Shuttle Planning for Link Closures in Urban Public Transport Networks

Evelien van der Hurk[*1], Haris N. Koutsopoulos[2,4], Nigel Wilson[2], Leo Kroon[1], and
Gábor Maróti[3]

[1]Rotterdam School of Management and ECOPT, Erasmus University Rotterdam, P.O. Box 1738,
NL-3000 DR Rotterdam, The Netherlands
[2]Civil & Environmental Engineering, Massachusetts Institute of Technology, 77 Massachusetts
Avenue, Cambridge, MA, 02139
[3]VU University Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, The Netherlands
[4]KTH the Royal Institute of Technology, Stockholm, Sweden

August 25, 2014

Abstract

Urban Public Transport systems must periodically close certain links for maintenance, which can have significant effects on the service provided to passengers. In practice, the effects of closures are mitigated by replacing the link with a simple shuttle service. However, alternative shuttle services could reduce inconvenience at lower operating cost. This paper proposes a model to select shuttle lines and frequencies under budget constraints. A new formulation is proposed that allows a minimal frequency restriction on any line that is operated, and minimizes passenger inconvenience cost, including transfers and frequency-dependent waiting time. This model is applied to a shuttle design problem based on a real world case study of the MBTA network of Boston (USA). The results show that additional shuttle routes can reduce passenger delay in comparison to the standard industry practice, while also distributing delay more equally over passengers, at the same operating budget. The results are robust under different assumptions about passenger route choice behavior. Computational experiments show that the proposed formulation, coupled with a preprocessing step, can be solved faster than prior formulations.

1   Introduction

Rail systems such as Boston's metro, London's underground, and Netherlands' passenger rail network must periodically deal with link closures and capacity limitations due to maintenance. Such link closures can cause significant delays for passengers which in turn can have long-term effects on their perceptions of the service. Passengers' appreciation of the public transport system is often an important performance measure in granting concessions to operate the network. Therefore limiting the negative effect of these disruptions is extremely important for public transport operators.

---

[*]E-mail corresponding author: ehurk@rsm.nl

Operators' standard procedure when facing these major disruptions is to replace the closed link with a shuttle service. However, additional shuttle services in the vicinity of the disrupted area may reduce passenger inconvenience at similar operating cost.

This paper studies the Shuttle Planning for Link Closures ($SPLC$) problem that concerns the location and frequency design of shuttle lines for link closures. The research is meant to support operators in minimizing passenger inconvenience under budget constraints when facing link closures. The aim is to develop a model that can solve realistic cases comprising a large number of Origin Destination pairs (OD-pairs) fast.

We propose a new mixed integer formulation for the $SPLC$. Key features of the model are that it includes a minimum operating frequency restriction for all candidate shuttle lines, and frequency-dependent passenger inconvenience costs such as transfers and waiting time costs, and it allows for a change in frequencies for both the existing network and the shuttle lines. A path reduction process is proposed that reduces the problem size significantly. Computational experiments indicate that the new formulation, together with the path reduction process, is able to solve realistic problems with large numbers of OD-pairs quickly.

The practical relevance of the proposed model is demonstrated based on a real world case study. The results indicate that (1) solutions for realistic sized problems with a large number of OD pairs can be generated fast, (2) allowing the selection of shuttles beyond the disrupted area, and allowing changes of frequency in the full network, can reduce both passenger inconvenience and operating cost, (3) inconvenience of the closure is distributed more evenly over passengers, and worst case delays are reduced, and (4) solutions are relatively robust with respect to different assumptions on passenger behavior.

The three key contributions of this paper are summarized as:

- a novel mixed-integer formulation for the $SPLC$ that a) allows to specify a minimum operating frequency for lines and b) includes frequency dependent passenger inconvenience costs.

- the proposition of a path reduction process that a) reduces problem size and therefore b) allows including large numbers of OD-pairs in the model.

- demonstration that the proposed methodology proposes practically relevant solutions quickly for realistic problem sizes based on a real world case study.

The proposed model and path reduction step may also be applicable for more general line planning problems. Moreover, computational experiments indicate that the solution speed may be high enough to use for real time occurrence of link closures due to disruptions as well.

The remainder of this paper is organized as follows. Section 2 provides a problem description, Section 3 presents related work, and the problem formulation is described in Section 4, together with the proposed model. Two important preprocessing steps, to reduce problem size and increase speed, are presented in Section 5. Section 6 discusses the results of the application of the proposed model to a real world case study. Finally Section 7 summarizes the paper and draws conclusions.

2 Problem description

Consider the public transport network in Figure 1 which consists of two lines: line 1 connecting stations A to G and line 2 connecting stations S to Z. A link closure between stations E and X disconnects the northern branch of line 1 from the rest of the network. Replacement shuttles are needed to restore the network connectivity while providing sufficient capacity and minimizing the inconvenience experienced by passengers due to the closure.

Standard practice introduces a single new shuttle line reconnecting stations affected by the closure (line 3 in Figure 1). This default route is easy to implement as the required capacity of the shuttle can be estimated from the expected demand on the closed link, and passengers can easily find the replacement shuttle by following their standard route. However, when the majority of trips originates beyond the disrupted area, this introduces two additional transfers for most passengers, where the additional travel time depends on the frequency of both the shuttle line and the regular line in addition to the extra running time of the shuttle.

Other shuttle lines could be more convenient for passengers and have similar operating cost. For example, if stations D and T are major demand generators, the opening of an additional shuttle line (line 4 in Figure 1b) could significantly reduce passenger inconvenience by providing a faster connection and reducing the number of transfers. The attractiveness of such a line will depend on the shuttle frequency which determines the waiting time of passengers boarding the line, and on the number of transfers of passengers using this line. The proposed model specifically includes frequency dependent transfer and waiting time cost.

(a) Existing network

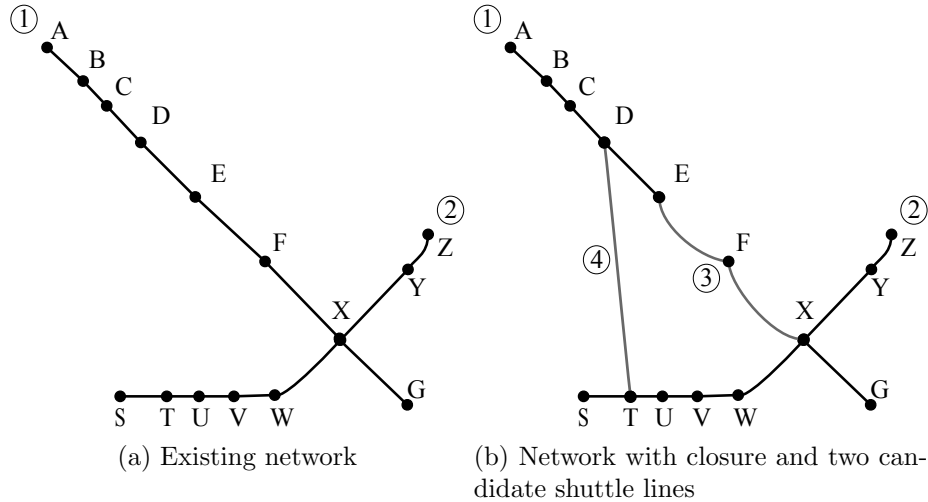(b) Network with closure and two candidate shuttle lines

Figure 1: Example of a public transport network graph

The *SPLC* model determines the optimal set of shuttle lines and their frequencies to minimize passenger inconvenience within a budget constraint given passenger demand, a transportation network, a set of candidate shuttle lines with minimum and maximum frequencies, and a link closure. Alternatively the operating cost could be weighted against the passenger inconvenience. The budget is defined as a maximum number of vehicles, equal to the number of vehicles needed for the standard solution. Passenger inconvenience is measured by the route assignment of passenger demand to paths, including transfers, in the public transport network.

The *SPLC* model simultaneously assigns passengers to paths and selects frequencies for the lines, as these are interdependent. The attractiveness of a path depends on the service frequency, while the required frequency depends on the demand for that line. However, including passenger assignment in the optimization model would lead to minimizing inconvenience for all passengers, instead of for each individual. In a system with free route choice and capacity constraints, the model's passenger assignment may therefore differ from the actual passenger flows. Consequently, the assignment of passengers to paths is restricted to a set of *reasonable* paths in the optimization model. A path is reasonable if it cost is within a small increment of the standard path. Moreover, the solution's quality is evaluated under several different assumptions about passenger behavior.

3   Related work

Link closures can cause significant disturbances in public transport networks. The different models proposed to increase the robustness of public transport networks and timetables to relatively small delays (Cicerone et al., 2009; Fischetti and Monaci, 2009; Liebchen et al., 2010) are aimed at the planning phase. Link closures however, are not, and cannot, be taken into account in this planning of standard operations, as they occur infrequently and require significant alterations from normal operations. Therefore, they are considered within the broader category of real-time disruption management, even if they are planned. At the same time, the problem of minimizing passenger inconvenience under planned closures through the introduction of shuttles also has a strong link to the strategic problem of line planning.

Disruption Management: Disruption management aiming at minimizing passenger delay was first studied in the context of airlines. Lan et al. (2006) examined the problem of reducing passenger delay through the rerouting and retiming of flights. Jespersen-Groth et al. (2009) discuss disruption management in rail transportation focussing on the three sub-problems of adjusting the timetable, rescheduling crews, and rescheduling rolling stock. Initially research in the area of disruption management in high frequency public transport focused on the complex questions of how to reschedule resources. For instance, Nielsen et al. (2012) and Cacchiani et al. (2012) focus on rolling stock rescheduling in the case of disruptions. Potthoff et al. (2010) and Veelenturf et al. (2014) present research on crew rescheduling.

Recent focus is shifting to using passenger service quality explicitly as the objective. Both Kroon et al. (2014) and Cadarso et al. (2013) incorporate passenger rerouting in the optimization of capacity rescheduling. Kroon et al. (2014) present a model for rolling stock rescheduling. Cadarso et al. (2013) also include timetabling decisions. Both studies use minimization of passenger delay as the objective. Both papers assume that arrival times of passengers are based on the schedule and thus delay is defined by the deviation from the planned timetable. Veelenturf et al. (2013) extend the passenger-oriented disruption management approach for resource rescheduling by studying the benefit of altering the stop sequence of a line. They allow adding or removing a stop of a line. Cacchiani et al. (2014) provide an extensive overview of real time rescheduling in passenger rail transport, noting that most research is focused on small delays with little attention given to major disruptions, such as link closures.

Line Planning: The introduction of shuttle services to minimize the negative effects of a link closure is essentially a network re-design problem. Therefore, it is strongly linked to the line planning problem. Ceder and Wilson (1986) present a framework that divides the problem into two parts: the generation of a line pool, and the selection of lines from this pool. This approach is followed by most of the research in this area. An excellent overview can be found in Schöbel (2011).

Claessens et al. (1998) solve the line planning problem for the Dutch passenger railway network. They assume the line pool is given and the demand per link is fixed. They propose a branch and price method for selecting lines. Their formulation is unique in introducing separate, binary decision variables that not only represent the decision of which line to include, but also at what capacity to operate it.

In contrast to Claessens et al. (1998), both Schöbel and Scholl (2006) and Borndörfer et al. (2007) include continuous frequency variables, and include the dynamic routing of passengers. Both suggest a column generation approach to solve the model, replacing the multi-commodity flow model for the routing of passengers by a path formulation. This greatly reduces the number of decision variables needed to solve the problem, as instead of one decision variable per OD-group per edge in the network, a decision variable per path is included. Column generation requires solving the LP relaxation of the main model. However, this may result in opening some lines at a very low frequency, which makes them less attractive to passengers. These models do not include a relation between the frequency of a line and the cost of passengers traveling on the line, nor do they include a minimum frequency restriction conditional on whether the line is operated. Both are part of the proposed model in this paper.

Finally Kaspi and Raviv (2012) present an alternative heuristic approach aiming at overcoming the rounding problem while including the dynamic routing of passengers. The heuristic solves the line planning problem simultaneously with the timetabling problem, thereby minimizing passenger travel time including waiting time and transfers, as well as operating cost.

Shuttle planning for link closures and link failures: Pender et al. (2013), in their survey of disruption management practices, note that bus bridging is the most common approach to link closure or failure in rail networks. Pender et al. (2009) evaluated crossovers in the context of bus bridging and link closures, but did not look at the optimal selection of shuttle routes, which is the focus of this work. Kepaptsoglou and Karlaftis (2009) present a methodological framework for what they call 'the bus bridging problem', which is similar to the problem of shuttle planning for link closures. As is customary in planning, they split the problem into two parts: the generation of possible bus bridging routes and the capacity assignment to those routes. Their heuristic approach changes possible routes found through a shortest path method.

The work of Jin et al. (2013) is closest to this work, presenting a three-step procedure: a) generating routes using column generation; b) selecting the feasible routes; and c) assigning capacity to the selected routes. They show that adding 'non-intuitive' routes can significantly reduce passenger delay. They focus on routes starting and ending at the edges of the disruption, although the approach can be extended to include other stations. In their more recent work, Jin et al. (2014), steps (b) and (c) are integrated into a single optimization model that includes a modest time-tabling component for shuttle buses, and includes transfer-to-shuttle bus times in the calculation of passenger delay. The method was applied in a network of around a 100 nodes, and a limited set of OD pairs (26).

The contribution of the current paper is the development of a method for link closures with dynamic passenger routing that can solve real life problem instances with a *large number of OD-pairs* (1397) thanks to a *path reduction* pre-processing step. The proposed formulation combines the path-formulation of Schöbel and Scholl (2006) and Borndörfer et al. (2007) with the capacity formulation of Claessens et al. (1998), and adds a flexible capacity assignment. The formulation allows specifying a frequency-dependent path cost, thus including frequency-dependent passenger waiting time and transfer times. Furthermore, a minimum frequency restriction can be included on the condition that a line is operated, preventing lines being included at very low frequencies. Finally, the concept of *reasonable paths* is used to prevent the assignment of passengers to overly altruistic paths, and indeed, solutions prove relatively robust under different passenger behavior assumptions. Complex cases can be solved in one minute, making the model also a candidate for online application.

4   Problem formulation

The *SPLC* problem is static, with the link closure lasting for the full planning horizon. Full information about the location and duration of the closure, as well as the substitute shuttle services is available to passengers in advance, which is a natural assumption in the context of planned closures. The pool of candidate shuttle lines is given. Moreover, a set of candidate frequencies is given for both shuttle lines and existing lines. The passenger capacity of a segment of a line is defined by the geographical route and the selected frequency of the line. The *SPLC* model solution selects those frequencies for existing lines and shuttle lines that minimize passenger inconvenience at reasonable operating cost.

The demand matrix is given and fixed, and passengers are assumed to arrive randomly over time. The assumption of random arrivals is consistent with high-frequency networks that do not operate according to a published timetable, for which Frumin and Zhao (2012) find empirical support. A valid solution to the *SPLC* problem should provide sufficient capacity for all passengers.

To define the *SPLC* problem formally, we use the Public Transport Network graph, $\text{PTN}_{|\mathcal{F}|}$, defined in Section 4.1. Section 4.2 defines the operating cost (OC) and Section 4.3 defines passenger inconvenience, *PI*. Finally the model formulation is presented in Section 4.4. An overview of notation and terminology is provided in Table 1.

4.1   The Public Transport Network

Consider a set of geographical transit lines, *geolines* $\mathcal{G}$, containing *both* candidate shuttle lines and existing lines, and the set of all frequencies $\mathcal{F}$. For each geoline $g \in \mathcal{G}$ an ordered list of stops S_g, and a set of potential frequencies $\mathcal{F}_g \subseteq \mathcal{F}$ are defined. We extend the concept of geoline to include the option of walking from one station to the next. Walk arcs have infinite capacity, infinite frequency, and zero operating cost. Based on the set of geolines $\mathcal{G}$ and the set of frequencies $\mathcal{F}$ the directed *Public Transport Network Graph* ($\text{PTN}_{|\mathcal{F}|}$), $G(\mathcal{V}, \mathcal{E})$, is defined as follows.

The node set $\mathcal{V}_{line}$ contains a node for each direction of a geoline $g \in \mathcal{G}$, each stop of this line $s \in S_g$ and each frequency $f \in \mathcal{F}_g$. The node set $\mathcal{V}_{OD}$ contains an entry node and an exit node for each unique geographical stop $s \in S_g, g \in \mathcal{G}$. Thus we define the node set of the graph as $\mathcal{V} := \mathcal{V}_{OD} \cup \mathcal{V}_{line}$.

Table 1: Notation and terminology

Symbol	explanation	Symbol	explanation
$\text{PTN}_{|\mathcal{F}|}$	public transport graph with frequency set $\mathcal{F}$	PTN_f	public transport graph with a single frequency per line
$\mathcal{G}$	set of geolines	S_g	ordered list of stops of line $g \in \mathcal{G}$
$\mathcal{F}$	set of frequencies	δ_{lg}	arc capacity for line $g \in \mathcal{G}$ prvided by single vehicle $l \in \mathcal{L}$
$\mathcal{Q}$	set of OD-groups	w_q	passengers in OD-group $q \in \mathcal{Q}$
$\mathcal{P}$	set of paths in the PTN-graph	$\mathcal{P}_q$	path set for OD-group $q \in \mathcal{Q}$
c_p	cost of path p	c_g	cost per vehicle for geoline g
$\mathcal{P}(e)$	paths traversing arc $e, \mathcal{P}(e) \subseteq \mathcal{P}$	$\mathcal{E}_{gf}$	arcs associated with geoline g at frequency f
M_g	maximum passenger capacity line g	ν_{gf}	minimum number of vehicles for geoline g at frequency f
$\mathcal{L}$	set of vehicle types	$\mathcal{L}_g$	set of vehicle types accepted for geoline g
β	maximum number of shuttles		

The set of directed arcs $\mathcal{E}$ is composed of $\mathcal{E} := \mathcal{E}_{line} \cup \mathcal{E}_{transfer} \cup \mathcal{E}_{OD}$. The arc set $\mathcal{E}_{line}$ contains an arc for each consecutive pair of stops in S_g, for each direction of each geoline $g \in \mathcal{G}$, and each frequency $f \in \mathcal{F}$. A separate arc for each direction of the geoline is needed as capacity constraints are direction specific. The set of entry and exit arcs $\mathcal{E}_{OD}$ contains an arc connecting every entry node in $\mathcal{V}_{OD}$ to any line node in $\mathcal{V}_{line}$ serving the same geographical station, and an arc from any line node in $\mathcal{V}_{line}$ to the exit node in $\mathcal{V}_{OD}$ in the same geographical station. The set of transfer arcs $\mathcal{E}_{transfer}$ contains an arc for any pair of geolines with a stop at the same geographical station. Note that because the line nodes $\mathcal{V}_{line}$ are frequency and geoline specific, the transfer arcs in $\mathcal{E}_{transfer}$ are also frequency and geoline specific.
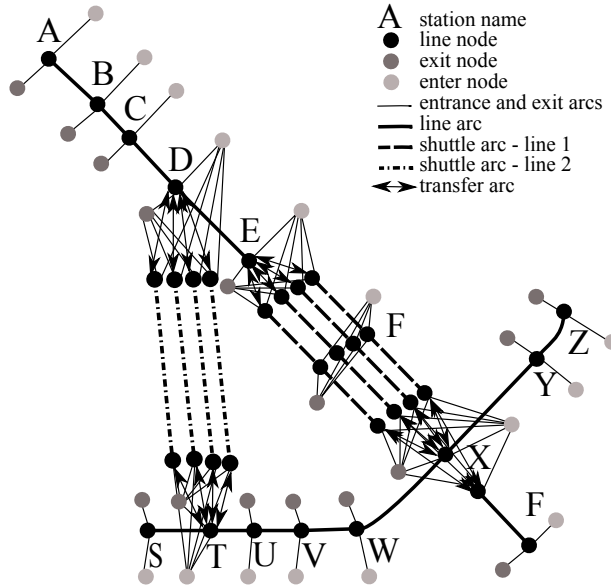

Figure 2: Example of a public transport network

Figure 2 displays a schematic representation of the public transport graph of Figure 1.

The network consists of 4 geolines. Each link in Figure 1 corresponds to *two* arcs: one for each direction of the line. There is a single frequency setting for the two regular lines, and four possible frequencies for the two shuttle lines. For each geographical shuttle line, each consecutive pair of stops on this line, and each possible frequency, a link is introduced in the public transport network graph. For each line node, there is an entry and exit arc from the respective entry and exit nodes of the physical station associated with the line node. Moreover, for nodes where transfers to other lines are possible, transfer arcs are introduced.

Throughout this paper $PTN_{|\mathcal{F}|}$ refers to the public transport network graph defined by the full set of frequency options: PTN_f refers to a graph with a single selected frequency for each geoline, which may be 0. The solution of the $SPLC$ defines a PTN_f.

4.2 Operating Cost

The operating cost (OC) is equal to the sum of the operating costs of all lines $g \in \mathcal{G}$. The operating cost of a line g is equal to the sum of the costs of all vehicles assigned to the line. Since geolines include existing lines, a change in frequencies on existing lines contributes to the OC. The objective of the $SPLC$ is to minimize Passenger Inconvenience using no more than the available operating budget.

Define $l \in \mathcal{L}$ as the set of vehicle types, and $\mathcal{L}_g$ as the subset of vehicle types that can be assigned to geoline g. Vehicle types distinguish between vehicles for different modes, vehicles with different seat capacities, and vehicles with different operating cost. Furthermore, define c_{lg} as the operating cost per vehicle of type l assigned to geoline g. These costs may be vehicle type and geoline specific.

Operating geoline g at frequency f requires a minimum number of vehicles per hour ν_{gf}, which depends on the run time of the geoline g, the turn around times for vehicles assigned to this line, and the frequency f. Therefore the operating cost of a geoline g at frequency f, defined as the sum of the cost over all assigned vehicles to this geoline, depends on the route, the frequency, and the type of vehicle assigned to it.

4.3 Passenger Inconvenience

The $SPLC$ model aims at reducing delay of passengers affected by the closure while providing sufficient capacity for all demand. Opened shuttle lines may attract passengers *not* affected by the closure if they provide a shorter route. The $SPLC$ model minimizes Passenger Inconvenience (PI) defined as the sum of the differences in costs of paths with and without the closure. As the costs of paths in the planned network are fixed, minimizing the costs of assigned paths in the network with the closure minimizes the *sum* of the differences in cost.

If the PI is minimized over all passengers, PI can be reduced by improving service for passengers *not* affected by the disruption. Therefore, in the $SPLC$ model 1) all passengers need to be assigned to a path, 2) this path needs to be *reasonable*, that is, the cost of the path is within a small increment of the standard solution's path cost, and 3) all candidate shuttle lines should benefit affected passengers. Consequently, passengers affected by the disruption are neither ignored nor significantly worse off than in the standard solution, and lines that only benefit passengers not affected by the disruption are excluded. The solutions for the case study successfully reduce delay of affected passengers. In the case

where solutions benefit passengers *not* affected by the disruption, the costs of paths could be adjusted to reflect delay instead of absolute costs. This could however, fail to capture all the demand attracted by shuttle lines in systems with free route choice.

The cost c_p of a path in the public transport network is equal to the sum of the costs of the arcs in the path representing waiting time, in-vehicle time, and transfers. Entry arcs in $\mathcal{E}_{OD}$ represent frequency dependent waiting time for the first vehicle given random arrivals of passengers. Costs of arcs in $\mathcal{E}_{line}$ are equal to the in-vehicle time between the two stops connected by the arcs. Arcs in $\mathcal{E}_{transfer}$ represent transfer costs, which are calculated as the expected waiting time to the transfer-to line as dependent on its frequency, plus a fixed transfer penalty.

Costs of arcs in a path can be mode, line, and station-specific. For instance, transfers at large stations can be penalized more than transfers at small stations, transfers to shuttles can be more costly than transfers to the same mode, and in-vehicle time cost can be geoline specific. Thus, the problem formulation allows for a realistic cost representation without additional complexity.

4.4 Model

Required input for the *SPLC* model is a set of geolines $\mathcal{G}$, a set of allowed frequencies $\mathcal{F}$, and a set of OD-groups Q. A single OD-group $q : (s_q, t_q, w_q, \mathcal{P}_q)$ is defined by an origin node s_q, a destination node t_q, a (demand) weight w_q, and a set of paths $\mathcal{P}_q$ connecting s_q to t_q in $\text{PTN}_{|\mathcal{F}|}$. The $\text{PTN}_{|\mathcal{F}|}$ is defined by $\mathcal{G}$ and $\mathcal{F}$. The set of paths $\mathcal{P}_q$ is obtained through the path generation method described in Section 5.1 and the path reduction method described in Section 5.2.

Continuous decision variables x_{pq} define the flow from OD-group q traveling on path p in the $\text{PTN}_{|\mathcal{F}|}$, with x_{pq} defined only for paths $p \in \mathcal{P}_q$, $q \in \mathcal{Q}$. Binary decision variables y_{gf} represent the decision of opening geoline g at frequency f. Decision variables v_{lg} define the number of vehicles of type l assigned to geoline g, alongside which we define $\mathcal{L}_g$ as the subset of vehicle types that can be assigned to geoline g. The choice of continuous or integer vehicle variables did not significantly affect computation time in our case study.

Some further notation: δ_{lg} is the maximum number of passengers that can be transported by a single vehicle of type l assigned to geoline g, dependent on the vehicle capacity of type l and the length of line g. M_g is the maximum number of passengers that can be transported on geoline g over all selections of frequency $f \in \mathcal{F}$ and assignment of vehicle types $l \in \mathcal{L}_g$, and β_l is the number of available vehicles of type l. Let $\mathcal{P}_e$ denote the set of paths traversing arc e, and $\mathcal{E}_{gf}$ denote the set of arcs representing geoline g at frequency f in $\text{PTN}_{|\mathcal{F}|}$.

The formulation of the *SPLC* problem is:

$$\min \sum_{q \in \mathcal{Q}} \sum_{p \in \mathcal{P}_q} c_p x_{pq} + \sum_{l \in \mathcal{L}} \sum_{g \in \mathcal{G}} c_{lg} v_{lg}$$

subject to:

$$\sum_{p \in \mathcal{P}_q} x_{pq} = w_q \qquad \forall q \in \mathcal{Q} \qquad (1)$$

$$\sum_{q \in \mathcal{Q}} \sum_{p \in \mathcal{P}(e)} x_{pq} \leq \sum_{l \in \mathcal{L}_g} \delta_{lg} v_{lg} \qquad \forall g \in \mathcal{G}, \forall e \in \mathcal{E}_g \qquad (2)$$

$$\sum_{q \in \mathcal{Q}} \sum_{p \in \mathcal{P}(e)} x_{pq} \leq M_g y_{gf} \qquad \forall g \in \mathcal{G}, \forall f \in \mathcal{F}_g, \forall e \in \mathcal{E}_{gf} \qquad (3)$$

$$\sum_{f \in \mathcal{F}_g} y_{gf} \leq 1 \qquad \forall g \in \mathcal{G} \qquad (4)$$

$$\sum_{l \in \mathcal{L}_g} v_{lg} \geq y_{gf} \nu_{gf} \qquad \forall g \in \mathcal{G}, \forall f \in \mathcal{F}_g \qquad (5)$$

$$\sum_{g \in \mathcal{G}} v_{lg} \leq \beta_l \qquad \forall l \in \mathcal{L} \qquad (6)$$

$$x_{pq} \geq 0 \qquad \forall q \in \mathcal{Q}, \forall p \in \mathcal{P}_q \qquad (7)$$

$$v_{lg} \geq 0 \qquad \forall l \in \mathcal{L}, \forall g \in \mathcal{G} \qquad (8)$$

$$y_{gf} \in \{0,1\} \qquad \forall y_{gf} \in \mathcal{L} \qquad (9)$$

Objective: The objective function minimizes expected passenger inconvenience and operating cost. By setting c_{lg} to zero, one can optimize passenger inconvenience, in which case the model selects the frequency for each line $f \in \mathcal{F}$ such that the route assignment in PTN_f has minimal passenger inconvenience over all possible $f \in \mathcal{F}$ under fleet size constraints.

Capacitated multi-commodity flow component: Constraint (1) requires that all passengers are assigned to a path. Constraint (2) restricts the number of passengers per line-arc not to exceed the capacity of vehicles assigned to a geoline g. Note that this constraint is not frequency specific. Therefore constraint (3) restricts passengers to only use geolines at their operated frequency f. This restriction is frequency dependent but not dependent on the number of vehicles assigned to the line, thus, both constraints (2) and (3) are needed to fully specify the capacity constraints.

Together constraints (1), (2) and (3) form the capacitated multi-commodity flow component of the model. A path formulation is chosen even though there exists an exponential number of paths in the graph and an arc-based formulation contains the large (but linear) number of $|\mathcal{E}| \times |\mathcal{Q}|$ decision variables. As the majority of existing paths will never be included in an optimal solution, for most practical applications the path-based formulation often leads to a significant reduction in the number of variables in comparison to the arc-based formulation. This however, requires the identification of the set of candidate paths, which we discuss in Section 5. The *SPLC* model could be solved through column generation, however, for the presented case study this was not needed to obtain optimal solutions fast.

Line selection component: Constraints (4) to (6) define restrictions on the selection of lines. Constraint (4) restricts the choice to one frequency f per geoline $g \in \mathcal{G}$. Thus, if a line is operated, it has to be at least at the minimum frequency in $\mathcal{F}$. Constraint (5)

forces the number of vehicles assigned to geoline g to be at least equal to the minimum number of required vehicles to operate the line at frequency f, ν_{gf}, which depends on the run time of geoline g. Lastly, equation (6) captures vehicle type dependent fleet size constraints.

The *SPLC* problem formulation contains two decision variables for the line selection: y_{gf} for the opening of geoline g at frequency f, and v_{lg} for the number of vehicles of type l assigned to geoline g. A single decision variable y_{gf} combines the choice of opening a geoline g with the selection of a frequency f, as proposed for railway line planning in Claessens et al. (1998). This enables the formulation of an MIP that 1) allows specifying a minimum frequency conditional on the opening of a line, and 2) can include frequency dependent passenger inconvenience costs, such as waiting and transfer costs. The usage of y_{gf} requires a discrete set of options, but there exists a continuous set of frequencies. Moreover, included paths p are frequency dependent. Therefore the problem size grows rapidly with the number of frequency options included.

The choice of the set of frequencies $\mathcal{F}$ can change the model solution in the formulation of the *SPLC* inspired by Claessens et al. (1998). Let us assume that the true optimal solution given a continuous set of frequencies includes a line $y_{g'f'}$. However, $\mathcal{F}$ includes only $y_{g'(f'-\varepsilon)}$ and $y_{g'(f'+\varepsilon)}$. Suppose $y_{g'(f'-\varepsilon)}$ provides insufficient capacity for all passenger demand, making it infeasible, and $y_{g'(f'+\varepsilon)}$ requires more vehicles than are available, making this solution also infeasible. In this case the model will propose a different solution, with different geolines than the true optimal solution.

Therefore, vehicle variables v_{lg} are introduced so that more capacity can be assigned to $y_{g'(f'-\varepsilon)}$, making this feasible. The y_{gf} variables define the *minimum* number of vehicles, and the waiting and transfer time of passengers boarding this line, and no longer the y_{gf} but the v_{lg} define the available (passenger) capacity and operating cost. The inclusion of different types of vehicles that can be assigned to one line can be included without the need to specify all possible combinations of assignments, as would be required in the formulation inspired by Claessens et al. (1998). Thus the problem of using vehicle variables can be solved using less frequency options, thereby greatly reducing the problem size, without the issues described above. This comes at the cost of slightly overestimating passenger transfer time and boarding time to lines where more vehicles are assigned than the minimum number required, as an assignment of more vehicles leads to higher frequency and thus lower *PI*, which is not included in the pre-computed path costs. By defining $\mathcal{F}$ based on small increments in headways, this difference could be kept small, and results in a more accurate estimation of *PI* than in the previous models of Borndörfer et al. (2007) and Schöbel and Scholl (2006) that do not include frequency-dependent path costs.

Decision variables: Decision variables for passenger flow assigned to a path and number of vehicles assigned to a line are restricted to be positive by constraint (7) and (8), respectively. The inclusion of a geoline g at frequency f is a binary decision variable due to constraint (9).

5 Solution Approach

This section defines two important pre-processing steps: Section 5.1 proposes an approach for the generation of a set of *reasonable* paths, which are required input for the model defined in Section 4.4. Section 5.2 presents a path reduction procedure that significantly

reduces the number of paths and increases the computational speed, without decreasing the quality of the solution. Together the generation of reasonable paths and the path reduction generate input for the *SPLC* model, which is then solved to optimality using CPLEX 12.6.

5.1 Path Generation

Path generation constructs the set of *reasonable paths* $\mathcal{P}_q$ for each OD-group in a given $\text{PTN}_{|\mathcal{F}|}$. The concept of *reasonable paths* follows Ceder and Wilson (1986), who limit the acceptable paths to some value above the absolute shortest path. We define a path as *reasonable* if its cost does not exceed the cost of the *standard path*, rather than the shortest path, by more than an increment α. The *standard path* is defined as the shortest path in the solution to the link closure closest to normal operations: a graph defined by the planned frequency of all lines, the closure, and the standard replacement shuttle around the closure at its maximum frequency. The standard path forms a natural reference point for both passenger path lengths and the operator's operating budget defined as the available fleet size.

The criteria and construction method that lead to a set of *reasonable* paths are defined for each OD-group, which contrasts with the global criteria generally used for column generation. A column generation approach will stop adding a new path p^* to the set of candidate paths $\mathcal{P}$ when there exists no path $p^* \in \text{PTN}_f$, $p^* \notin \mathcal{P}$ that would reduce the *overall* passenger inconvenience. This global condition would allow adding paths that are purely in the interest of the global social optimum but not in the interest of the OD-group itself. Because passengers are allowed to freely choose their route in the network (within certain limits), some of these altruistic paths p^* may be unrealistic in practice. Using the incremental cost α, one could consider the set of *reasonable paths* to be the set of paths between which passengers are indifferent, and thus exclude purely altruistic paths. Defining the path set per OD-group, and not at the system level, is a better reflection of the assumption of free route choice for each passenger. As an additional advantage, this specification allows the decoupling of the problem: the path set $\mathcal{P}_q$ can now be specified for each OD-pair separately.

The construction uses the concept of a *geopath*. Given a path $p \in \text{PTN}_{|\mathcal{F}|}$, the translation of this path to a *geopath* p^γ is defined by storing only the geoline information without the frequency information for each arc in the path. The translation of a path p^γ to the corresponding set of paths in $\text{PTN}_{|\mathcal{F}|}$ is defined by the set of all possible paths in $\text{PTN}_{|\mathcal{F}|}$ that contain the same geoarcs as path p^γ. These paths can be constructed by finding all frequency-specific arcs that match the geoarcs in p^γ, and generating from these arcs all possible paths that have the exact same ordering of geoarcs as p^γ. Thus given a set of *geopaths*, these paths can be translated into a set of paths in $\text{PTN}_{|\mathcal{F}|}$.

The intuition behind our approach is the following. For each candidate shuttle geoline we construct a graph consisting of the existing network, the line closure, and the candidate shuttle service at its maximum frequency. Note that we do *not* include the standard shuttle line in these graphs. Shortest paths for all OD-groups are calculated. The *reasonable* geopaths are then added to the candidate set, that is, all geopaths for which the estimated cost does not exceed the cost of the standard path by more than α units. Finally, the set of geopaths is translated to $\text{PTN}_{|\mathcal{F}|}$ to arrive at the full candidate set $\mathcal{P}_q$.

The concept of reasonable paths prevents passengers from being assigned to paths that make them considerably worse off than in the standard solution. However, it is not

guaranteed that passengers are assigned to their shortest path. For instance, when both the standard path and a shorter path exist in the final solution for an OD-group q, a passenger from q may be assigned to either path based on what results in the lowest global passenger inconvenience given certain budget constraints. Moreover, limiting the choice set to reasonable paths, may result in an overall more expensive solution. However, when choosing an appropriately small α, passengers can be considered to be indifferent between these paths, as they are all part of the reasonable path set, and their costs differ by at most α.

5.2 Path reduction

An OD-group consists of an OD-pair, a weight, and a path set. Given a set of OD-groups $\mathcal{Q}$ for a PTN$_{|\mathcal{F}|}$, the *path reduction* constructs a new set of OD-groups $\mathcal{Q}'$. The path reduction aims at reducing the number of paths and OD-groups contained in a new OD-group set, without changing the outcome of the route assignment. Computational results for the case study show that the path reduction reduces the set of OD-groups and paths by at least a factor of two, and decreases the computation time even more.

The *SPLC* model includes demand for *all* OD-groups, as any of these may be affected by the closure: some OD-groups have multiple paths to choose from to traverse the closure, other OD-groups do not traverse the closure but find a faster alternative in one of the candidate shuttles. Moreover, OD-groups can be affected by a change in demand resulting from the closure or a change in frequency on the existing line. Thus, in order to estimate passenger inconvenience and required capacity correctly, demand for all OD-groups needs to be included.

Each path of each OD-group introduces a new decision variable in the *SPLC* formulation. However, for the majority of OD-groups the path choice in terms of the geographical lines and stops is fixed, but still several paths are included for the different frequencies. For these OD-groups including demand per link leads to the same demand assignment as including a set of candidate paths the demand of the groups needs to be assigned to. The intuition behind the path reduction is to split passenger demand into demand that can be geolink-based since there is only one geopath, and demand for which there are multiple geopaths available, which requires a path based assignment. This is done *within* passenger groups. Although this does not necessarily reduce the number of paths and passenger groups, our case study results show that the practical benefits are significant.

The path reduction process is based on the geopath-translation, including entrance and transfer arcs, of the path set of OD-group q, which, to improve readability, we will still denote by $\mathcal{P}_q$ in this section. Note that any geopath can be translated into a new set of paths in PTN$_{|\mathcal{F}|}$, as discussed in Section 5.1.

5.2.1 Definition and properties

For an OD-group $q \in \mathcal{Q}$ we define s_q^* as the last common *node* among geopaths $\mathcal{P}_q$ before a change in the stop sequence, and t_q^* as the first common *node* after s_q^*, meaning that all geopaths in $\mathcal{P}_q$ contain the arcs from s to s_q^* and the arcs from t_q^* to t. Let $\mathcal{A}^*$ be the set of all arcs between s_q and s_q^* and all arcs between t_q^* and t_q. Furthermore, we define a new geopath set $\mathcal{P}^*$ by adding the remaining sub-path of each geopath $p \in \mathcal{P}_q$ after removing all arcs $a \in \mathcal{A}^*$ from this path. Note that by the construction of $\mathcal{A}^*$ and $\mathcal{P}^*$ all paths $p^* \in \mathcal{P}^*$ connect s_q^* to t_q^*. A new set of OD-groups $\mathcal{Q}'$ is constructed by defining:

- a new OD-group $q' \in \mathcal{Q}'$ as $s_{q'} = u, t_{q'} = v, w_{q'} = w_q, \mathcal{P}_{q'} = a$ for each arc $a = (u, v) \in \mathcal{A}^*$

- a new OD-group $q' \in \mathcal{Q}'$ as $s_{q'} = s_q^*, t_{q'} = t_q^*, w_{q'} = w_q, \mathcal{P}_{q'} = \mathcal{P}^*$

These new OD-groups q' are only defined for non-empty $\mathcal{A}^*$ and $\mathcal{P}^*$. $\mathcal{A}^*$ is empty when $\mathcal{P}_q$ contains multiple geopaths with fully disjoint arc sets, and therefore $\mathcal{P}_q = \mathcal{P}^*$. $\mathcal{P}^*$ is empty when all arcs in $\mathcal{P}_q$ are contained in all geopaths of $\mathcal{P}_q$, thus when $\mathcal{P}_q$ contains only one geopath. Paths are compared on an arc basis, thus including geoline-specific transfers, and new OD-groups are defined on a node basis. Thus transferring passengers, arriving passengers, and in-train passengers can be distinguished.

Observation I: the path set $\mathcal{P}_{q'}$ is uniquely defined by the origin and destination node of the new OD-groups $q' \in \mathcal{Q}'$. The cost of a path c_p is the sum of the cost of all arcs in the path. The cost of a single arc is independent of the cost of other arcs in the path, consisting of geoline specific entrance arcs, line arcs, transfer arcs and exit arcs. Thus, for any new $q', q'' \in \mathcal{Q}'$ that have the same s_q^*, t_q^*, the additional inconvenience of an assignment to any path p in the subset of paths $\mathcal{P}^*$ of either q', q'' is equal for both q', q'' independent of the OD-pairs of q', q''. Moreover any path $p^* \in \mathcal{P}^*$ will be part of both q', q'' as the concept of a *reasonable* path is defined as a fixed *incremental* cost on top of the standard path. Thus, any reasonable path of q' is a reasonable path of q'' and vice versa, and therefore the path set $\mathcal{P}_{q'}$ and $\mathcal{P}_{q''}$ are the same. It is straightforward to see that the same holds true for any OD-group defined by a single arc $a \in \mathcal{A}^*$.

Therefore, whenever there are two OD-groups q', q'' with $s_{q'} = s_{q''}, t_{q'} = t_{q''}$ we define a new combined OD-group $q''' := \{s_{q'''} = s_{q'}, t_{q'''} = t_{q'}, w_{q'''} = w_{q'} + w_{q''}, \mathcal{P}_{q'''} = \mathcal{P}_{q'}\}$ replacing q', q'' in $\mathcal{Q}'$.

Observation II: The path reduction process will not increase the number of paths by more than the number of arcs in $\text{PTN}_{|\mathcal{F}|}$. By construction, paths are added for arcs in $\mathcal{A}^*$ and paths in $\mathcal{P}^*$. The maximum number of paths resulting from $\mathcal{A}^*$ is smaller than, or equal to the number of arcs in $\text{PTN}_{|\mathcal{F}|}$ (because of the first observation). Paths resulting from $\mathcal{P}^*$ are subpaths of the original path set of OD-group q, and therefore this number is smaller than or equal to the number of paths in the original passenger group.

However, path reduction is likely to reduce the number of paths and the number of OD-groups because of the first observation. The number of decision variables is determined by $\sum_{q \in \mathcal{Q}} |P_q|$. Thus, both a reduction in the number of groups and the number of paths will lead to a reduction in the size of the mathematical programming problem defined in Section 4.4.

Observation III: a minimum inconvenience route assignment is the same for $\mathcal{Q}$ and $\mathcal{Q}'$ given a $\text{PTN}_{|\mathcal{F}|}$. By construction, paths are only split into multiple portions for those arcs that occur in all paths of the passenger group. Fixing this part of the assignment does not limit the path assignment model. Moreover, any path assignment of $\mathcal{Q}'$ can thus always be translated to a path in $\mathcal{Q}$, and vice versa. The cost of a path is defined as the sum of its arc costs, which contains transfer arcs and entrance arcs. The arc costs are independent of their position in the path. Therefore the cost of assigning a passenger to the full path in $\mathcal{Q}$ is equal to the cost of assigning a passenger to all disjunct subsets of the path included in $\mathcal{Q}'$. Therefore, the minimum inconvenience route assignments of $\mathcal{Q}$ and $\mathcal{Q}'$ are the same.

Remark: The path reduction process could be used independent of the concept of reasonable paths. However, in that case the path sets may not be uniquely defined by the origin and destination nodes s, t, possibly leading to a higher number of OD-groups.

5.2.2 Example

Consider the previously introduced public transport network and the associated graph given in Figure 2. For this network all reasonable geopaths for the OD-groups in the set $\mathcal{Q}$, where $\mathcal{Q}$ contains all passengers traveling to node V, are drawn in Figure 3a. This graph is again a schematic representation and arcs, nodes, and path segments representing entry and exit are omitted for reasons of clarity. Path reduction will compare the set of paths P_q for each OD-group $q \in \mathcal{Q}$ and then introduce the set of new OD-groups based on the comparison.

Take for example $(s_q, t_q) = (A, V)$. There are two reasonable paths: One passing through X and the other through T. Comparing these two paths, the last common node is $s_q^* = D$, while the first common node is $t_q^* = V$. Thus new OD-groups are introduced for (A, B), (B, C), (C, D) and (D, V) that collectively replace the original group A, V. Moreover an OD-group for the entry arc at station A is introduced (not shown in Figure 3a). A separate group for the exit arc at V will not be introduced as passengers may arrive from different directions at V. Distinguishing boarding and transfer arcs is essential to take into account the waiting time and transfer time of passengers.

For each of the groups in $\mathcal{Q}$ we follow this procedure and find a new OD-group set $\mathcal{Q}'$ for which the resulting paths are shown in Figure 3b. The number of paths is reduced by a factor of 5 in this example. Also in our case study we find that the number of paths is significantly reduced, and as a result the computational speed is increased.

(a) Geographical paths to source node V   (b) Geographical paths to source node V
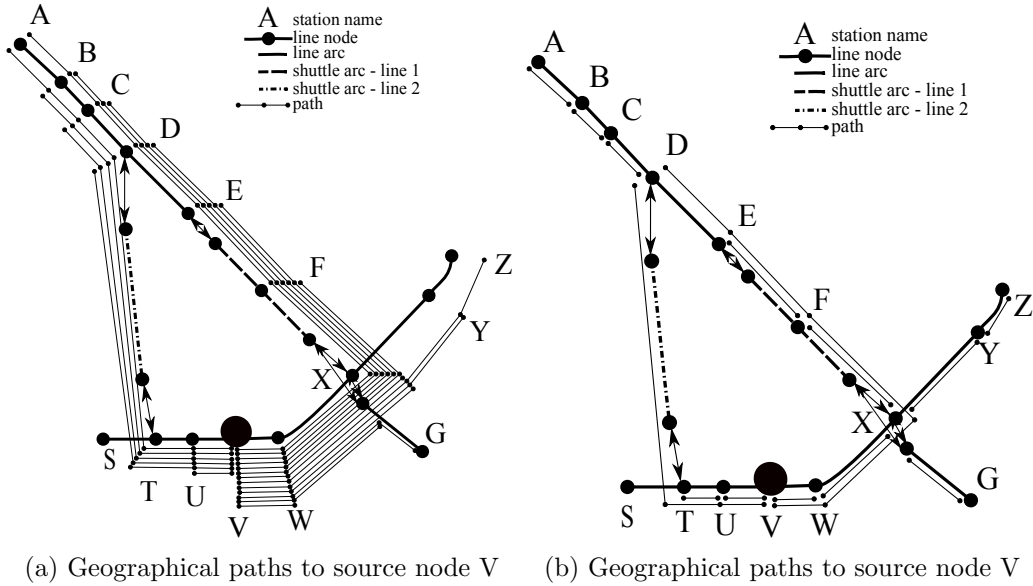
Figure 3: Example of path reduction

A quick observation of the network in Figure 3a may lead to the idea to simply cut off all branches, such as branch A-D. However, there are several reasons why this should *not* be done. First of all, it is important for the measurement of passenger inconvenience to distinguish passengers transferring at D from passengers who are in a vehicle at D and passengers who originate at D. Secondly, when one allows changes in the frequency on existing lines, this affects all passengers on the line, and thus requires considering passengers traveling outside the area of the disruption. Finally, an existing line could experience an increase in demand, which could lead to shortages of capacity for passengers

not affected by the disruption. Therefore it is important that any translation will lead to the same passenger inconvenience as the full OD matrix. The path reduction process meets this requirement.

6   Application

In this section we apply the *SPLC* model to a real life case study of a network closure in the urban rail network of the Massachusetts Bay Transportation Authority (MBTA) in Boston, Massachusetts, in the United States. We show that the *SPLC* model is able to find better solutions than the current standard practice at the MBTA which uses a single shuttle service to replace the track section being closed for repair. Moreover, through a sensitivity analysis we show that the proposed model's solutions are relatively robust against changes in demand and changes in the trade off of passenger inconvenience and operating cost. Finally we demonstrate that the model's solutions are of good quality under different assumptions about passenger behavior. Although the model formulation does not specifically optimize for robustness, the use of reasonable paths, as defined in Section 5, contributes to this property.

The case study is based on actual data which is described in Section 6.1. The experimental design is presented in Section 6.2. The model's solutions are discussed in Section 6.3, and finally Section 6.4 reports computational performance demonstrating the improvement in computation time of the model compared to prior formulations.

6.1   Data

The case study uses actual data from the MBTA urban rail network in Boston. Part of the 'Red Line' will be periodically suspended on weekends between 2013-2017 due to major maintenance work on the Longfellow Bridge. On such occasions, three stations central in the network will no longer be connected by the rail line. *One in eight* of the passengers traveling by metro in Boston on these days will be affected by this closure. The following discusses the input data consisting of the Origin-Destination (OD) matrix, public transport network, and candidate shuttle lines. The problem size is indicated in Table 2.

OD-matrix: The OD matrix is estimated based on fare payment data on five Saturdays between August 31, 2013 and September 28, 2013. Peak hour demand is estimated from the average peak demand between 2pm and 4pm on these days. The OD-demand matrix contains 1397 different OD-pairs.

Public Transport Network: The public transport network analyzed contains 113 stations, representing the full subway system of the MBTA with just two exceptions: some stations on the Green line were excluded due to lack of OD-information, together with some stations at the southern end of the Red Line, where very few passengers travel on the weekends. Stations are connected by 300 directed geographically distinct links under normal operations.

The core of the MBTA network is depicted in Figure 4. Due to the closure of the line between Kendall/MIT, Charles/MGH and Park Street, the Red Line is divided into two
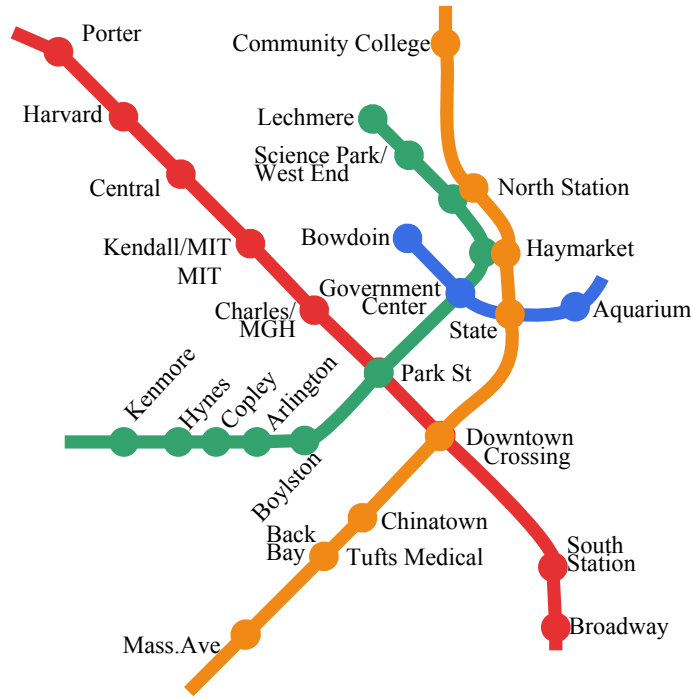
Figure 4: Core of MBTA network

parts: from Kendall north to Porter and Alewife, and from Park Street south to Broadway and Braintree. Thus the northern end of the Red Line will be disconnected from the rest of the network.

The travel times are based on the planned schedule. The translation of the network to $PTN_{|\mathcal{F}|}$ for *normal operations* results in a graph with 405 nodes and 1543 directed arcs. This graph includes nodes and arcs for transfers, entrances and exits for a single frequency for each line. The size of $PTN_{|\mathcal{F}|}$ including shuttles is larger and depends on the size of the candidate shuttle pool and the set of frequencies per line, as discussed in Section 4.1.

Candidate shuttle pool: The current strategy of the MBTA is to maintain connectivity along the original Red Line route by connecting the Kendall, Charles, and Park stations through a shuttle bus service running at 1 minute headways. We show that considering a broader set of candidate shuttle lines can significantly reduce passenger inconvenience with no increase in MBTA budget.

The candidate shuttle pool results from a demand analysis and an exploration of the network. In the demand analysis we identified the main attractor stations of demand for passengers traversing the link closures. The shuttle lines were constructed by connecting all attractor stations within a maximum distance of each other and requiring the shuttle lines to form one new direct connection between a station of the line affected by the closure and a station of another line, or the other part of the split line.

This procedure resulted in 14 different geolines, connecting 2 to 5 stations per line. Candidate shuttle lines are shown in Figure 5. Shorter versions of these lines were also included in the problem. The standard shuttle, that replaces the closed section of the line, is included in this set. The 14 candidate lines also include shuttles between stations beyond the closure. Such shuttle lines were selected in the solution of the *SPLC* model. Travel

Table 2: Problem size.

Network	113 stations, 300 trips
Shuttles	14 geolines
Passengers	1397 OD pairs

times for shuttle routes were estimated from vehicle location (AVL) data when available. The other connections were estimated based on Google maps driving times in dense traffic.

Finally, in addition to shuttle connections, we included the option of walking between selected stations that are close together, that is, less than 10 minutes walking time. Allowing passengers to walk between these stations acknowledges that the OD matrix is just an estimate of the passenger's true origin and destination, and prevents the assignment of passengers to public transport routes for parts of their journey that are easily walkable.
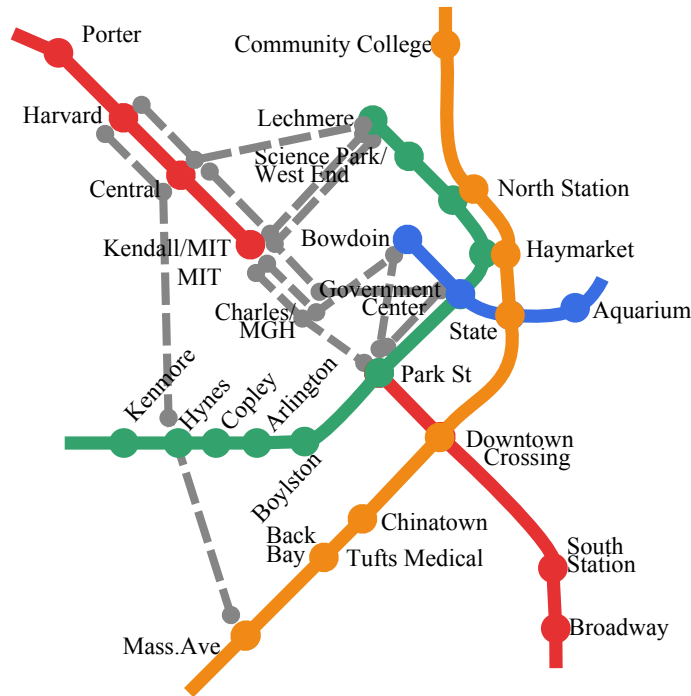

Figure 5: Core of MBTA network with candidate shuttle lines.

6.2 Experimental Design

All experiments are based on two cases: one scheduling shuttles, the second scheduling both shuttles *and* allowing a redistribution of vehicles on the Red Line, as well as changing the frequencies on both parts. We evaluate different objective functions, different demand levels, and the robustness of the solution under different passenger behavioral assumptions. The quality of solutions is measured by the passenger inconvenience (PI), compared to PI of the standard solution (Sf_{st}) of a single shuttle replacing the closed links. OC is presented as the required number of shuttles for the solution, which defines OC as all shuttles in our case study have equal cost. An overview of the complete experimental design is provided in Table 3.

Shuttle Scheduling and Network Scheduling: We study two cases for the Longfellow Bridge closure. The first, *Shuttle Scheduling* (C_{shut}), considers only the introduction of shuttle lines in the current network. The second, *Network Scheduling* (C_{netw}), considers both the introduction of new shuttle lines in the current network, *and* the redistribution of vehicles and the frequency setting on both sections of the Red Line. The number of vehicles assigned to the two disconnected sections of the Red Line is limited to the current number of available vehicles for the line. Moreover, a cost is included for the change in the number of drivers a new frequency requires in comparison to the current 8 minute headway schedule. A decrease in frequency of one section of the line can compensate the cost for an increase in frequency of the other section of the line. Finally, a budget constraint is added so that the total operating cost do not exceed the budget available for C_{shut}. The frequency candidate set contains headways between 4 and 10 minutes in 1 minute increments.

For each shuttle line we introduce a frequency such that the headway varies between 1 and 10 minutes in 1 minute increments. The operating cost increases nonlinearly with the decrease in headways. The *SPLC* problem formulation allows assigning more vehicles to a line, however, *PI* depends on the selected frequencies. Therefore it is more important to gradually change *PI* than *OC* for the included set of frequencies. Although a continuous set of frequencies are feasible per line, in Section 6.4 we establish that indeed the inclusion of 10 frequency options is sufficient, that is, the optimal solution does not change when the set of frequencies is defined at to a finer level.

Objective function: The objective is to minimize *PI* due to the closure within a budget for *OC*. However, possibly there are good solutions using *less* vehicles, or the maximum fleet size has yet not been determined. Therefore three different objectives are compared:

- Ob_{PI}^{L}: minimizing *PI* with a *limited* shuttle pool

- $Ob_{PI,OC}^{L}$: minimizing *PI* and *OC* with a *limited* shuttle pool

- $Ob_{PI,OC}^{N}$: minimizing *PI* and *OC* with an *unlimited* shuttle pool.

Both *PI* and *OC* are expressed in monetary terms. MBTA bus operating costs are approximately \$140 per hour. Passenger inconvenience is expressed in terms of *value of time*, defined as the weighted total travel time, including waiting time, transfers, and in-vehicle time, multiplied by an estimate of the wage per hour. We conduct a sensitivity analysis for different assumptions about the value of time and different levels of demand. This includes three levels for the wage per hour $\omega := \{\$8, \$16, \$32\}$, ranging between the minimum wage and average wage in Massachusetts, and four demand levels to scale the OD matrix by $\sigma := \{0.6, 0.8, 1.0, 1.2\}$, with $\sigma = 1$ equal to the current average (peak) demand. These options reflect both a reduction in demand as the expected inconvenience may result in some passengers postponing their trips or changing mode, and an increase in demand that may be caused by major sporting events, such as a Boston Red Sox game. Finally, the shuttle pool is limited to the number of shuttles required for the standard solution Sf_{st} under any demand scenario.

The *PI* is defined as the weighted sum of the costs of the arcs in the paths assigned to passengers. The waiting time to board a vehicle is defined as the expected waiting time for the first arrival based on random passenger arrivals and the scheduled frequency.

For transfers between any metro or shuttle line, the waiting time is defined as the waiting time for a random arrival at the transfer-to line plus a fixed transfer-penalty, which reflects average transfer inconvenience in the MBTA network where arrivals of lines are not synchronized to enable easy transfers and penalizes additional transfers (and waiting time) due to he closure. We define the *costs* of waiting time to be three times the cost of in-vehicle time, as first proposed in Quarmby (1967). Wardman (2004) lists more studies that find broadly similar results, although his results suggest that the actual costs may depend on other variables such as the trip purpose.

The *SPLC* model minimizes overall *PI*, and consequently not necessarily all passengers are assigned to their preferred path. The true preferred path of a passenger is unknown. Common approaches to modeling path choice are that each passenger follows the shortest path or that passengers choose one of a set of reasonable alternatives, based on some probabilistic rule. To address this we evaluate the robustness of the obtained solution under two different passenger behavior models. The first model, RA_{sp}, assigns passengers to their shortest path in the proposed network. The second model, RA_{prob}, assigns passengers probabilistically, based on a logit model to one of two candidate routes in the proposed network: (1) the shortest path and (2) the geographical route that follows most closely the standard route. The probability of choosing path p_i, $i \in \{1,2\}$, is defined as:

$$P(p_i) = \frac{e^{\theta c_{p_i}}}{e^{\theta c_{p_1}} + e^{\theta c_{p_2}}} \qquad\qquad \forall i \in \{1,2\}$$

Where c_{p_i} is the weighted travel time of the path, and $\theta = -0.2$ such that even if the difference between the p_i's is 15 minutes, 5% of the passengers still travel along the (longer) standard path (for example, reflecting inertia).

Both assignments, unlike the route assignment in the *SPLC* model, do not consider capacity constraints, but depend only on the 'optimal' set of geolines and their frequencies. In order to make a fair comparison between the quality of the solutions, the *OC* is recalculated such that sufficient capacity is provided for all passengers to follow their preferred route, and that all lines are operated at a frequency no less than specified in the solution of the *SPLC* model. However, the *PI* is not adjusted for the resulting increases in frequency. Thus, conservative estimates are presented of both *OC* and *PI*,

Table 3: Notation for model applications.

Category	Notation	Description
Solutions	Sf_{mod}	PTN_f solution from *SPLC* model
	Sf_{st}	PTN_f standard solution with single shuttle bridging closure
Cases	C_{shut}	Optimizing shuttle lines operated
	C_{netw}	Optimizing shuttle lines operated and reassignment of Red Line capacity
Passenger	RA_{mod}	Model solution, model route assignment
Behavior	RA_{sp}	Model solution, shortest path assignment
Models	RA_{prob}	Model solution, probabilistic path assignment
Objectives	$Ob_{PI,OC}^{L}$	*PI* and limited shuttle pool
	Ob_{PI}^{L}	*PI*, *OC*, and limited shuttle pool
	$Ob_{PI,OC}^{N}$	*PI*, *OC*, and unlimited shuttle pool
parameters	ω	Value of time
	σ	Demand factor (increase or decrease)

6.3 Results

This section discusses the results of the case study according to the experimental design defined in Section 6.2. First the model solution is discussed in Section 6.3.1, then Section 6.3.2 reports on the sensitivity analysis.

6.3.1 Solution

A solution for the *SPLC* problem defines the set of lines to be operated and their frequencies given $PTN_{|\mathcal{F}|}$ defined by the geolines, the set of closed links, and the set of possible frequencies per line. Given the demand, a set of reasonable paths, and the number of available vehicles, the model identifies the set of frequencies that minimize the objective function: the weighted sum of passenger inconvenience and operating cost. A solution's quality is compared against current practice, and measures *PI* relative to normal operations, so that it is possible to conclude that a solution is not just better than the standard solution for line closures, but also a good overall solution in the event of a line closure. This section presents results for the cases C_{shut} and C_{netw} with objective $Ob_{PI,OC}^{L}$, wage $\omega =$ \$16, and demand scale $\sigma = 0.8$, which we define as the *base case*[1].
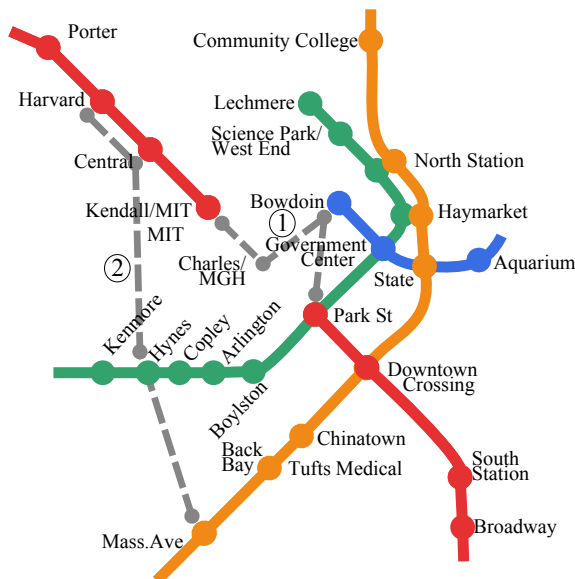

Figure 6: MBTA core network with shuttle lines operated in base case C_{shut} solution.

Two shuttle lines are operated in the C_{shut} solution. The resulting network is shown in Figure 6. The solution includes shuttle (1) operated at a headway of slightly less than 3 minutes, and shuttle (2) at a headway of 2 minutes. Shuttle (2) provides a direct high frequency connection between Harvard, Central, and Hynes: three of the major trip attractors in the network, thereby eliminating two transfers for these passengers and reducing waiting time (headways on the metro lines are 8 minutes, the shuttle runs at 2 minutes). This illustrates how shuttles beyond the closed area can significantly reduce delay, without increasing *OC*.

The solution for C_{netw} operates the Northern branch of the Red Line at 4 minute headways instead of 8 minutes. The 4 minute headway comes at the cost of a lower frequency of

---

[1]Note: the MBTA expects a 20% decrease in demand due to the closure. So $\sigma = 0.8$ is selected over $\sigma = 1$ as the base case.

10 minute headways on the southern branch of the Red Line and the additional need for about 4 drivers. The total number of vehicles on the Red Line stays the same. The solution operates shuttle (1) in Figure 6 at a 2 minute headway, and a modified version of shuttle (2) starting at Central (i.e. omitting the Harvard section) at a 3 minute headway. Although this introduces an additional transfer for all Harvard passengers, the higher frequency on this part of the Red Line compensates for this increase in inconvenience.

Table 4: PI percentage increase over normal operations, OC fleet size (shut= 64), $\text{Ob}_{\text{PI,OC}}^{\text{L}}$, ω= 16, σ=0.8

	C_{shut}		C_{netw}	
	PI	OC	PI	OC
Sf_{mod}	15.7	61	11.1	55
Sf_{st}	25.7	64	25.7	64

Table 4 shows PI and OC for the Sf_{mod} solution and the standard Sf_{st} solution for the C_{shut} and C_{netw} base cases. PI is expressed as the percentage increase in passenger inconvenience over standard operations. The OC is presented as the number of shuttle buses required, given a fixed cost of \$140 per bus the OC follows directly from this. The OC of C_{netw} represents the total cost for shuttles and additional drivers on the Red Line in terms of shuttles. Lower values for PI and OC represent better solutions.

The model's solution for C_{shut}, allowing bus deployment to shuttle services only, has a lower OC than Sf_{st} (61 instead of 64 buses) and a PI of 15.7%, reducing PI by 40% compared to the standard solution. For example, for a passenger who makes a 40 minutes trip under normal conditions and travels over the links which will be closed, in the standard solution the journey would be 10 minutes longer (on average), and 6 minutes (on average) longer in the model solution.

In the C_{netw} solution, which allows redistribution of vehicles and selecting a new frequency within a feasible range on the two Red Line sections, the passenger inconvenience is reduced to 11.1% - a reduction of 57% compared to the standard solution. This results in a delay of 4 minutes instead of 10 minutes in the previous example. The solution has a lower operating cost than both the standard solution and C_{shut} (55 shuttle equivalent instead of 64 or 61, respectively), while using the same number of cars in the Red Line.

The distribution of PI over passengers for the standard C_{shut} and the C_{netw} solutions respectively are given in Figure 7 as the weighted additional travel time per passenger in comparison to normal operations. Passengers with no change in inconvenience in comparison to normal operations are excluded from the graph. A unique translation of the path reduction passenger assignment to an assignment per original OD-group does not exist whenever a passenger group is assigned to multiple paths. Therefore we base the analysis of per passenger inconvenience on a shortest path assignment in the network, which has a similar PI as discussed in Section 6.3.2.

Both solutions for C_{shut} and C_{netw} reduce worst case delays, and distribute delay more evenly over passengers than the standard approach Sf_{st}. In the standard solution most passengers have a weighted delay of 25-30 minutes. The C_{shut} and C_{netw} greatly reduce the number of passengers experiencing a delay of 25 minutes or more, and reduce delay to under 15 minutes for most passengers. The C_{netw} solution reduces worst case delays for more passengers than the C_{shut} at a lower OC, although it also has a higher number of passengers with small delays (between 0 and 5 minutes) than C_{shut} solution.

Moreover, the C_{netw} case benefits non-affected passengers by operating a higher frequency on the Northern Branch of the Red Line, explaining the -5 minutes delay bar in the graph. The percentage of passengers experiencing inconvenience from the closure is 28% for Sf_{st}, 23% for C_{shut} and 22% for C_{netw}. Thus the C_{netw} reduces PI not only by providing better connections for non-affected passengers, but also by reducing inconvenience for affected passengers
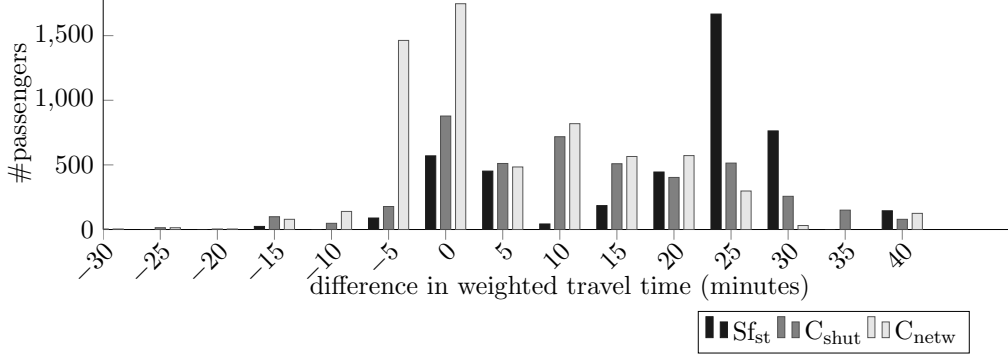

Figure 7: Distribution of per passenger PI increase for Sf_{st}, C_{shut} and C_{netw}.

6.3.2 Sensitivity analysis

This section discusses the quality of the solution for different objective functions, demand levels, and its robustness under different passenger assignment models according to the experimental design presented in Section 6.2.

Objectives: The $SPLC$ model has the objective of minimizing passenger inconvenience at a constrained operating cost. We compared three different formulations of the objective function: $Ob_{PI,OC}^{L}$, Ob_{PI}^{L} and $Ob_{PI,OC}^{N}$. In $Ob_{PI,OC}^{L}$ and in $Ob_{PI,OC}^{N}$ PI is weighted by the value of time, assuming three different values: $\omega := \{8, 16, 32\}$ \$/hr. For the different objectives Table 5 presents the range of PI and OC over these three different values of time for both the C_{shut} and C_{netw} cases. PI is expressed as the percentage increase over normal operations (with the same value of time), and OC as the number of shuttle buses. The Ob_{PI}^{L} solutions do not depend on ω.

The $Ob_{PI,OC}^{N}$ uses a significantly higher number of shuttles for a relatively small increase in PI: 69 busses for a reduction in PI of 0.8 percent point (Ob_{PI}^{L}) or 0.9 percent point ($Ob_{PI,OC}^{L}$), at a cost of respectively 5 and 7 shuttles for C_{shut}. The C_{netw} has a higher difference in OC for a smaller difference in PI. An unconstrained fleet size thus can lead to a large variation in the number of shuttles in a solution given the relative weights assigned to PI and OC in the objective. The advantage of $Ob_{PI,OC}^{L}$ is that, because the fleet size is constrained, different weights lead to smaller differences in solutions. Not assigning any weight at all to the OC, as in Ob_{PI}^{L}, solutions may contain unnecessarily many shuttles: 2 (C_{shut}) to 6 (C_{netw}) shuttles could be saved by increasing PI with 0.1% point, an OC reduction that is 5 to 9 times as high as the resulting increase in PI. Therefore, $Ob_{PI,OC}^{L}$ can assist operators best in making the right trade-off between the PI and OC without a high sensitivity to chosen weights, and the results for the case study clearly illustrates that redesigning the network can reduce both PI and OC.

Table 5: Different objective functions for C_{shut} and C_{netw} cases, $\omega := \{8, 16, 32\}$ \$/hr, Ob_{PI}^L and $Ob_{PI,OC}^L$, shut= 64.

	C_{shut}		C_{netw}	
	PI	*OC*	*PI*	*OC*
Ob_{PI}^L	15.5	64	10.4	64
$Ob_{PI,OC}^L$	15.6 - 17.7	50 - 62	10.5 - 12.0	54 - 58
$Ob_{PI,OC}^N$	14.7 - 17.7	50 - 69	10.1 - 11.7	54 - 81

Demand: The base case assumes a demand loss of 20% due to the closure. Although this is MBTA's best estimate, demand could be lower when passengers can easily postpone their trips, or have access to other modes. However, when the added passenger inconvenience is small the resulting decrease in demand may also be small. Major sports events could even increase demand. Because the demand is uncertain, we evaluate the sensitivity of the solutions for demand scaled by $\sigma = \{0.6, 0.8, 1.0, 1.2\}$.

The *PI* and *OC* for all scenarios are given in Table 6 for the mixed objective with limited fleet size, $Ob_{PI,OC}^L$. Both an increase and a reduction of demand (slightly) increases *PI* in comparison to the base scenario (0.8) for both C_{shut} and C_{netw}. The increase in *PI* at higher demand levels is caused by the limited fleet size of 64 shuttles, which is not enough to provide a similar service, especially at a 20% demand increase. Increasing the fleet size reduces *PI*, but requires a significant extension of the fleet (77 and 84 busses for $\sigma = \{1.0, 1.2\}$ for C_{shut}). *PI* also increases when demand decreases, as a decrease in demand reduces the relative benefit of operating lines at a higher frequency, and less vehicles are therefore employed than at higher demand levels, resulting in lower *OC* and also a slightly lower *PI* for $Ob_{PI,OC}^L$ and $Ob_{PI,OC}^N$.

In terms of shuttle services for C_{shut}, at $\sigma = 0.6$ the same lines are proposed as in the base case ($\sigma = 0.8$), but operating at lower frequencies. The resulting public transport network is shown in Figure 6. For $\sigma = \{1.0, 1.2\}$ a third shuttle connecting the Red Line to Lechmere station is proposed in addition to shuttle (1), and a shorter version of shuttle (2) operating from Central Square to Hynes/Massachusetts Avenue is included. The short lines (2) and (3) help cope with the increased demand by using the same number of shuttle buses. The *PI* and *OC* of the solutions with original shuttles (1) and (2), and the solution with three shuttle lines, are fairly similar. Thus the base solution is satisfactory even with higher demand, as long as the frequencies are adjusted accordingly. Results for C_{netw} are similar, with higher frequencies of shorter lines, and a connection to Lechmere at high demand.

Table 6: Different demand levels for C_{shut} and C_{netw} cases, shut=64.

	C_{shut}		C_{netw}	
OD scale	*PI*	*OC*	*PI*	*OC*
0.6	16.7	47	12.6	42
0.8	15.7	61	11.1	55
1.0	16.9	64	13	62
1.2	19	64	15.2	64

Route Assignment: The model optimizes over all *PI*. The path assignment is restricted to *reasonable* paths, between which, are could argue, passengers are indifferent - but still
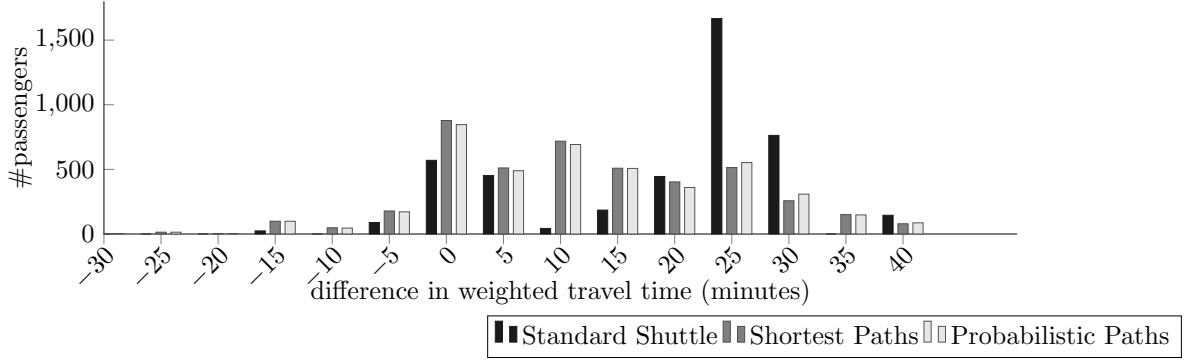
Figure 8: Distribution of per passenger *PI* increase for assignment in the Standard Shuttle solution, and Shortest Path and Probabilistic Path assignment for the C_{shut} (shut= 64, $Ob_{PI,OC}^{L}$, $\omega= 16$, $\sigma=0.8$) solution.

does not fully reflect free route choice. This section therefore investigates the robustness of the solution under different assumptions on passenger route choice behavior.

The *PI* and *OC* for the model's route assignment RA_{mod}, the shortest path route assignment RA_{sp}, and the probabilistic route assignment RA_{prob} in the network of the model's solution are shown in Table 7 for the base case C_{shut}. *PI* is not very sensitive to the route assignment model. Actually *PI* is within 0.4% points from the model's estimated *PI*. The *OC* for the two different route assignment models increases. However, the *OC* is a conservative estimate (upper bound) of costs, as it estimates costs of lines based on the maximum vehicles per line needed for either the *SPLC* solution or the new passenger assignment solution. If the frequency of lines is reduced to mirror the decrease in demand due to the changed route assignment models, the increase in *OC* would be lower.

The delay distributions of RA_{sp} and RA_{prob}, shown in Figure 8, are similar. In comparison to the standard solution, both route assignment models reduce passenger inconvenience by reducing the worst case delays. Results are similar for both the C_{shut} and C_{netw} cases at different levels of ω and σ, and for other objective functions.

Table 7: Route Assignment C_{shut}

| | C_{shut} | |
	PI	*OC*
RA_{mod}	15.7	61
RA_{sp}	15.3	63
RA_{prob}	15.8	65

6.4 Computational considerations

The computational benefits of the new formulation using *vehicle variables* and the *path reduction* preprocessing step are evaluated in this section. Two formulations for the *SPLC* problem are compared: 1) *SPLC*(y, v), containing both vehicle variables v and line variables y, and 2) *SPLC*(y), containing line variables y only, which is a more straightforward adaption of Claessens et al. (1998) to the *SPLC* problem. All computational experiments were run using CPLEX version 12.6 on an Intel I7 3.07GHz processor using the default settings assigning 0.5GB of memory to the program.

Path reduction: The path reduction process reduces the number of OD groups by 75 % (from 1397 to 320). The number of different paths is reduced by more than 50 %. For example, with 10 frequency options per shuttle line the number of paths is reduced from 59,237 to 25,179 (reduction of 57%), while with 40 frequency options per shuttle line the number of paths is reduced from 224,666 to 99,726 paths (reduction of 56%). The number of x_{pq} decision variables is thus greatly reduced, thereby almost equally reducing problem size as the great majority of decision variables result from x_{pq}. Moreover, many of the new paths are 'simple' paths that traverse a single geo-arc. This might further contribute to decreasing the computation time.

The CPLEX optimization running time (T_{opt}) and the total running time including pre-processing (T_{tot}) for solving the problem with and without path reduction are given in Table 8. The candidate frequency set for both the $SPLC(\text{y})$ and $SPLC(\text{y}, \text{v})$ formulations are chosen such that the resulting solutions are comparable. For both formulations, the path reduction process reduces computational time significantly: it is 58 times faster for the $SPLC(\text{y})$ case and 10 times faster for the $SPLC(\text{y}, \text{v})$ case, without changing the problem definition or optimal solution value. Therefore, in the discussion that follows, it will be used in all reported computation times.

Table 8: Computation times (seconds) (C_{shut} , $\text{Ob}_{\text{PI}}^{\text{L}}$, ω=16, σ=0.8, shut = 64)

| Model | Number of | Path reduction: | | | |
| | | No | | Yes | |
	Frequencies	T_{opt}	T_{tot}	T_{opt}	T_{tot}
$SPLC(\text{y})$	40	3297	3517	56.6	82
$SPLC(\text{y}, \text{v})$	10	96.4	115	9.53	12

Vehicle variables: We search for the number of frequency settings per model to arrive at comparable solutions. We find that at 40 options the $SPLC(\text{y})$ and $SPLC(\text{y}, \text{v})$ models propose the same shuttle lines, with a difference in objective value of only 0.1%. The computation times for optimizing through CPLEX (T_{opt}) and the total run time including preprocessing (T_{tot}) shown in Table 9 are for different frequency options of the two formulations $SPLC(\text{y})$ and $SPLC(\text{y}, \text{v})$. The $SPLC(\text{y}, \text{v})$ solves the model 5 times faster, in 9.53 seconds instead of 56.6 seconds respectively, as it requires only 10 frequencies to arrive at the same solution as the $SPLC(\text{y})$ formulation with 40 frequencies. The addition of vehicle variables itself does not speed up the computation time, solving $SPLC(\text{y}, \text{v})$ with 40 frequencies in 871 seconds is slower than $SPLC(\text{y})$. But the $SPLC(\text{y}, \text{v})$ formulation with 10 frequencies finds the same solution as the $SPLC(\text{y}, \text{v})$ with 40 frequencies. Thus, because the $SPLC(\text{y}, \text{v})$ requires less frequencies it is significantly faster than the $SPLC(\text{y})$ formulation. The $SPLC(\text{y}, \text{v})$ formulation together with the path reduction are able to solve the model 345 times faster, from 3297 seconds to 9.53 seconds as shown in Table 8.

Table 9: Computation times (seconds) (C_{shut} , $\text{Ob}_{\text{PI}}^{\text{L}}$, ω=16, σ=0.8)

Model	Frequencies	T_{opt}	T_{tot}
$SPLC(\text{y})$	40	56.6	82
$SPLC(\text{y}, \text{v})$	40	871	894
$SPLC(\text{y}, \text{v})$	10	9.53	12

Cases and Objectives: Computation times for the cases C_{shut} and C_{netw} are presented in Table 10. For all objective functions, the $SPLC(y, v)$ formulation is faster than the $SPLC(y)$ formulation. It computes C_{shut} in a minimum of 9.53 seconds, versus 39.48 for the $SPLC(y)$ formulation, thus 4 times faster. For the more complex case C_{netw} computing time is 28.2 seconds and 1192 seconds respectively, thus $SPLC(y, v)$ computation time is around 40 times faster than the $SPLC(y)$ formulation, 10 times more than the speed increase for the C_{shut} case. These results suggest that the benefit of the $SPLC(y, v)$ formulation increases as the problem complexity increases.

There are significant differences in computation times for different objective functions as shown in Table 10. Ob_{PI}^{L} is fastest, $Ob_{PI,OC}^{L}$ is 1.5 times as long and the $Ob_{PI,OC}^{N}$ is 3 times longer than the Ob_{PI}^{L}. We were not able to solve the $SPLC(y)$ formulation with objective $Ob_{PI,OC}^{N}$ due to insufficient memory. By estimating an appropriate upper bound on the number of vehicles, the $Ob_{PI,OC}^{N}$ could be solved using the $Ob_{PI,OC}^{L}$ for different fleet sizes, thus increasing the computation speed.

Table 10: Computation times (seconds)(Sf_{mod}, $\sigma = 0.8$, $\omega = 16$). *-instances were not solved due to insufficient memory.

Settings		$SPLC(y, v)$		$SPLC(y)$	
Instance	Cost Shuttle	T_{opt}	T_{tot}	T_{opt}	T_{tot}
C_{shut}	Ob_{PI}^{L}	9.53	12.27	39.48	57.46
C_{shut}	$Ob_{PI,OC}^{L}$	15.4	18.08	74.68	92.74
C_{shut}	$Ob_{PI,OC}^{N}$	34.2	37.03	1436	1461
C_{netw}	Ob_{PI}^{L}	28.2	52.00	1192	2588
C_{netw}	$Ob_{PI,OC}^{L}$	36.2	60.33	1421	2790
C_{netw}	$Ob_{PI,OC}^{N}$	65.6	91.26	*	*

7 Conclusion and Discussion

This paper presents a new analytic model for the design of shuttle services for planned closures in high frequency urban transportation networks. A new formulation is proposed that allows the selection of the frequency for new and existing lines, while minimizing passenger inconvenience, including frequency dependent transfer and waiting costs, within a budget for operating cost. A sensitivity analysis including different demand levels, different objective functions, and different assumptions about passenger route choice behavior is included. It is shown that the proposed formulation is able to find good solutions quickly for real life cases of planned closures, making this model a candidate for real time applications. The solutions are better both in terms of passenger inconvenience and operating cost in comparison to the standard practice.

The proposed model formulation combines the ideas of Borndörfer et al. (2007) of a path formulation with the idea of Schöbel and Scholl (2006) to include transfers, and the idea of Claessens et al. (1998) to combine the capacity assignment and the opening of lines into one decision variable. We extend these ideas with the introduction of additional vehicle decision variables that relax the assignment of capacity to a line, while still including integer decision variables of which lines to operate. This allows the specification of a minimum frequency for a line that is operated, as well as frequency dependent costs of path assignments.

The model is somewhat limited by the assumption that demand and frequencies are not time dependent. Although this assumption is fair and common in high frequency networks with cyclical timetables, it might not be valid for all public transportation networks. Future research could focus on extending the model formulation to time dependent problems. For more complex problems, the proposed model could be combined with a dynamic column generation approach. The combination of column generation with the proposed path reduction process could be an interesting area of research for any multi-commodity flow model. The proposed path reduction process assumes there is sufficient capacity available to transport all demand. In real time disruptions capacity shortages may arise, leading to competition amongst passengers for seats. How to include this competition in the model and path reduction, specifically preventing unrealistical altruistic behavior of passengers to reach a social optimum, is identified for future research.

Finally, the introduction of the concept of *reasonable* paths is a first step towards including realistic passenger behavior in systems with free route choice. Future research could extend on this, for instance extending the ideas proposed in Schmidt (2012) for an arc-formulation to the path formulation used in this paper.

ACKNOWLEDGEMENT

References

R. Borndörfer, M. Grötschel, and M. E. Pfetsch. A Column-Generation Approach to Line Planning in Public Transport. *Transportation Science*, 41(1):123–132, February 2007.

V. Cacchiani, A. Caprara, L. Galli, L. Kroon, G. Maroti, and P. Toth. Railway Rolling Stock Planning: Robustness Against Large Disruptions. *Transportation Science*, 46(2): 217–232, December 2012.

V. Cacchiani, D. Huisman, M. Kidd, L. Kroon, P. Toth, L. Veelenturf, and J. Wagenaar. An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B: Methodological*, 63:15 – 37, 2014.

L. Cadarso, A. Marín, and G. Maróti. Recovery of disruptions in rapid transit networks. *Transportation Research Part E: Logistics and Transportation Review*, 53:15–33, July 2013.

A. Ceder and N. H. M. Wilson. Bus network design. *Transportation Research Part B*, 20 (4):331–344, 1986.

S. Cicerone, G. D'Angelo, G. Stefano, D. Frigioni, A. Navarra, M. Schachtebeck, and A. Schöbel. Recoverable robustness in shunting and timetabling. In R. K. Ahuja, R. H. Möhring, and C. D. Zaroliagis, editors, *Robust and Online Large-Scale Optimization*, volume 5868 of *Lecture Notes in Computer Science*, pages 28–60. Springer Berlin Heidelberg, 2009.

M. T. Claessens, N. M. van Dijk, and P. J. Zwaneveld. Cost optimal allocation of rail passenger lines. *European Journal of Operational Research*, 110:474–489, 1998.

M. Fischetti and M. Monaci. Light robustness. In R.K. Ahuja, R.H. Möhring, and C.D. Zaroliagis, editors, *Robust and Online Large-Scale Optimization*, volume 5868 of *Lecture Notes in Computer Science*, pages 61–84. Springer Berlin Heidelberg, 2009.

M. Frumin and J. Zhao. Analyzing passenger incidence behavior in heterogeneous transit services using smartcard data and schedule-based assignment. *Transportation Research Record: Journal of the Transportation Research Board*, 2274:52–60, 2012.

J. Jespersen-Groth, D. Potthoff, J. Clausen, D. Huisman, L. Kroon, G. Maróti, and M. Nyhave Nielsen. Disruption management in passenger railway transportation. *Robust and Online Large-Scale Optimization, Lecture Notes in Computer Science*, 5868:399–421, 2009.

J. G. Jin, K. M. Teo, and L. Sun. Disruption response planning for an urban mass rapid transit network. In *Proceedings of the Transportation Research Board 92nd Annual Meeting - Washington DC 2013*, 2013.

J. G. Jin, K. M. Teo, and A. R. Odoni. Optimizing bus bridging services in response to disruptions of urban transit rail networks. *Working Paper, Massachusetts Institute of Technology (under review for journal publication)*, 2014.

M. Kaspi and T. Raviv. Service-Oriented Line Planning and Timetabling for Passenger Trains. *Transportation Science*, 47(3):295–311, June 2012.

K. Kepaptsoglou and M. G. Karlaftis. The bus bridging problem in metro operations: conceptual framework, models and algorithms. *Public Transport*, 1(4):275–297, 2009.

L. G. Kroon, G. Maróti, and L.K. Nielsen. Rescheduling of railway rolling stock with dynamic passenger flows. *Transportation Science*, 2014. To aprear.

S. Lan, J. Clarke, and C. Barnhart. Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions. *Transportation Science*, 40(1):15–28, 2006.

C. Liebchen, M. Schachtebeck, A. Schöbel, S. Stiller, and A. Prigge. Computing delay resistant railway timetables. *Computers & Operations Research*, 37(5):857–868, May 2010.

L. K. Nielsen, L. Kroon, and G. Maróti. A rolling horizon approach for disruption management of railway rolling stock. *European Journal of Operational Research*, 220(2): 496–509, July 2012.

B. Pender, G. Currie, A. Delbosc, and Y. Wang. Proactive recovery from rail disruptions through provision of track crossovers and bus bridging. *Transportation Research Record: Journal of the Transportation Research Board*, 2275-1:68–76, 2009.

B. Pender, G. Currie, A. Delbosc, and N. Shiwakoti. Disruption recovery in passenger railways: International survey. *Transportation Research Record: Journal of the Transportation Research Board*, 2353:22–32, 2013.

D. Potthoff, D. Huisman, and G. Deaulniers. Column generation with dynamic duty selection for railway crew rescheduling. *Transportation Science*, 44:493–505, 2010.

D.A. Quarmby. Choice of travel mode for the journey to work. *Journal of Transport Economics and Policy*, (1):273–314, 1967.

M. Schmidt. Line planning with equilibrium routing. Technical report, Institut für Numerische und Angewandte Mathematik, Georg-August Universität Göttingen, 2012. `http://num.math.uni-goettingen.de/preprints/files/2012-09.pdf`.

A. Schöbel. Line planning in public transportation: models and methods. *OR Spectrum*, 34(3):491–510, May 2011.

A. Schöbel and S. Scholl. Line Planning with Minimal Traveling Time. *ATMOS Proceedings*, 2006.

L.P. Veelenturf, L. K. Nielsen, G Maróti, and L.G. Kroon. Passenger oriented disruption management by adapting stopping patterns and rolling stock schedules. In *Proceedings of CASPT - Santiago, Chile*, 2013.

L.P. Veelenturf, D. Potthoff, D. Huisman, L.G. Kroon, G. Maróti, and A.P.M. Wagelmans. A quasi-robust optimization approach for resource rescheduling. *Transportation Science, Accepted*, 2014.

M. Wardman. Public transport values of time. *Transport Policy*, 11(4):363–377, 2004.