations for modeling purposes; thus, we have a total of 43 stations in the Red Line model.

Each of the problem instances we tested is based on a disruption occurring at 8:15 AM, when passenger volumes are high and operating headways are short. Under these conditions, a relatively minor disruption can have serious consequences if appropriate control actions are not taken. At 8:15 AM, there are 11 Ashmont trains and 16 Braintree trains operating on the

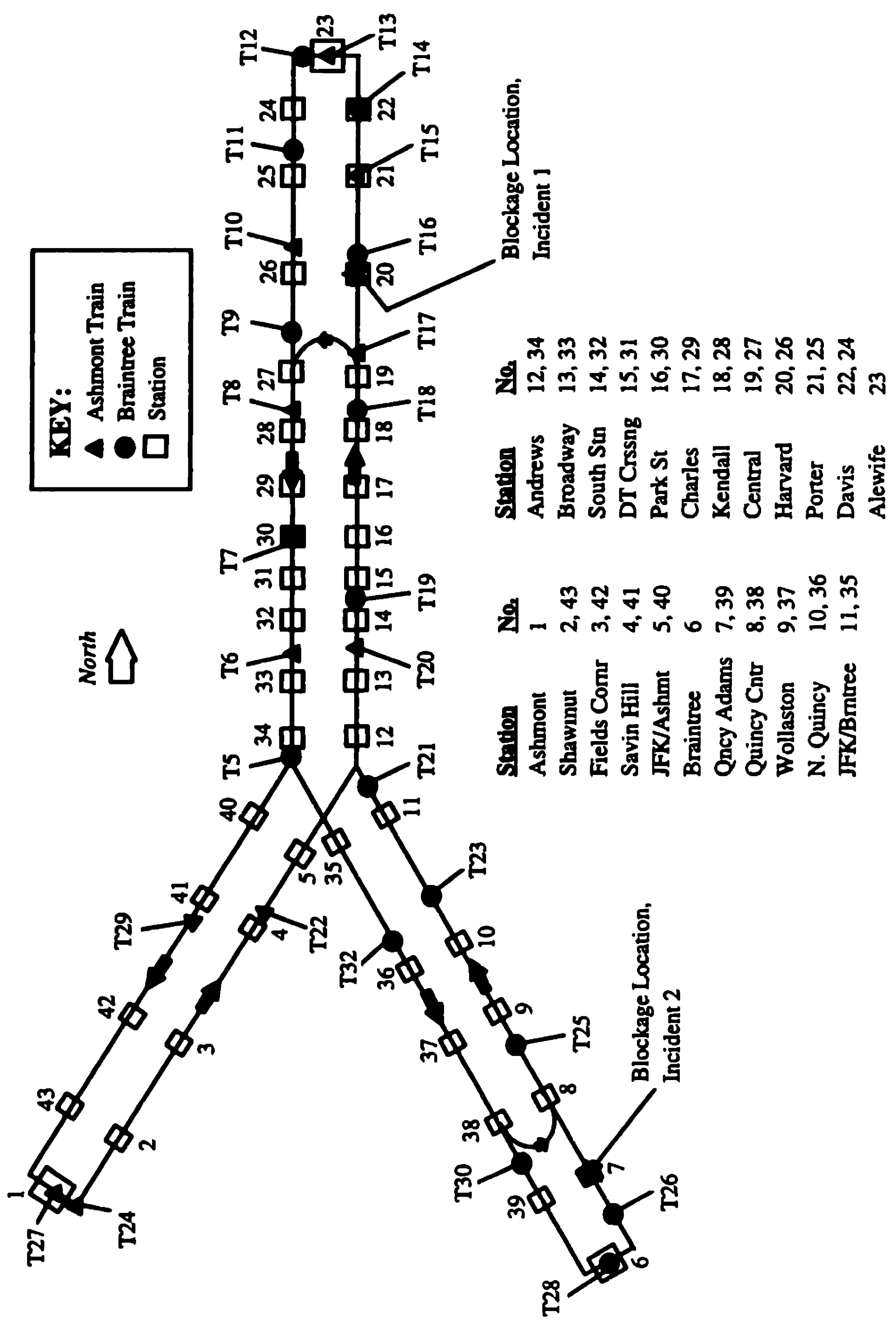


Figure 4-1: **MBTA Red Line, 8:15 AM**

Red Line. Each six car train has a capacity of approximately 1200 people. The Ashmont and Braintree trains are dispatched at headways of approximately 8 and 6 minutes, respectively, resulting in a headway of approximately 3.4 minutes on the northern trunk. Since we tested the model formulations off-line, we did not use "real-time" information on train locations, but instead located the trains based on the dispatching schedule and nominal running times. The trains were assigned numbers based on their departure times at Alewife. The train numbers and locations are indicated in Figure 4-1.

In addition to the constraints given in the Chapter 3 formulations, some additional system-specific constraints are required to model the Red Line:

- Separate RunTime and Headway constraints are required for Ashmont trains passing through the junction (JFK Station to Andrews, or vice versa), because the numbering of consecutive stations is not contiguous.

- Different dwell time parameters are required for the RunTime constraints at Park Street Station. There are platforms on both sides of the train at Park Street, so that passengers board and alight through twice as many doors.

- The Headway constraint is modified at several stations in the central business district, where stations are located so close together that train departures are affected by the dwell times of trains at two stations ahead.

- Trains entering Alewife Station can stop on either of two platforms. The Headway constraint at Alewife is replaced by two constraints: one requires a safe following distance behind the preceding train, and the other requires that a platform is free upon arrival. The latter constraint depends on the departure time of the train which is two trains ahead.

- The dwell time of trains at Alewife is independent of the dwell time function. If a train is on or ahead of schedule, the train remains at Alewife for a five minute layover time. If a train entering Alewife is behind schedule by no more than three minutes, the layover is reduced so that the train departs on time. For trains which are behind by more than three minutes, the layover is reduced to two minutes (the minimum time required for the driver to "change ends"). The RunTime constraint at Alewife is therefore replaced by a pair of constraints in which $x = 1$ if train i is on or ahead of schedule at Alewife, and equals 0 otherwise:

$$d_{i,k} \geq Sch_i - Mx_i$$
$$d_{i,k} \geq d_{i,k-1} + R_k + 2 + 3x_i$$

- Since there are relatively long layovers at the dispatching terminals, we do not hold trains at stations located before the dispatching terminals for blockages located beyond the terminals. Instead, we simply delay their dispatching times. In addition, we add a constraint that trains can not be dispatched before their scheduled time. Since the impact sets we consider in our test cases do not extend on both sides of the terminal stations, and since we have "sufficiently" long layovers at the terminals, we do not need to model the complicated train movements which take place there.

- At Alewife Station, trains do not depart as soon as the passengers have alighted or boarded; rather, the departure times are controlled by a dispatcher. Therefore, a train can be "held" at Alewife without implementing an additional control action. When the Hold Once and Hold at First formulations are applied to the Red Line, holding at Alewife is permitted *in addition* to the one holding station.

- The Red Line control system requires that several track circuits behind a switch be clear before a train can cross over for a short-turn. Therefore, we add an additional constraint to our short-turning formulations. For example, in Incident 1 we require that $d_{i_f,k'_{st}-2} + R_{k'_{st}-1} \geq d_{i_{st},k_{st}} + T_x$, where T_x is the time it takes for the short-turning train to cross-over, the driver to change ends, and the train to begin moving towards station k'_{st}. Thus, the following train can not enter Harvard until the short-turning train has completed the cross-over and is moving towards Central Station.

4.1.2 Input Data

The data requirements for the System G formulations were described in Section 3.1.2. In order to apply our model formulations, we needed actual data or good estimates for each of these items.

To estimate the arrival rates and alighting fractions at each station, we used an extensive data set collected by Massachusetts' Central Transportation Planning Staff (CTPS)[5]. The data set consists of detailed counts of passengers arriving and alighting at each station, for fifteen minute intervals throughout the day. The counts were collected on a train-by-train basis, but not all were collected on the same day. In fact, the Park Street Station counts are a composite of 1985 and 1989 counts, because only limited information was collected there in 1989. As a result, the total number of boardings on the line does not equal the total number of alightings. Discontinuities were handled at the stations in the Central Business District where the volumes are highest, so that the percentage error is minimized. We smoothed the data by averaging the counts for a one hour period around 8:15 AM. The estimated boarding rates and alighting fractions are given in Appendix A. The boarding rates and alighting fractions were used, in conjunction with the scheduled headways, to estimate the current load on each train at the time of the disruption.

The CTPS data was also used in our estimation of a dwell time function. The MBTA provided us with some data on dwell times during the PM peak hours at Park Street Station and South Station, in both the northbound and southbound directions. The data includes the times of train arrivals and departures, as well as estimates of the passenger load in the car for which the doors must remain open longest (typically the most crowded car). From the measured train loads and headways, we used the CTPS data to estimate the number of passengers boarding and alighting and the passenger load for each train in the dwell time data set.

Unfortunately, the dwell time data set we obtained was rather small, and the approximations we used are subject to "measurement" errors. Therefore, we estimated only a crude dwell time function, in which the boarding and alighting rates are constrained to be equal. The dwell time functions we used are given below, for time in minutes. The variable *num* is equal to the total number of passengers boarding and alighting.

Uncrowded Conditions, Typical Stations:	$dwelltime = 0.44 + 0.00150 * num$
Uncrowded Conditions, Park Street Station:	$dwelltime = 0.71 + 0.00052 * num$
Crowded Conditions, Typical Stations:	$dwelltime = 0.44 + 0.00195 * num$
Crowded Conditions, Park Street Station:	$dwelltime = 0.71 + 0.00078 * num$

After we obtained detailed control line information for the Red Line, a simulation model was used to determine minimum safe headways and free running times. The short-turning parameters we used are estimates based on discussions with MBTA personnel.

4.1.3 Comparison of Control Strategies

We solved each of four problem instances using the "Hold All", "Hold Once", and "Hold at First" strategies of the Fixed Order Holding Problem with Capacity constraints (FOHPC), and as a Short Turning Problem with predetermined train order (STPP). We also solved the problems for a "Do Nothing" strategy in which no active controls were applied. Unfortunately, we do not have a simulation package which can correctly model train movements on the MBTA Red Line under its existing control system. Therefore, we assumed that a train will not depart from a station until it can travel to the next station at maximum permissible (free running) speed, with no inter-station stopping or deceleration. With this assumption, we may be overestimating the passenger waiting time for trains behind the blockage, because the minimum safe headway is slightly larger for the free running conditions than for the existing control system. However, we may be underestimating the passenger waiting time for trains ahead of the blockage, because the trains ahead are held back a little. We believe that passenger waiting times under the Do Nothing strategy are reasonably close to the actual waiting times.

To compare the effectiveness of various control strategies for a particular problem instance, we must compare the waiting times for the same set of passengers, i.e. the passengers arriving at each station k between a start time (call it T_{k1}) and a finish time (T_{k2}). Unfortunately, we can not simply compare the objective function values for the control strategies directly. The objective function includes the waiting time for passengers arriving at each station k between the last train departure prior to the disruption, and the departure time of the last train in the impact set. The former time is the same for each strategy; we will designate this time as T_{k1}. However, the departure times of the last train in the impact set are different for the various strategies, so that the sets of passengers considered in the objective functions are not always the same.

We divide the impact set of trains into two sets: the trains "ahead" are the trains ahead of and including the blocked train, while the trains "behind" follow the blocked train. Once the blockage has been cleared, the blocked train moves through the impact set of stations without being blocked (again) or held at any of the stations, using any of the strategies we considered. (There is one exception to this; see the discussion in Section 4.2). The departure times of the blocked train are approximately equal for each of the strategies, therefore they can be used as the finish times T_{k2} to compare the waiting time for passengers boarding the trains ahead. We are also interested, however, in studying the effectiveness of controls on trains behind the blockage. Fortunately, the last train in the impact set is not held for the Hold All strategy, and departs from each station in the impact set at approximately the same time as for the uncontrolled Do Nothing strategy, in each problem instance. We can compare the waiting times for the trains behind using these two strategies by setting T_{k1} and T_{k2} equal to the departure times of the blocked train and the last train from station k, respectively.

Since the Hold Once and Hold at First formulations were developed by adding constraints to the Hold All formulation, the two strategies can be no more effective than the Hold All strategy, and the Hold All results provide a lower bound on the passenger waiting time for the trains behind. The Hold All formulation provides no such bound for the short-turning strategy, however. To determine the effect of short-turning on the waiting time for the trains behind, we compared the departure times of the last train in the impact set for the short-turning strategy with the times for the last train in Do Nothing (T_{k2}). The times are either approximately equal, or the time is slightly earlier for the short-turning strategy (again, with one exception as discussed below). When the short-turning time is earlier than T_{k2}, we calculate the time interval between them and the expected number of passengers arriving during that interval, at each station k. Since trains behind the impact set are not controlled, we can then calculate the waiting time for these passengers. By adding this waiting time to the results from STPP, we can compare times for the same set of passengers as in the Hold All and Do Nothing strategies. Note that since the headways for trains following a blockage are typically small, the waiting time that must be added to STPP to equalize the passenger sets is also relatively small.

4.2 Solution of FOHPC and STPP

We used the modeling software package GAMS (V. 2.25) [4] to enter the model formulations and data sets, and the MIP solver in CPLEX (V. 3.0) [6] to find solutions using branch-and-bound, for the Fixed Order Holding Problem with Capacity constraints, and the Short-Turning Problem with Predetermined train order. The problems were solved on a Sun SPARC 20 workstation. The solution process was terminated when the objective function was guaranteed to be within 0.1% of the optimal value. In order to speed up the execution, we changed one of the CPLEX parameters from its default setting, and used a front-end algorithm to fix some of the binary variables; these modifications are discussed in Section 4.3.2.

We tested several different piecewise linear approximations for the objective function, varying n from 4 to 7. Since there was no significant penalty in running time, we selected the more accurate 7 segment approximation, which overestimates the true waiting time function by no more than 4% for headways between 1.5 and 21 minutes. At headways of 21-27 minutes, however, the approximation underestimates the true function by 0-11.5%.

4.2.1 Presentation of the Results

We considered four problem instances, based on delays of ten and twenty minutes at each of two incident locations. The incidents are illustrated in Figure 4-1.

Tables A-D below summarize the results for the four problem instances. "Total Waiting Time" is the passenger waiting time which corresponds to the departure times output by the model; it is calculated using the quadratic objective function (not the piecewise linear approximation used for the optimization). Times are given in passenger-minutes. "Time Ahead" is the waiting time for the trains up to and including the blocked train, and "Time Behind" is the waiting time for the trains behind the blocked train. For Hold Once and Hold at First, the sets of passengers boarding the trains behind are not the same as for the other strategies, therefore the "Total Waiting Time" and "Time Behind" values can not be compared and have been omitted from the tables. The "Savings" entries give the passenger-minutes saved, and the percentage savings, over the Do Nothing strategy. "Maximum Load" is the maximum number of passengers wishing to board a train in the impact set; any number higher than 1200 indicates that a train is unable to clear a platform, and passengers are left behind. The "Problem Size" is measured by the number of $d_{i,k}$ variables in the problem (train/station pairs in the impact set). "CPU Time" is the solution time in seconds, using the software and hardware described above. "Nodes" is the number of nodes visited during the branch-and-bound solution process, and "Iterations" is the number of simplex iterations required for that solution.

The train headways for each problem solution are given in Appendix B.

4.2.2 Incident 1

Incident 1 causes a delay in the time that Train 17, which is traveling northbound, can leave Station 20 (Harvard). We consider an impact set consisting of the 13 trains between Train 9 and Train 21, inclusive. The resulting station impact set includes Stations 12 (Andrews Northbound) through 27 (Central Southbound). For the Short-Turning Problem, Train 18 crosses over from Station 19 to Station 27. The objective function for Incident 1 includes waiting time at stations southbound from the impact set to Station 43 (from the impact set to the dispatching terminals). Weighting factors $U_o = 0.5$ and $U_b = 0.25$ are used for the trunk and branch stations beyond the impact set. We use $U_{pen} = 0.1$ to weight the penalty for headway variation, and $U_{ahd} = 10$ and $U_{bhd} = 5$ for the people left behind on the platforms by overcrowded trains.

Table A: Incident 1, Ten Minute Delay

Incident 1, Ten Minute Delay	Do Nothing	FOHPC			STPP
		Hold All	Hold Once	Hold at First	Hold All
Total Waiting Time	15993	13627			14609
Time Ahead	11202	8863	8931	8961	9997
Savings (minutes)		2338	2270	2240	1204
Savings (percentage)		15%	14%	14%	8%
Time Behind	4791	4763			4753
Savings (minutes)		28			39
Savings (percentage)		0%			0%
Maximum Load	988	603	614	666	603
Problem Size ($d_{i,k}$)		95	95	95	88
CPU Time (seconds)		22	37	21	16
Nodes		93	247	84	114
Iterations		2779	5180	2355	1848

As indicated in Table A, each of the active control strategies we considered results in significant savings of passenger waiting time for Incident 1 with a ten minute delay. However, the savings are realized entirely by passengers who board either the blocked train or one of the trains ahead of it. There are virtually no savings for passengers boarding trains behind, for either the Hold All or the short-turning strategy. Since Hold Once and Hold at First can be no more effective than Hold All, there can be no savings behind for these strategies either.

It is quite possible that Hold Once and Hold at First actually result in negative savings, or a penalty, for the trains behind. This is because we assumed "free running" conditions in the model; trains travel between stations at maximum permissible speed with no inter-station stopping or deceleration. Since the current MBTA Red Line control system does not operate this way, the model uses holding to meet this requirement. In the Hold Once and Hold at First strategies, trains must be held only once so that there is no blocking at any point within the impact set; this results in excessive holding of trains behind the blockage. In the Do Nothing strategy, however, trains are blocked (in a sense, "held") for just long enough at each station that they can run freely to the next station. Thus, we have imposed a constraint in the Hold Once and Hold at First strategies, which could conceivably result in waiting time *increases* over the Do Nothing strategy at stations behind the blockage. This problem can be avoided by including only trains ahead of the blockage in the impact set.

As shown in Table A, the short-turning strategy is far less effective than simply holding for this problem instance. The delay at Station 20 is only ten minutes, and the time required to short-turn Train 18 is six minutes after it unloads its passengers at Station 19. As a result, Train 10 is delayed at Station 26 for longer than its optimal holding time because of Train 18's short-turn. Note that if the time to notify Train 18's operator of the short-turn (T_{not}) is reduced from 3 minutes to 1 minute, the total waiting time for STPP is reduced to 14437 minutes - still greater than without the short-turn.

The Hold All strategy is only slightly more effective than the Hold Once or Hold at First strategies. The passenger waiting time for trains ahead is nearly equivalent for the three strategies, with savings of 14-15% of the total waiting time for the impact set. The Hold Once and Hold at First strategies require far less holding - only once per train - than the Hold All strategy. Since there is no benefit to holding trains behind the blockage for Incident 1, then the impact set should be reduced to only the blocked train and the trains ahead, in which case the majority of benefits can be realized with a minimum of controls for the Hold at First and Hold Once strategies. Note that if we consider only passenger waiting time within the reduced impact set, the active control strategies save 20-21% over Do Nothing.

The train headways are given in Appendix B. As indicated, the active control strategies result in far more even headway distributions than the Do Nothing strategy, particularly at the stations just beyond the blockage (21,22, and 23). The headway distribution for the Hold All strategy becomes somewhat less "regular" at Station 27; this may occur because the headways are adjusted to minimize waiting times on the branches as well.

The execution times for this problem instance are short enough to be considered "real-time" in a decision suppo[illegible]pplication.

Table B: Incident 1, Twenty Minute Delay

Incident 1, Twenty Min. Delay	Do Nothing	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	STPP Hold All
Total Waiting Time	46087	24767			21523
Time Ahead	36868	16934	17306	17385	16836
Savings (minutes)		19934	19563	19483	20032
Savings (percentage)		43%	42%	42%	43%
Time Behind	9218	7833			6842
Savings (minutes)		1386			2377
Savings (percentage)		3%			5%
Maximum Load	1646	666	759	805	651
Problem Size ($d_{i,k}$)		95	95	95	88
CPU Time (seconds)		25	82	27	17
Nodes		123	220	95	76
Iterations		3018	10977	3308	2020

Table B indicates that the active control strategies we considered result in very large waiting time savings for Incident 1 with a twenty minute delay. As before, the vast majority of savings is for passengers boarding the trains ahead, for whom the three holding strategies are, again, nearly equivalent. The waiting time savings for these passengers range from 42-43% for the three holding strategies, and the short-turning strategy as well. In contrast to the ten minute delay, there are also small savings for the trains behind the blockage using either the Hold All or short-turning strategies.

The time to execute the short-turn is not as critical a factor for the twenty-minute delay as it was for the ten minute delay. The short-turn does not provide any significant benefit, however, because the number of stations (and passengers) outside the short-turning loop is large, whereas the number of passengers benefitting from the short-turn (i.e. only those waiting at Station 27 and beyond) is small. The largest savings for the trains behind occurs for the short-turning strategy, where the delay behind the blockage is effectively shortened by one headway when Train 18 is short-turned off the northbound track.

When any of the active strategies is used, the trains are always able to clear the platforms of passengers. In contrast, there are over 400 passengers left behind Train 17 at Station 26 for the Do Nothing strategy, and passengers are left behind at Stations 25 (by Train 17) and 27 (Trains 17 and 18) as well.

For Incident 1 with a twenty minute delay, the majority of benefits can be realized with a minimum of controls if the impact set is reduced to the trains ahead, and Hold at First or Hold Once is used.

The running time for this problem instance using the Hold Once strategy was significantly longer than for the other strategies. A discussion of execution time issues is given in Section 4.3.2.

4.2.3 Incident 2

Incident 2 causes a delay in the time that Train 26, which is traveling northbound, can leave Station 7 (Quincy Adams). We consider an impact set consisting of the 11 trains between Train 19 and 29 inclusive. The resulting station impact set includes Stations 1 (Ashmont) through 15 (Park Street Northbound). Recall that we do not hold any trains before they reach the dispatching terminal, for a blockage located past the dispatching terminal. For the Short-Turning Problem, Train 32 (from outside the impact set) crosses over from Station 38 to 8. Thus, the additional waiting time for the passengers at Stations 38 and 39, as a result of the short-turn, is not included in the objective function value output by the model. An estimate of this time must be added in order to compare the strategies. The objective function for Incident 2 includes waiting time at stations northbound from the impact set up to Station 23 (Alewife), with a weight $U_o^{NB} = 0.5$, and southbound stations on the trunk with $U_o^{SB} = 0.25$. We use $U_{pen} = 0.1$ to weight the penalty for headway variation, and $U_{ahd} = 10$ and $U_{bhd} = 5$ for the people left behind on the platforms by overcrowded trains.

If the blockage occurs on a branch line heading towards the junction, the division between "trains ahead" and "trains behind" is unclear for trains located on an unblocked branch. For Incident 2, the trains on the unblocked Ashmont line are considered "behind" until they enter the trunk line, at which point they are clearly either "ahead" or "behind" Train 26. One can view the junction as the effective blockage location for these trains.

Table C: Incident 2, Ten Minute Delay

Incident 2, Ten Minute Delay	Do Nothing	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	STPP Hold All
Total Waiting Time	38088	28421			28419
Time Ahead	32495	23101	24465	25327	23016
Savings (minutes)		9394	8031	7186	9479
Savings (percentage)		25%	21%	19%	25%
Time Behind	5593	5320			5404
Savings (minutes)		273			189
Savings (percentage)		<1%			<1%
Maximum Load	1336	1137	964	985	776
Problem Size ($d_{i,k}$)		69	69	69	78
CPU Time (seconds)		17	274	23	12
Nodes		110	1587	138	37
Iterations		3365	50254	3985	1489

In Section 4.1.3, we referred to an exception in which the blocked train does not proceed without blocking immediately after the blockage is cleared. This exception occurs for the short-turning strategy in Incident 2 with a ten minute delay. Train 26 is delayed for 1.7 minutes after the incident has been resolved, because it must wait for Train 32 to execute its short-turn. As a result, the departure times for Train 26 (the blocked train) and Train 28 (the last train in the impact set for the Braintree branch stations) are later for the short-turning strategy than for Hold All or Do Nothing. The waiting time for the passengers who arrive during the incremental time interval must be subtracted from the short-turning results in order to equalize the passenger sets for comparison.

Table C indicates that the active control strategies result in large waiting time savings for Incident 2 with a ten minute delay. Once again, the time is saved almost entirely by passengers boarding one of the trains ahead; there are negligible benefits from applying controls to the trains behind the blockage.

The effectiveness of the three holding strategies varies more for Incident 2 than for Incident

1; savings range from 19-25% for the trains ahead. The impact set for Incident 2, unlike that for Incident 1, includes trains which merge together from the two branches. The Hold All strategy can be more effective than the other holding strategies because the headways for trains can be continuously adjusted both before and after the merge. In contrast, if trains can be held only once, it is not always possible to have even headways both before and after the merge (refer to Appendix B for the headway distributions). We re-ran this problem instance with a reduced impact set including only the trains ahead of the blockage. For this case, the Ashmont trains which have not entered the junction prior to the disruption (Trains 22 and 24) are not held until they pass the junction (Station 12), i.e. the "first" station in Hold at First is Station 12. The passenger waiting time savings for the trains ahead with the reduced impact set is 22% for either Hold at First or Hold Once, as compared to 25% for Hold All.

For this problem instance, there is no waiting time benefit to short-turning. Although few people are inconvenienced by the short-turn, the time it takes Train 32 to execute the short-turn is too long when compared to the length of the delay at Station 7. Train 26 is delayed even further as it waits for Train 32 to complete its crossover. When Train 32 finally heads northbound, it has a large headway in front of it, with Train 26 following tightly behind.

When any of the active strategies is used, the trains are always able to clear the platforms of passengers. In contrast, there are over 100 passengers left behind Train 26 at Station 12 for the Do Nothing strategy, and passengers are also left behind at Stations 10, 11, and 13.

These results suggest that, again, the majority of benefits can be realized with a minimum of controls if either the Hold Once or Hold at First strategy is implemented on a reduced impact set which includes only the trains ahead.

For the implementation used here, the execution time for Hold Once on this problem is too long to be used in a real-time decision support system. A discussion of execution time issues is given in Section 4.3.2.

Table D: Incident 2, Twenty Minute Delay

Incident 2, Twenty Min. Delay	Do Nothing	FOHPC			STPP
		Hold All	Hold Once	Hold at First	Hold All
Total Waiting Time	94977	55102			44208
Time Ahead	88204	48978	52620	55487	38244
Savings (minutes)		39226	35584	32717	49960
Savings (percentage)		41%	37%	34%	52%
Time Behind	6773	6124			5964
Savings (minutes)		649			809
Savings (percentage)		<1%			<1%
Maximum Load	1653	1422	1343	1307	1200
Problem Size ($d_{i,k}$)		69	69	69	78
CPU Time (seconds)		25	2458	763	62
Nodes		328	8134	5121	832
Iterations		4684	426384	123302	11115

As indicated in Table D, the active control strategies result in very large waiting time savings for Incident 2 with a twenty minute delay. As for all of the problems we considered, the time is saved almost entirely by passengers boarding one of the trains ahead. The effectiveness of the three holding strategies varies for the reasons discussed above, with savings ranging from 34-41% for the trains ahead. When the impact set is reduced to include only the trains ahead, so that the "first" control station for Trains 22 and 24 is Station 12, then the savings for either Hold Once or Hold at First are 40% - nearly equivalent to the Hold All savings.

The short-turning strategy is by far the most effective for this problem instance. There is ample time for Train 32 to cross over ahead of the blockage; the large headways ahead of the blockage are reduced significantly (see the headway distribution in Appendix B). Furthermore, the load on Train 32 before the short-turn and the demand at Stations 38 and 39 is quite small, so few people are inconvenienced by the short-turn. The short-turning strategy is the only one for which the trains could always clear the platforms. However, the maximum number of people left behind with any of the holding strategies is less than half that for Do Nothing.

As for all of the test problems, the majority of benefits can be realized with a minimum of controls if the impact set is reduced to the trains ahead, and if the short-turning strategy is implemented with the modification that trains are held only once.

With the implementation used here, the execution times for the Hold Once and Hold at First strategies are too long to be used in a real-time decision support system.

4.2.4 Findings

The following items summarize the key findings from our model application:

- The results of the tests on the four problem instances are very encouraging. The active control strategies result in significant savings of passenger waiting time.
- The vast majority of benefits are realized by passengers boarding either the blocked train, or one of the trains ahead of it. In fact, the Hold Once and Hold at First strategies could result in increased waiting time for the trains behind. Therefore, the impact set should include only the trains ahead.
- When the impact set includes only the trains ahead of the blockage, the Hold Once and Hold at First strategies are nearly as effective as the Hold All strategy, and require significantly fewer control actions.
- Short-turning should be considered if (1) the length of the blockage is significantly longer than the time required to execute the short-turn, and (2) the number of stations outside the short-turning loop is small, so that the number of passengers who will benefit from the short-turn is large in comparison to the number who will be inconvenienced.
- As implemented above, FOHPC and STPP were not always solved quickly enough to be considered "real-time". This issue is addressed in Section 4.3.2.

4.3 Implementation Issues

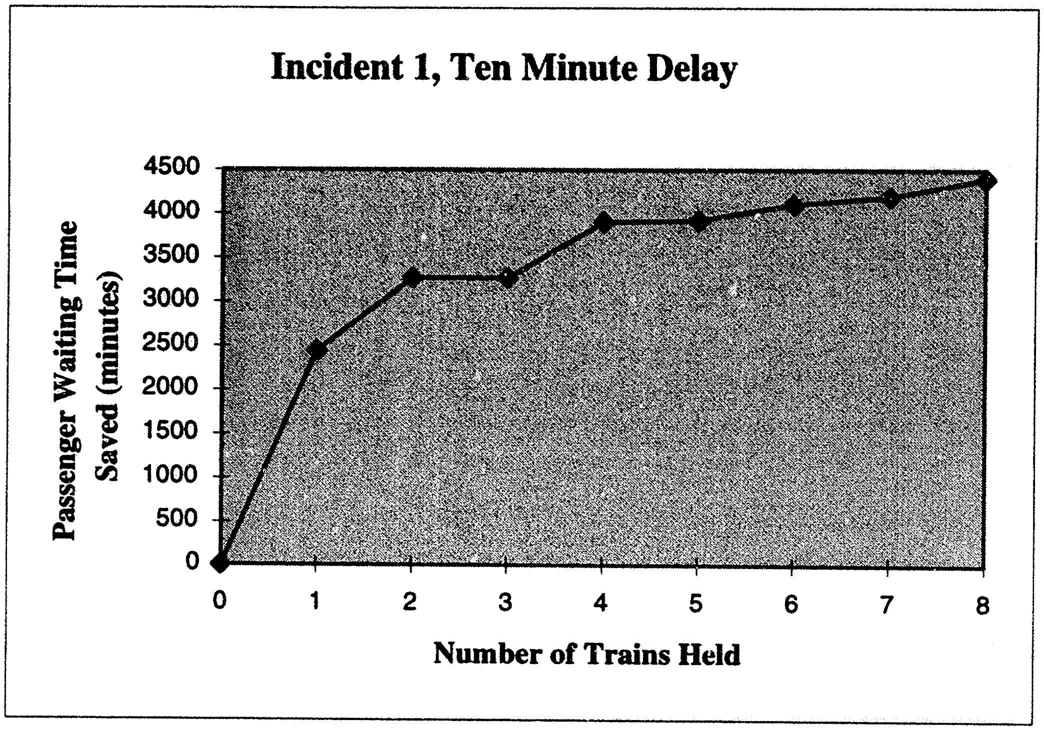
4.3.1 Impact Set Size

As discussed in Section 3.1.3, there are several factors which can be used to define an upper bound on the impact set size used in applying the System G model formulations. These include the necessity of a real-time solution, the infeasibility of controlling a large set of trains, and the stochastic nature of the system. However, larger impact sets should result in more effective control strategies. Therefore, we investigated the effect of varying impact sizes on the passenger waiting time savings for Incident 1, using the "Hold All" strategy of the Fixed Order Holding Problem. Since we demonstrated in Section 4.2 that there is little or no benefit to holding trains behind the blockage, we considered holding only the trains ahead.

We varied the number of trains held ahead of the blockage from 0 to 8. For example, we first held no trains, then held only Train 16, then held only Trains 15 and 16, etc. until we held Trains 9-16. Under each scenario, we calculated the passenger waiting time for Trains 9 through 17 (the blocked train) at Stations 20 and beyond, for both ten and twenty minute delays. We applied the same weighting factors beyond Station 27 as we applied in Section 4.2. The passenger waiting time savings for varying impact set sizes in Incident 1 are shown in Figure 4-2. The majority of the benefits are realized if only a few trains are held. The waiting time savings for two trains and for four trains are approximately 75% and 90% of the savings for eight trains, respectively. The shape of the plots in Figure 4-2 suggests that the marginal return for holding more than eight trains is small.

4.3.2 Execution Time

Tables A-D in Section 4.2 indicate the execution times for each problem instance using GAMS (V. 2.25) and CPLEX (V. 3.0) on a Sun SPARC 20 workstation. When we first solved the problems, we directly implemented the formulations given in Chapter 3, and used the default settings in CPLEX. Unfortunately, the execution times were often much longer than those shown in Tables A-D. To speed up the execution, we made two modifications as described below.



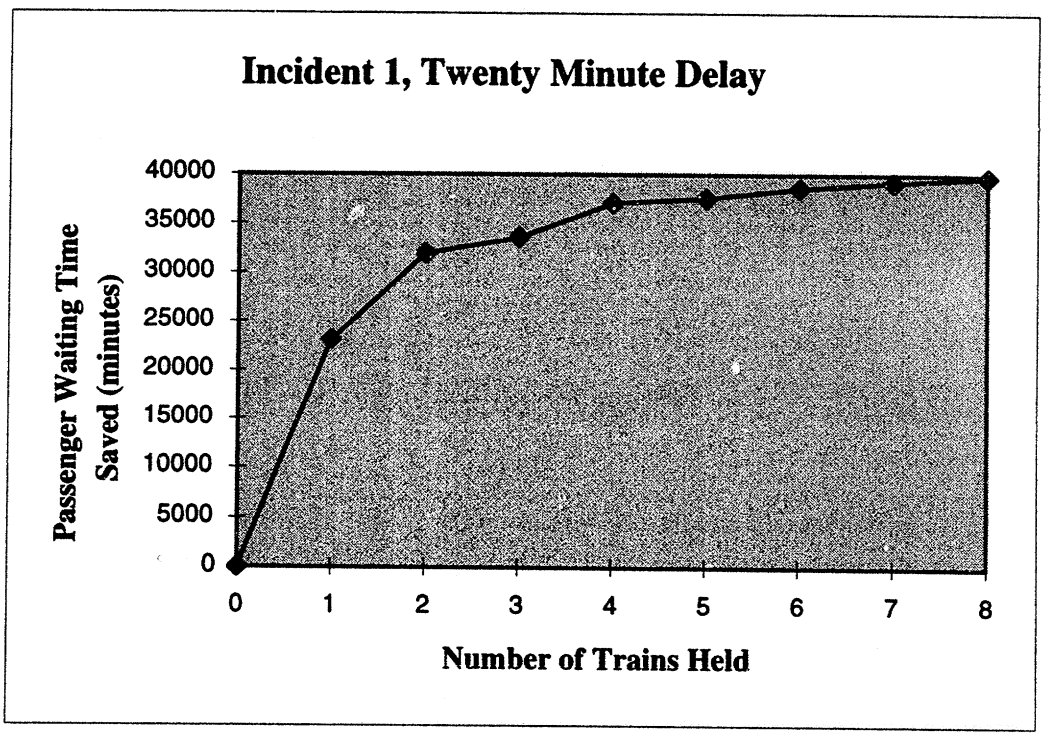


Figure 4-2: Varying Impact Set Sizes

First, we set the CPLEX MIP branch strategy parameter to -1. This parameter controls which branch (up or down) is taken first at each node in the branch-and-bound. By setting the parameter to -1, the down branch is taken first. There are only two sets of integer variables in FOHPC and STPP: $v_{i,k}$, which is used in each of the strategies, and $w_{i,k}$, which is used only in the Hold Once strategy. Recall that $v_{i,k} = 0$ unless train i is loaded to capacity as it departs station k, and $w_{i,k} = 0$ unless station k is the one holding station for train i. Thus, these two binaries are equal to zero far more often than they are equal to one, and we typically reach optimality more quickly by branching on zero first.

Second, we used a simple front-end algorithm to fix many of the $v_{i,k}$ variables before performing the optimization. The idea behind the algorithm is that for a given delay, we can identify a number of train/station pairs such that train i will not be loaded to capacity as it departs station k, no matter which control strategy - including Do Nothing - is applied. We iteratively calculate an upper bound for each $l_{i,k}$ as:

$$l_{i,k}^{\max} = A_k\left(h_{i,k}^{\max}\right) + (1 - Q_k)\, l_{i,k-1}^{\max} + p_{i',k} \tag{4.1}$$

where $h_{i,k}^{\max}$, the maximum possible headway for train i, is equal to the (estimated) length of the delay at the blockage plus the scheduled operating headway. If the upper bound $l_{i,k}^{\max}$ is less than the train capacity $L_{\max}$, we set $v_{i,k} = 0$ before optimizing. Thus, we effectively reduce the number of binary variables, to which the MIP execution time is extremely sensitive.

The execution times given in Tables A-D in Section 4.2 reflect the two modifications described above. Although ten of the sixteen cases given in Tables A-D ran in less than 30 seconds, several ran too slowly to be used in a real-time decision support system. The Hold Once formulation, which has the most binary variables, ran particularly slowly. However, the problems we ran are very large; the impact sets of 11 - 13 trains result in 69 - 95 $d_{i,k}$ variables. Since we demonstrated above that there is little benefit to holding trains behind the blockage, we reduced the impact sets to include only trains ahead of the blockage, and re-ran the six cases for which the execution time exceeded 30 seconds. The execution times were reduced to less than 30 seconds in all but one case (Hold Once for Incident 2 with a twenty minute delay) which ran in 53 seconds. For this most difficult case, we used a refined version of the front-end algorithm to fix several additional $v_{i,k}$ variables, and reduced the execution time to 34

Inc.	Delay	Problem	Strategy	Prob. Size ($d_{i,k}$)	No. of $v_{i,k}$ Fixed	CPU Time (sec)	Prob. Size ($d_{i,k}$)	No. of $v_{i,k}$ Fixed	CPU Time (sec)
1	10 Min.	FOHPC	Hold All	95	58	22			
		FOHPC	Hold at First	95	58	21			
		FOHPC	Hold Once	95	58	37	43	18	11
		STPP	Hold All	88	53	16			
1	20 Min.	FOHPC	Hold All	95	49	25			
		FOHPC	Hold at First	95	49	27			
		FOHPC	Hold Once	95	49	82	43	16	29
		STPP	Hold All	88	45	17			
2	10 Min.	FOHPC	Hold All	69	49	17			
		FOHPC	Hold at First	69	49	23			
		FOHPC	Hold Once	69	49	274	41	32	24
		STPP	Hold All	78	54	12			
2	20 Min.	FOHPC	Hold All	69	29	25			
		FOHPC	Hold at First	69	29	763	41	20	16
		FOHPC	Hold Once	69	29	2458	41	20/22	53/34
		STPP	Hold All	78	30	62	50	21	11

Table 4.1: Execution Times

seconds. (In the refined version of the algorithm, we consider that the sum of the headways for train i and its predecessor train ii at station k can be no greater than $h_{i,k}^{\max}$ plus the scheduled operating headway.) A summary of execution time results is given in Table 4.1.

4.3.3 On-Board Time

The objective in the System G formulations is to minimize the passenger waiting time. An alternate objective would be to minimize the total passenger time, including the on-board time. We chose to minimize only waiting time for several reasons. First, the effect of holding and short-turning on on-board time is relatively small, when compared to the effect on passenger waiting time. With the exception of the time spent by passengers who are on-board trains while they are being held, the effects of the control strategies on on-board time are generally second order. Second, on-board time is generally assumed to be less onerous than waiting time. Finally, the on-board time terms must be calculated as the (typically) non-linear product of load ($l_{i,k}$) and time, which greatly reduces the tractability of the problem.

Fortunately, if we neglect second order effects, we can include the on-board time in our

Incident	Delay	Objective Function	Passenger Time		
			Waiting	On-Board	Total
1	10 Min.	PWT	8961	1543	9578
		TPT	9074	271	9182
1	20 Min.	PWT	17385	2372	18334
		TPT	17659	806	17982
2	10 Min.	PWT	23411	8666	26877
		TPT	23702	5920	26070
2	20 Min.	PWT	50018	17617	57065
		TPT	51201	10488	55396

Table 4.2: Passenger On-Board Time

objective function when the Hold at First strategy is used. The load on each train at the time of the disruption is a constant, given as input data to the problem. We multiply this constant by $(1 - Q_k)$ and the holding time at the first station to calculate the on-board time as a linear function of the holding time.

For the Hold at First formulation of FOHPC, we solved each of the four problem instances with two different objective functions: (1) the minimization of passenger waiting time (PWT), and (2) the minimization of total passenger time (TPT), including the on-board time. In each case, we calculated the resulting passenger waiting, on-board, and total time. Since waiting time is considered more onerous than on-board time, the latter was weighted by a (widely-used) factor of 2.5 in the objective function. Our results are summarized in Table 4.2; the total time is calculated as the weighted sum of the waiting and on-board time. Note that the impact set was reduced to include only trains ahead of the blockage.

When the on-board time was included in the objective function, the total weighted passenger time was reduced by less than 5%. Although the on-board time was reduced by 40 -80%, the larger passenger waiting time was increased by 1 - 9%.

Chapter 5

Conclusions and Future Research

5.1 Summary and Conclusions

5.1.1 Analysis of a Simplified System

In Chapter 2 of this thesis, we considered a simple, idealized model of a rail system. "System S" consists of a set of trains operating on a one-way loop line at even headways prior to a disruption which temporarily blocks traffic. Several assumptions were used in the analysis of System S. First, the passenger arrival rate was assumed constant and equal at each station. In addition, the dwell time at each station and the passenger load on each train were assumed constant. Finally, we assumed that train capacity is not a binding constraint in System S; this is reasonable for some systems during non-peak hours when the disruption is not very long.

We considered two bounding cases in System S. In the first, there are long layovers at the terminal station, implying that delays do not propagate beyond the terminal, while in the second, there are no layovers at the terminal station, implying that delays propagate without any damping. We considered four holding strategies, ranging from "do nothing" (Strategy IV) to holding all trains in the system for a time equal to the blockage duration (Strategy I). We also analyzed a strategy in which trains are either short-turned from behind a blockage, or into the gap in front of a blockage. For the various strategies, we determined the increases in passenger waiting time and on-board time which result from the disruption, as a function of various system parameters.

We found that the "do nothing" strategy is always less effective than at least one of the active control strategies considered. For systems with layovers between round-trips, we found that the most successful holding strategy is one in which an optimal number of trains is held in front of, and behind, the blockage in order to even out consecutive headways as much as possible (Strategies II or III). This optimal number is a function of the number of stations between the blockage and the terminal, and the ratio of trains to stations in the system; it is independent of the blockage duration. We found that there is a minimum number of trains which must be held behind the blockage or they will be blocked; this number depends on the ratio of the minimum safe headway to the standard operating headway. For systems without layovers, the total passenger waiting time becomes less for Strategy I than for Strategy II or III as the length of the blockage or number of remaining round trips increase, or as the number of trains in the system decreases.

In systems for which Strategy II is the preferable holding strategy, short-turning is beneficial only if the number of stations skipped by the short-turning train is quite small. For systems without layovers in which Strategy I is preferable to Strategy II, the passenger waiting and on-board time can be further reduced by short-turning.

5.1.2 Generalized System Model

In Chapter 3, we developed a more realistic (albeit deterministic) generalized model of a rail system. The "System G" model applies to transit systems with branches as well as to simple loop lines. The passenger arrival rate and alighting fraction, train running times, and minimum safe headway between trains are all station-specific parameters. The dwell time is modeled as a function of the number of passengers boarding and alighting and the crowding conditions on the train. The order of trains need not remain constant, but may change due to merging at a junction or short-turning. Finally, the capacity of the trains may be a binding constraint in System G, thus peak hour operations can be reasonably modeled.

In order to make the optimal control problem more tractable, System G is subject to a number of limitations. The number of passengers boarding, alighting, and on-board the trains are estimated from stored data. Many stochastic elements are approximated by their expected values, and the dwell time function is assumed to be linear. Only a limited "impact set" of trains and stations is considered. Finally, the controls output by the model constrain the trains to operate in a "free-running" manner, so that they are separated by sufficient headways to travel at maximum permissible speed between stations, with no inter-station stopping or deceleration.

The objective function used in the System G problem formulations is the minimization of passenger waiting time, a quadratic function of the headways between trains. We developed a piecewise linear approximation to the quadratic function so that the problems can be solved as linear and mixed integer programs.

We developed a series of mathematical programming formulations for the holding problem, in which one of several holding strategies is used to mitigate the effects of a short-term blockage. The strategies we considered include: (1) Hold All, in which each train can be held at any or all of the stations in the impact set, (2) Hold at First, in which each train is held at the next station in reaches after the disruption occurs, and (3) Hold Once, in which each train can be held at only one optimally chosen station.

We developed a formulation for the Fixed Order Holding Problem without capacity constraints (FOHP). In the fixed order problem, the order of trains entering a junction from separate branches is given. Without capacity constraints, this formulation applies only when the passenger load on the train is well below the train capacity, as might occur during off-peak operations. When the piecewise linear approximation is used for the objective function, FOHP is a linear program. We then extended the Fixed Order Holding Problem to include capacity constraints (FOHPC). The introduction of these constraints necessitated the use of binary variables; FOHPC is thus a mixed integer programming problem. We also developed a formulation for the Variable Order Holding Problem, in which we allow trains to deviate from their scheduled order as they merge at a junction. All versions of the Variable Order Holding Problem are mixed integer programs.

We developed a formulation for the Short-Turning Problem with Predetermined Train Order (STPP), in which the final train order after the short-turn is given. We also discussed extending this formulation to the problem of undetermined train order, in which the decision of when and if a train short-turns is determined during the optimization. These formulations are based on combining short-turning with holding, and can be written using any of the holding strategies discussed earlier. Without capacity constraints, STPP can be written as a linear program. The undetermined order problem is a rather complicated integer program, with or without capacity constraints.

5.1.3 Model Application

In Chapter 4, we applied the generalized system formulations to four problem instances based on the MBTA Red Line. For each problem, we determined the passenger waiting time using five different control strategies: the Hold All, Hold Once, and Hold at First version of FOHPC; the Hold All version of STPP; and a Do Nothing strategy.

The results of our tests are very encouraging; the active control strategies result in significant passenger waiting time savings. For the two incidents we considered, we were able to reduce waiting time by 15-25% for a ten minute delay, and by more than 40% for a twenty minute delay. The vast majority of the time was saved by passengers boarding either the blocked train, or one of the trains ahead of it. Therefore, we recommend that the impact set include only trains ahead of the blockage.

For trains ahead of a blockage, the Hold Once and Hold at First strategies are virtually as effective as the Hold All strategy, and require significantly fewer control actions. Among holding strategies, we recommend that the simplest Hold at First strategy be used. Short-turning should be considered if the number of stations outside the short-turning loop is small, and the length of the blockage is significantly longer than the time required to execute the short-turn. Note that our results for short-turning in System G reaffirm what we found for System S.

We experimented with various impact set sizes, and found that the majority of the waiting time savings can be realized if only a small set of trains is held. For the incidents we studied, there is little marginal benefit to holding more than eight trains; in fact, the waiting time savings for two trains is approximately 75% of the savings for eight trains.

We investigated an alternate objective function in which on-board time is minimized along with waiting time. Such an objective function is non-linear for all problems unless the Hold at First strategy is used. For the four problems we tested, the total weighted passenger time is reduced by less than 5% when the on-board time was included in the objective function, thus there is little penalty for omitting it. However, since the Hold at First strategy is the recommended strategy, the alternate objective function can be easily used.

When we solved a direct implementation of the four test problems using the default settings in CPLEX, several of the cases ran too slowly to be considered real-time. However, by reducing the impact sets to a reasonable size, and using a simple front-end algorithm to fix some of the binary variables in advance, we were able to reduce the execution time of each problem to approximately 30 seconds or less. It appears that the mathematical programming formulations we developed may be solved quickly enough to be used as part of a real-time decision support system for rail transit systems. We believe this thesis represents a significant step towards the development of such a system.

5.2 Recommendations for Future Research

There is still a tremendous amount of work to be done in developing real-time decision support systems for rail transit. This thesis suggests four major areas of research which are direct extensions:

1. The models we developed are deterministic; many elements were approximated by their expected values. These include the train running and dwell times, passenger arrival rates and alighting fractions, and the estimated length of the disruption. Intuitively, we expect that the solutions to the System G formulations are not particularly sensitive to variance in these elements, with the possible exception of the length of the disruption. Before significant work is done towards developing stochastic versions of our models, we recommend that sensitivity analyses be performed on each of the stochastic elements to assess the validity of a deterministic approximation.

2. We investigated only two of the most commonly used control strategies: holding and short-turning. This work can be extended to include other strategies, in particular expressing.

3. Although our work indicates that the holding and short-turning problems we considered can be solved as mathematical programming problems in real-time, the solutions we found are rather simple and intuitive. It is quite likely that near-optimal results can be achieved easily through heuristics, requiring a smaller investment in software (i.e. CPLEX) and,

perhaps, hardware. Furthermore, heuristics may be required when controls such as expressing are used.

4. We developed formulations for use during and after minor service disruptions. However, the formulations (in particular, the fixed order holding problem without capacity constraints) can easily be modified for use during routine operations control.

Appendix A

Arrival Rates / Alighting Fractions

Table A.1 in this Appendix gives the estimated arrival rates (in passengers per minute) and alighting fractions at 8:15 AM, for each station on the MBTA Red Line.

	Northbound			Southbound				
	Stn.	Arrival	Alighting	Stn.	Arrival Rates			Alighting
Station Name	No.	Rate A_k^o	Fraction	No.	A_k^o	A_k^1	A_k^2	Fraction
Ashmont	1	27.30	0.000	1	0.00			1.000
Shawmut	2	3.97	0.001	43	0.03			0.158
Fields Corner	3	9.27	0.023	42	0.25			0.415
Savin Hill	4	3.73	0.005	41	0.18			0.067
JFK / Ashmont	5	1.78	0.070	40	0.43			0.528
Braintree	6	12.52	0.000	6	0.00			1.000
Quincy Adams	7	25.30	0.003	39	0.00			0.660
Quincy Center	8	21.02	0.005	38	0.10			0.626
Wollaston	9	15.13	0.002	37	0.38			0.096
North Quincy	10	19.83	0.060	36	0.55			0.494
JFK / Braintree	11	4.22	0.032	35	1.15			0.282
Andrews	12	10.22	0.009	34	0.00	0.37	0.70	0.065
Broadway	13	3.95	0.017	33	0.10	0.52	0.96	0.113
South Station	14	24.50	0.256	32	0.49	0.82	1.54	0.645
Downtown Crossg	15	18.45	0.442	31	14.30	2.08	3.89	0.439
Park Street	16	19.37	0.577	30	17.38	1.20	2.24	0.378
Charles	17	5.43	0.166	29	3.77			0.060
Kendall	18	1.05	0.420	28	3.83			0.105
Central	19	2.70	0.268	27	21.17			0.038
Harvard	20	1.67	0.766	26	31.67			0.125
Porter	21	0.50	0.239	25	23.85			0.005
Davis	22	0.77	0.386	24	29.77			0.002
Alewife	23	0.00	1.000	23	31.33			0.000

Table A.1: Arrival Rates and Alighting Fractions, 8:15 AM

Appendix B

Headways

The tables in this Appendix give the preceding headways (in minutes) for the impact set of trains, for each of the four problem instances discussed in Chapter 4.

Figure B-1: Incident 1, Ten Minute Delay

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
27	9	3.25	2.74	2.88	1.74	4.40
	10	3.63	3.12	3.17	2.93	4.40
	18					3.25
	11	3.25	3.25	3.25	2.51	3.05
	12	4.50	4.53	3.25	3.97	3.05
	13	3.25	3.45	3.25	1.96	2.25
	14	3.63	3.53	4.25	5.26	3.22
	15	3.44	3.89	4.16	2.58	3.22
	16	3.44	3.87	4.17	3.05	3.22
	17	3.25	3.25	3.25	8.52	3.22
	18	2.65	2.86	3.04	1.86	
	19	3.09	2.87	3.05	2.05	3.22
	20	2.25	2.25	2.25	2.51	3.22
	21	2.25	2.25	2.25	3.05	2.25
26	10	3.86	4.07	4.25	2.89	5.30
	11	3.86	3.25	3.25	2.53	6.86
	12	3.86	4.49	3.25	3.92	3.22
	13	3.86	3.48	3.26	2.04	3.22
	14	3.86	3.53	4.22	5.14	3.22
	15	3.86	3.88	4.17	2.68	3.22
	16	3.44	3.87	4.17	3.03	3.22
	17	3.26	3.27	3.28	8.32	3.22
	18	2.68	2.87	3.05	2.11	
	19	2.68	2.87	3.05	2.05	3.22
	20	2.68	2.27	2.28	2.50	3.22
	21	2.25	2.25	2.25	3.03	2.28
25	11	4.04	4.42	4.61	2.56	8.25
	12	4.04	4.42	3.24	3.84	3.25
	13	4.04	3.52	3.26	2.16	3.25
	14	4.04	3.52	4.17	4.94	3.25
	15	4.04	3.87	4.17	2.83	3.25
	16	3.46	3.87	4.17	3.01	3.25
	17	3.28	3.31	3.34	7.99	3.25
	18	2.71	2.90	3.06	2.43	
	19	2.68	2.87	3.05	2.11	3.25
	20	2.68	2.31	2.33	2.48	3.25
	21	2.28	2.25	2.25	3.00	3.10
24	12	4.21	3.79	5.26	3.79	7.62
	13	4.21	4.75	3.26	1.98	3.74
	14	4.21	4.75	4.13	5.08	3.41
	15	4.21	3.86	4.17	2.90	3.41
	16	4.21	3.87	4.17	3.01	3.41
	17	3.29	3.33	3.37	7.81	3.41
	18	2.71	2.91	3.07	2.18	
	19	2.71	2.87	3.05	2.56	3.25
	20	2.68	2.33	2.35	2.47	3.25
	21	2.29	2.26	2.26	2.98	3.25

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
23	13	4.31	4.71	4.75	1.96	5.79
	14	4.31	4.75	4.10	5.04	3.74
	15	4.31	3.90	4.17	3.00	3.74
	16	4.31	3.87	4.17	3.00	3.74
	17	3.33	3.35	3.40	7.59	3.74
	18	2.72	2.93	3.09	2.44	
	19	2.72	2.88	3.05	2.55	3.74
	20	2.49	2.36	2.38	2.47	3.25
	21	2.49	2.26	2.26	2.96	3.25
22	14	5.89	5.89	5.89	5.25	5.59
	15	5.89	5.89	5.89	1.96	5.59
	16	5.83	5.82	5.82	5.04	5.59
	17	4.75	4.76	4.76	10.12	5.59
	18	2.51	2.45	2.45	2.44	
	19	2.49	2.55	2.64	2.55	2.62
	20	2.49	2.56	2.56	2.47	2.47
	21	2.49	2.78	3.13	2.52	2.47
21	15	6.88	6.88	6.88	2.00	6.27
	16	5.82	5.82	5.82	2.43	5.59
	17	4.75	4.75	4.75	13.02	5.59
	18	2.51	2.46	2.46	2.46	
	19	2.49	2.54	2.63	2.54	2.62
	20	2.49	2.56	2.57	2.48	2.48
	21	2.49	2.78	3.13	2.51	2.48
20	17	13.00	13.00	13.00	13.00	13.00
	18	2.48	2.48	2.48	2.48	
	19	2.54	2.54	2.62	2.54	2.62
	20	2.49	2.56	2.58	2.48	2.48
	21	2.50	2.79	3.13	2.51	2.48
19	18	12.45	12.45	12.45	12.45	
	19	2.48	2.48	2.48	2.48	12.45
	20	2.54	2.54	2.71	2.54	2.62
	21	2.49	2.86	3.13	2.48	2.48
18	19	12.10	12.42	12.38	12.10	6.15
	20	2.48	2.53	2.78	2.48	5.95
	21	2.54	2.90	3.12	2.54	2.62
17	19	7.13	6.30	12.41	6.30	6.30
	20	7.13	8.76	2.96	7.96	3.98
	21	2.48	3.00	3.12	2.48	3.98
16	19	6.18	6.18	12.21	6.18	6.18
	20	3.99	8.73	3.10	2.97	3.81
	21	5.72	3.08	3.12	6.74	3.26
15	19	5.68	5.68	11.61	5.68	5.68
	20	3.87	8.70	3.25	3.05	3.87
	21	3.25	3.17	3.12	4.14	3.25
14	20	3.99	8.66	9.28	3.25	3.99
	21	3.25	3.32	3.11	4.02	3.25
13	21	4.55	9.28	9.67	4.55	4.55
12	21	4.80	4.80	9.89	4.80	4.80

Figure B-2: Incident 2, Twenty Minute Delay

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
27	9	4.75	3.28	3.25	1.74	4.40
	10	4.75	3.28	3.25	2.93	4.40
	18					3.25
	11	4.75	4.28	3.25	2.51	4.05
	12	4.75	4.75	4.75	3.97	4.05
	13	3.92	4.75	4.77	1.96	3.25
	14	4.75	4.75	6.04	5.26	4.69
	15	4.75	6.14	5.94	2.58	4.69
	16	4.75	5.96	5.94	3.05	4.69
	17	4.75	4.75	4.75	18.89	4.40
	18	2.25	2.32	2.71	3.17	
	19	2.92	2.67	3.00	2.15	2.47
	20	2.25	2.50	2.50	1.86	2.47
	21	2.25	2.25	2.25	2.05	2.25
26	10	5.09	4.75	4.70	2.89	5.30
	11	5.09	4.25	3.25	2.53	6.86
	12	5.09	4.73	4.69	3.92	4.69
	13	5.09	4.75	4.76	2.04	4.69
	14	5.09	4.75	6.00	5.14	4.69
	15	5.09	6.09	5.94	2.68	4.69
	16	4.75	5.97	5.94	3.03	4.69
	17	4.75	4.80	4.79	18.82	4.42
	18	2.35	2.41	2.78	2.81	
	19	2.59	2.66	2.99	2.33	2.54
	20	2.58	2.51	2.52	2.15	2.47
	21	2.25	2.26	2.26	2.05	2.26
25	11	5.51	6.05	5.05	2.56	8.25
	12	5.51	4.68	4.60	3.84	4.91
	13	5.51	4.75	4.75	2.16	4.91
	14	5.51	4.75	5.94	4.94	4.75
	15	5.51	6.02	5.94	2.83	4.75
	16	4.77	5.97	5.94	3.01	4.75
	17	4.77	4.87	4.87	18.43	4.75
	18	2.49	2.56	2.91	2.50	
	19	2.58	2.65	2.98	2.81	2.66
	20	2.58	2.52	2.55	2.33	2.48
	21	2.27	2.28	2.28	2.15	2.27
24	12	5.93	8.09	7.02	3.79	7.62
	13	5.93	4.75	4.74	1.98	6.88
	14	5.93	4.75	5.90	5.08	5.22
	15	5.93	5.97	5.94	2.90	5.22
	16	5.93	5.97	5.94	3.01	4.75
	17	4.77	4.91	4.90	18.32	4.75
	18	2.58	2.64	2.98	2.06	
	19	2.58	2.64	2.98	2.50	2.74
	20	2.58	2.52	2.56	2.81	2.49
	21	2.28	2.28	2.29	2.33	2.28

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
23	13	6.45	6.88	7.93	1.96	6.46
	14	6.45	6.88	5.85	5.04	6.46
	15	6.45	5.92	5.94	3.00	6.46
	16	6.45	5.97	5.94	3.00	6.46
	17	4.82	4.96	4.95	17.63	4.75
	18	2.63	2.75	3.07	2.61	
	19	2.63	2.64	2.98	2.38	2.83
	20	2.58	2.53	2.58	2.50	2.50
	21	2.30	2.29	2.30	2.60	2.29
22	14	8.50	7.46	7.46	5.25	8.50
	15	8.50	7.46	7.46	1.96	8.50
	16	8.50	10.00	10.00	5.04	8.50
	17	6.88	7.47	7.47	20.16	6.88
	18	2.63	2.47	2.47	2.44	
	19	2.63	2.55	2.80	2.55	2.83
	20	2.58	2.74	2.47	2.50	2.50
	21	2.30	2.45	3.19	2.60	2.29
21	15	10.00	10.00	10.00	2.00	10.00
	16	10.00	10.00	10.00	2.43	10.00
	17	7.45	7.45	7.45	23.03	7.45
	18	2.65	2.49	2.49	2.48	
	19	2.62	2.55	2.78	2.55	2.84
	20	2.58	2.72	2.49	2.50	2.51
	21	2.29	2.47	3.19	2.60	2.29
20	17	23.00	23.00	23.00	23.00	23.00
	18	2.51	2.51	2.51	2.51	
	19	2.55	2.55	2.76	2.55	2.66
	20	2.52	2.71	2.51	2.51	2.49
	21	2.57	2.49	3.19	2.59	2.51
19	18	22.45	22.45	22.45	22.45	
	19	2.51	2.51	2.51	2.51	22.45
	20	2.55	2.55	2.76	2.55	2.66
	21	2.52	2.71	3.20	2.51	2.49
18	19	22.10	22.45	22.35	22.10	9.46
	20	2.51	2.47	2.88	2.51	12.65
	21	2.55	2.82	3.20	2.55	2.66
17	19	9.76	6.30	22.19	6.30	6.30
	20	14.50	18.53	3.21	17.96	5.32
	21	2.51	3.10	3.22	2.51	12.65
16	19	6.18	6.18	21.84	6.18	6.18
	20	5.21	18.34	3.47	2.97	3.81
	21	14.50	3.34	3.23	16.74	3.26
15	19	5.68	5.68	21.09	5.68	5.68
	20	4.06	18.12	3.72	3.05	3.87
	21	4.33	3.60	3.25	4.14	3.25
14	20	4.22	17.68	19.07	3.25	3.99
	21	4.24	4.12	3.34	4.02	3.25
13	21	4.55	19.05	19.64	4.55	4.55
12	21	4.80	4.80	19.79	4.80	4.80

Figure B-3: Incident 2, Ten Minute Delay

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
15	19	5.68	5.68	5.68	5.68	5.68
	20	4.83	4.06	4.75	2.37	4.75
	21	4.83	4.75	4.75	4.87	4.75
	22	4.83	4.74	4.75	2.84	4.75
	23	4.83	4.75	4.75	3.01	4.75
	24	4.83	4.75	4.75	2.82	4.75
	25	4.83	4.75	4.90	4.90	4.58
	32					3.25
	26	4.01	4.76	3.72	13.21	3.10
	27	2.87	2.94	3.16	3.09	2.86
	28	2.84	2.89	3.18	3.00	2.86
	29	2.82	2.82	2.79	2.85	2.82
14	20	4.75	4.22	4.88	2.60	4.75
	21	4.75	4.65	4.66	4.69	4.75
	22	4.75	4.83	4.86	3.04	4.75
	23	4.75	4.66	4.65	2.91	4.75
	24	4.75	4.87	4.82	2.93	4.75
	25	4.75	4.46	4.65	4.71	4.55
	32					3.34
	16	4.28	4.75	3.79	12.54	3.13
	17	3.29	3.26	3.29	3.81	3.00
	18	2.85	2.92	3.17	3.04	2.86
	29	2.83	2.86	3.01	2.93	2.84
13	21	5.32	4.55	6.79	4.55	5.25
	22	5.01	6.46	4.93	3.19	4.75
	23	5.01	4.59	4.58	2.84	4.75
	24	5.01	4.96	4.87	3.02	4.75
	25	4.75	4.25	4.46	2.81	4.75
	32					4.67
	16	4.75	4.73	3.85	13.75	3.16
	17	3.62	3.51	3.39	4.36	3.10
	18	2.87	2.94	3.16	3.09	2.86
	29	2.84	2.89	3.18	3.00	2.86
12	21	5.15	4.80	7.02	4.80	4.94
	22	5.15	6.45	4.95	2.16	4.94
	23	5.15	4.60	4.57	3.19	4.94
	24	5.15	4.97	4.87	3.71	4.94
	25	4.75	4.24	4.45	2.15	4.75
	32					4.67
	16	4.74	4.73	3.85	14.36	2.86
	17	2.54	3.53	3.40	2.36	2.47
	18	3.25	2.94	3.16	4.21	3.10
	29	2.87	2.90	3.19	3.09	2.86

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
11	23	9.20	11.05	11.75	2.93	8.74
	25	9.20	9.20	9.31	5.94	8.74
	32					6.88
	26	6.88	4.72	3.86	16.73	2.89
	28	2.87	2.88	6.57	2.93	2.88
10	25	12.32	17.26	18.07	5.92	11.39
	32					10.00
	26	9.97	4.75	3.90	16.65	2.91
	28	2.94	2.92	6.57	2.99	2.89
9	25	11.99	16.78	17.56	5.79	11.10
	32					10.05
	26	10.00	5.10	4.30	16.34	2.51
	28	2.55	3.02	6.51	2.69	2.91
8	32					15.25
	26	16.10	16.10	16.10	16.10	2.40
	28	2.40	3.07	6.46	2.54	2.51
7	26	16.00	16.00	16.00	16.00	17.69
	28	2.35	3.48	6.77	2.35	2.21
6	28	12.12	13.61	16.77	12.12	13.81
5	22	10.39	11.43	12.17	7.19	9.97
	24	10.39	9.59	9.44	6.89	9.97
	27	11.69	12.52	11.72	18.68	14.47
	29	5.79	5.85	6.36	6.56	5.56
4	24	10.00	6.94	14.46	6.94	10.00
	27	7.50	10.00	11.69	9.06	9.44
	29	10.00	15.30	6.42	11.79	10.00
3	24	9.98	6.94	14.41	6.94	9.98
	27	6.88	9.98	11.70	9.04	6.88
	29	7.08	7.02	6.45	7.98	7.08
2	24	9.92	6.91	14.28	6.91	9.92
	27	6.88	9.94	11.73	9.01	6.92
	29	7.12	7.06	6.54	7.99	7.08
1	27	9.00	9.92	19.07	9.00	9.00
	29	8.00	7.08	6.57	8.00	8.00

Figure B-4: Incident 2, Twenty Minute Delay

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
15	19	6.40	5.68	5.68	5.68	5.82
	20	6.40	4.61	5.53	2.37	5.82
	21	6.40	4.75	5.74	4.87	5.82
	22	6.40	4.66	6.82	2.84	5.82
	23	6.40	6.04	6.88	3.01	5.82
	24	6.40	6.73	6.84	2.82	5.82
	25	6.40	7.56	4.95	3.01	4.75
	32					4.75
	26	3.95	9.16	6.54	25.54	3.99
	27	2.92	3.16	3.55	3.93	2.89
	28	3.39	3.28	3.57	3.62	2.94
	29	2.86	2.78	2.85	2.94	2.84
14	20	6.36	4.75	5.63	2.60	5.84
	21	6.36	4.66	5.63	4.69	5.84
	22	6.36	4.75	6.88	3.04	5.84
	23	6.36	5.87	6.76	2.91	5.84
	24	6.36	6.82	6.85	2.93	5.84
	25	6.36	7.13	4.75	2.89	4.77
	32					4.77
	16	4.75	9.08	6.47	24.45	3.81
	17	3.35	3.67	3.72	4.68	3.32
	18	3.12	3.20	3.56	3.62	2.90
	29	3.16	3.06	3.26	3.33	2.90
13	21	7.63	6.67	8.46	4.55	6.60
	22	6.88	4.82	6.91	3.19	6.60
	23	6.88	5.75	6.68	2.84	6.60
	24	6.88	6.88	6.88	3.02	6.60
	25	6.88	6.81	4.61	2.81	4.75
	32					4.78
	16	4.78	9.01	6.40	23.57	3.68
	17	3.69	4.07	3.86	4.98	3.64
	18	2.92	3.16	3.55	3.93	2.89
	29	3.39	3.28	3.57	3.62	2.94
12	21	7.07	4.80	8.68	4.80	6.66
	22	7.07	6.94	6.93	2.16	6.66
	23	7.07	5.74	6.67	3.19	6.66
	24	7.07	6.88	6.88	3.71	6.66
	25	7.07	6.77	4.61	2.15	4.75
	32					4.78
	16	4.79	9.02	6.38	24.17	3.68
	17	2.62	4.08	3.88	2.51	2.56
	18	3.25	3.18	3.56	4.67	3.25
	29	2.92	3.29	3.59	3.93	2.89

Station	Train	FOHPC Hold All	FOHPC Hold Once	FOHPC Hold at First	FOHPC Do Nothing	STPP Hold All
11	23	13.14	12.66	17.45	2.93	10.00
	25	13.14	13.62	11.55	5.94	10.00
	32					10.00
	26	9.10	9.10	6.38	26.55	5.04
	28	2.91	2.90	7.41	3.32	2.88
10	25	17.88	23.25	25.96	5.92	11.86
	32					11.54
	26	14.50	9.12	6.42	26.46	8.66
	28	2.96	2.96	7.42	3.32	2.93
9	25	17.38	22.59	24.96	5.79	11.55
	32					11.55
	26	14.81	9.60	7.22	26.39	8.69
	28	2.56	2.92	7.31	2.83	2.53
8	32					15.25
	26	26.52	26.52	26.52	26.52	10.68
	28	2.18	2.65	6.99	2.18	2.37
7	26	26.00	26.00	26.00	26.00	26.00
	28	2.77	3.36	7.56	2.77	2.18
6	28	22.12	23.49	27.54	22.12	22.12
5	22	13.44	11.91	15.78	7.19	13.40
	24	13.44	12.62	13.55	6.89	13.40
	27	15.65	19.79	14.77	28.49	15.43
	29	5.87	6.61	7.30	7.18	5.81
4	24	11.91	6.94	22.14	6.94	11.91
	27	10.00	12.62	14.72	9.06	10.00
	29	15.65	24.23	7.39	21.61	15.34
3	24	10.00	6.94	22.05	6.94	10.00
	27	6.88	12.58	14.75	9.04	6.88
	29	7.06	4.36	7.44	7.98	7.06
2	24	9.93	6.91	21.82	6.91	9.93
	27	6.88	12.50	14.84	9.01	6.88
	29	7.11	4.49	7.56	7.99	7.11
1	27	9.00	12.47	29.69	9.00	9.00
	29	8.00	4.54	7.60	8.00	8.00

Bibliography

[1] Araya, S., K. Abe and K. Fukomori, 1983. An Optimal Rescheduling for Online Train Traffic Control in Disturbed Situations, *Proceedings of the 22nd IEEE Conference on Decision and Control*, Vol.1, p.489-494.

[2] Barnett, A., 1974. On Controlling Randomness in Transit Operations. *Transportation Science*, Vol. 8, No.2, p.102-116.

[3] Barnett, A., 1978. Control Strategies for Transport Systems with Nonlinear Waiting Costs. *Transportation Science*, Vol. 12, No.2, p.119-136.

[4] Brooke, A., D. Kendrick and A. Meeraus, 1992. *Gams: A User's Guide, Release 2.25*, Scientific Press, South San Francisco, California.

[5] Central Transportation Planning Staff, 1991. *1989 Passenger Counts: MBTA Rapid Transit and Commuter Rail*, Boston, MA.

[6] CPLEX Optimization, Inc., 1994. *Using the CPLEX Callable Library, Version 3.0.*

[7] Eberlein, Xu Jun, 1995. *Real-Time Control Strategies in Transit Operations: Models and Analysis*, Ph.D. Dissertation, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

[8] Komaya, K. and T. Fukuda, 1989. Estrac III: An Expert System for Train Traffic Control in Disturbed Situations, *IFAC Control, Computers, Communications in Transportation*, Paris, France.

[9] Lin, Tyh-ming, 1990. *Dwell Time Relationships for Urban Rail Systems*, S.M. Thesis, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

[10] Union Signal & Switch, 1992. *OCC Final Design Review Document - T92COS.*