Getting Charlie off the MTA: A multiobjective optimization method to account for cost constraints in public transit accessibility metrics

Matthew Wigginton Conway[*] and Anson F. Stewart[†]

April 30, 2019

The Version of Record of this manuscript has been published and is available in *International Journal of Geographical Information Science* 2019 https://www.tandfonline.com/doi/full/10.1080/13658816.2019.1605075

Abstract

Most analyses of accessibility by public transit have focused on travel time and not considered the cost of transit fares. It is difficult to include fares in shortest-path algorithms because fares are often path-dependent. When fare policies allow discounted transfers, for example, the fare for a given journey segment depends on characteristics of previous journey segments. Existing methods to characterize tradeoffs between travel time and monetary cost objectives do not scale well to complex networks, or they rely on approximations. Additionally, they often require assumed values of time, which may be problematic for evaluating the equity of service provision. We propose a new method that allows us to find Pareto sets of paths, jointly minimizing fare and travel time. Using a case study in greater Boston, Massachusetts, USA, we test the algorithm's performance as part of an interactive web application for computing accessibility metrics. Potential extensions for journey planning and route choice models are also discussed.

Keywords: **Pareto-optimal solution; multiobjective optimization; fares; accessibility; public transit; equity**

1 Introduction

Public transit systems provide access to opportunities. Most transit operators charge fares, which makes this access contingent on users' ability to pay. To evaluate changes in fares or service, especially the equity of such changes with respect to users with different incomes, transit authorities should evaluate the range of travel time and monetary cost tradeoffs provided by different route choices within transit networks.

Fares affect both social equity and urban economies. Fare policy may include free transfers between vehicles within a time limit, discounted transfers between services, distance-based fares, and parallel services offering varying fares. Equity concerns are potentially magnified when there are multiple transit services charging different fares. For example, there might be a faster, more expensive train paralleling a slower, cheaper bus; those with less available income may resort to the bus, even if the train would provide a faster and more reliable trip. Distance-based fares or fare structures that require payment on each boarding

[*]School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA (formerly Conveyal, Washington, DC, USA); mwconway@asu.edu

[†]Conveyal, Watertown, MA, USA; astewart@conveyal.com

may disproportionately burden low-income travelers if they are more likely to travel long distances or make journeys with multiple transfers. Conversely, flat fare structures may disadvantage lower-income populations that live in central areas and take shorter trips. National policies often require such impacts to be evaluated (*e.g.* Martens and Golub 2018). From an economic standpoint, individuals whose access is curtailed by unaffordable fares may not be able to participate fully in the metropolitan economy, limiting its productive potential.

Analytical tools to assess the spatial and economic impacts of changes to transit service typically rely on shortest-path algorithms. A key computational challenge in applying such algorithms to fare changes is that fares are often path-dependent; that is, the fare for one ride often depends on past rides and whether they confer discounts on future rides. Thus, a more expensive way to reach an intermediate point may still produce the cheapest journey, causing issues for network analysis algorithms. Generalized costs, relying on values of time assumed *a priori*, or other simplifying assumptions, are often used to transform problems into ones solvable with typical optimization algorithms. Müller-Hannemann *et al.* (2007) conclude that given complex and non-additive fares, 'there is no hope to solve the cheapest connection problem exactly and simultaneously efficiently' and outline methods for simplification or approximation of fares within an optimization algorithm (p. 79).

This article presents an exact method for solving the cheapest-connection problem in a multiobjective optimization also considering travel time, accounting correctly for complex transfer incentives in large networks. Rather than combining the time and monetary cost dimensions into a single dimension, our multiobjective optimization technique produces the journeys that are *Pareto-optimal* in terms of travel time and cost. For a given origin and destination, the Pareto-optimal set represents the frontier beyond which no journey is both faster and cheaper. To calculate these optimal journeys, we formulate an abstraction of fare policies, which we call a *transfer allowance*. This abstraction makes it feasible to calculate cumulative-opportunity accessibility metrics with cutoffs for both time and fare (*e.g.* the number of jobs reachable from a given origin with limits of 60 minutes and $5).

After reviewing prior work, we prove our formulation of transfer allowance returns correct results. The method is implemented within an open-source web application for rapid-turnaround sketch planning.[1] A case study of a change to the transit fare structure of the Massachusetts Bay Transportation Authority (MBTA) in Greater Boston (Massachusetts, USA) is presented in the penultimate section.

2 Prior Work

2.1 Accessibility, Social Justice, and Fares

There is increasing interest in using accessibility metrics in transport planning (*e.g.* Miller 2018). Accessibility metrics have been used to evaluate the effects of potential land use or network changes (*e.g.* Anderson *et al.* 2013), as well as their equity implications (*e.g.* Manaugh and El-Geneidy 2012, Grengs 2012, Foth *et al.* 2013, Karner 2018). Geurs and van Wee (2004) distinguish four types of accessibility metrics; we focus on location-based metrics, which measure how many opportunities can be reached from a location. Such metrics are attractive because they require few behavioral assumptions and can effectively inform stakeholder deliberations (Stewart 2017).

Accessibility is key to numerous conceptions of a just transport network. Pereira *et al.* (2017) argue that, in a just transport system, all individuals should receive a minimum level of access, and that improvements to access should preferentially benefit those who are most disadvantaged by the transport system. Martens (2017) similarly uses the concept of 'accessibility insurance' to argue that society should provide some support for those individuals who experience insufficient accessibility. Even a strict utilitarian approach,

[1]Conveyal Analysis; source code available at `https://github.com/conveyal`.

without regard to equity, finds accessibility useful to define the effective size of the labor market and thus the overall productivity (*i.e.* the aggregate utility) of the urban agglomeration (Bertaud 2018, ch. 2).

Typical accessibility metrics evaluate access based on the travel time to various opportunities (*e.g.* Geurs and van Wee 2004). A commonly used metric, the cumulative opportunity accessibility metric, sums opportunities reachable from a given point within a chosen travel time cutoff. Fare costs have not typically been considered in the computation of accessibility, with two notable exceptions to date. El-Geneidy *et al.* (2016) implemented an accessibility metric in which travel time and fares are considered sequentially. They calculated fastest paths between centroids of analysis areas in Montréal at 7 AM, then used the transit agency's online fare calculator to calculate the fare for the fastest trip. However, the sequential approach can fail to find all journeys that meet the cost and time constraints. For instance, a low-cost bus parallel to a higher-cost rail line might meet the time and cost constraints, but not be returned from the shortest-path search because it is slower than the rail line. If the rail line does not meet the cost constraints, the sequential algorithm will incorrectly conclude that the destination is not reachable within the constraints.

Rodriguez *et al.* (2017) compared two fare systems in Bogotá using a similar time- and cost-constrained accessibility metric, computed using an earlier iteration of the algorithm used in the present paper. This allowed the retention of lower-cost paths even when they were not as fast as alternatives. However, the algorithm relied on the assumption that any path to a destination that was optimal in terms of fare was also optimal at every intermediate point along the path. While this is true of many transit systems, and in many others it is true for the vast majority of practical trips, it may be violated in systems with transfer incentives that provide discounts on future rides. (An example of such a violation is shown in Figure 1.) The method introduced herein relaxes this assumption, making it applicable to a broader variety of transit fare schemes.

2.2 Multiobjective Shortest-Path Searches

There has been significant research on multiobjective searches in transport networks. Many authors address the issue by constructing combinations of the multiple objectives and then using methods developed for single-objective optimization (*e.g.* Antsfeld and Walsh 2012). Cui and Levinson (2018) propose using the sum of a variety of costs in the shortest-path search component of accessibility metrics. In the context of a shortest-path search involving both transit fares and travel time as objectives, however, such algorithms depend on assumed values of time. This may raise equity concerns and lead to incorrect estimation; value of travel time varies by person and trip purpose (Hess *et al.* 2017), and in particular by income (Börjesson and Eliasson 2019). Additionally, combinations of many indicators are difficult to interpret and compare (O'Sullivan *et al.* 2000).

To address this criticism, others have developed algorithms to produce Pareto sets of co-optimal solutions that do not assume any defined functional form for tradeoffs between the objectives (*e.g.* Xie and Waller 2012, Kujala *et al.* 2018). Delling *et al.* (2015) introduce the RAPTOR algorithm and its multi-criteria extension (McRAPTOR) for multiobjective public transit routing. They demonstrate an application to minimizing the number of fare zones traversed, a simplified version of the problem this article addresses. However, none of these algorithms can account for the path-dependent property of many fare structures, as described in the introduction.

Others have taken a time-space approach, measuring accessibility in terms of the the opportunities that can be reached within a time-space prism. This represents a person-based accessibility measure, as defined by Geurs and van Wee (2004), as it is based on the space-time constraints of the individual. An early example from the GIS literature is Miller (1991), who documents an algorithm to compute space-time prisms and their two-dimensional corollary potential path areas, and proposes its use as an accessibility metric. Recently, Mahmoudi *et al.* (2019) extended the concept of space-time prisms to incorporate constraints other than travel time. They mention monetary costs, but focus on carbon budgets and the availability of en-route recharging stations for electric vehicle drivers.

Lo *et al.* (2003) present an exact solution to the issue of path-dependent fares, dividing a network into a set of modes and corresponding modal subnetworks (or, more generally, arbitrary groupings of transit vehicles; these groupings could be based on fare policy). They then define a small graph of probable transfers between these modes, based on assumptions about travel behavior; each node or 'state' in this graph implies not only the current travel mode, but also the previous travel modes and number of transfers. All nodes and edges in each modal subnetwork are duplicated for each state, and transfer edges are defined between these different state subnetworks. Since the nodes have been duplicated, two journeys which arrive at the same physical location via different modal sequences are at different nodes in the transformed network, and thus cannot be compared; more paths will be retained, potentially causing performance issues. In order to prevent these performance issues, Lo *et al.* suggest keeping the number of allowable sequences of modes (states) small. In contrast, the algorithm presented herein does not require behavioral assumptions about probable transfers, and allows comparisons between paths that use different modes.

3 Method

Our method is based on the McRAPTOR algorithm (Delling *et al.* 2015). Starting from the implementation of single-objective searches demonstrated in Conway *et al.* (2017), we add monetary cost as an objective function, which requires calculating fare paid at intermediate points *within* the routing algorithm, rather than after a full journey has been identified. The method we introduce allows the calculation of Pareto sets of time and monetary cost required to reach each destination in a region from a specified origin, using a public transit system with a complex fare policy.

Shortest-path algorithms start from a specified origin and build a tree-like structure of 'journey prefixes,' which are optimal ways to reach points in the network and represent a portion of a full journey between the origin and any intermediate point. These journey prefixes are then built upon, adding 'suffixes' to specific destinations as needed until an optimal journey to every point in the network has been found. The McRAPTOR algorithm proceeds in 'rounds' given a specified departure time. During round k it finds all optimal journey prefixes to each stop using rides on up to k transit vehicles (*i.e.* allowing $k-1$ transfers). During round $k+1$, it extends the journey prefixes found in round k with one additional ride, and updates the optimal journey prefix to reach each stop. Only 'dominant' journey prefixes—*i.e.* as good as or better than the alternatives in terms of the optimization criteria —are retained and used to explore onward travel in subsequent rounds. Others are pruned in order to make the algorithm tractable (Delling *et al.* 2015).

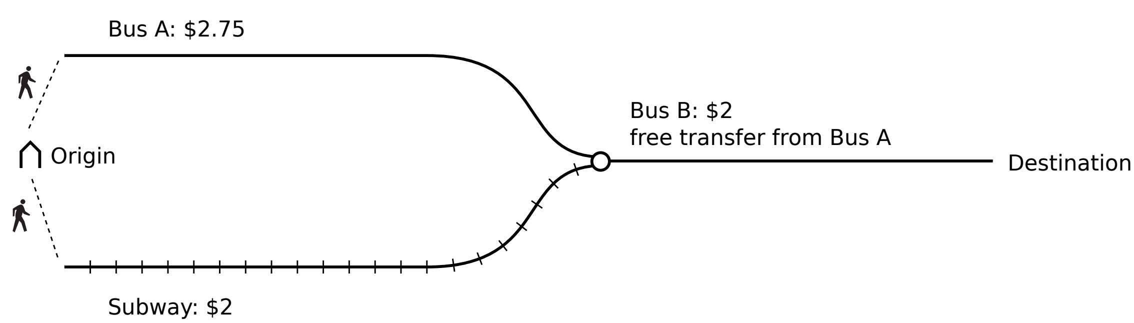

Figure 1: A simple transit system with transfer incentives illustrating why the lowest-cost path is not necessarily the lowest-cost at every intermediate point

One key advantage of the McRAPTOR algorithm for our use case is that journey prefixes are only compared after alighting from vehicles; there is no need for the fare on a vehicle to be a simple linear combination of the links traversed, since (partial) fares need not be compared for journey prefixes that end on board vehicles. The fare can be computed when alighting from the vehicle, based on what has been traversed so far. This obviates the need to connect every stop to every other stop with a link representing the appropriate fare, as was proposed by Lo *et al.* (2003).

The key challenge with performing a shortest-path search using fare paid as an objective function is that a portion of the least expensive journey is not necessarily the least expensive way to reach an intermediate point along that journey, due to transfer incentives. For example, consider the contrived transit system in Figure 1. In this system, the subway and Bus B both cost \$2. Bus A costs \$2.75 and has a free transfer to Bus B, whereas the subway has no free transfer. Suppose that the subway is faster than the buses. To get from the origin to the destination, one can take either Bus A or the subway to a transfer point, and then take Bus B to the destination. Clearly, the cheapest way to get to the destination is to take Bus A followed by Bus B. However, at the transfer point (prior to boarding Bus B), the subway is both cheaper *and* faster; algorithms that do not properly account for the path-dependence introduced by transfer incentives would consider the subway journey prefix to be strictly better than, *i.e.* to dominate, the Bus A journey prefix at the transfer point, leading the latter to be pruned. A standard shortest-path algorithm would not find the correct, lowest-cost full journey because of this premature pruning.

More generally, correct algorithms will only prune one journey prefix in favor of another if the latter is equal or better in terms of the full journey fare, regardless of what onward path ('journey suffix') is taken. Using the notation in Table 1, for journey prefix P reaching point x to be better than or equal to another journey prefix Q reaching x, it must be shown that the fare for full journey PS is less than or equal to the fare for full journey QS, for all possible journey suffixes S. Expressed mathematically, this 'domination

Symbol	Definition	First Used
P, Q	Journey prefixes (*i.e.* combinations of transit rides) reaching a particular point x	(1)
S	Journey suffix (*i.e.* combination of transit rides) reaching from x to any other destination in the network	(1)
PS, QS	Full journeys from the origin to a destination formed by linking P or Q with S	(1)
U	Set of all journey suffixes originating at point x	(1)
f_P, f_Q	Cumulative fare for journey prefix P or Q, respectively	(5)
t_P, t_Q	Arrival time at point x for journey prefix P or Q, respectively	(2)
f_{PS}, f_{QS}	Fare paid for journey prefix P or Q, respectively, followed by journey suffix S	(1)
f_S	Full fare for journey suffix S, assuming no discounts due to previous rides	(5)
a_{PS}, a_{QS}	Transfer allowance (discount from full fare f_S) due to previously riding journey prefix P or Q, respectively	(3)
a_P, a_Q	Maximum discount on any journey suffix that can be obtained by riding journey prefix P or Q, respectively	(4)

Note: P and Q are arbitrary journey prefixes; they can be exchanged in any definition herein and the definition holds.

Table 1: Nomenclature

rule'—a rule used to determine when one journey suffix dominates another—for fare paid is:

$$f_{PS} \leq f_{QS} \quad \forall\, S \in U \tag{1}$$

where f_{PS} and f_{QS} are the fare for full journeys PS and QS, and U is the set of all journey suffixes originating at point x. Note that it is not critical that Q always be rejected if (1) holds; if it is not, the algorithm will run more slowly due to exploring suboptimal results, but will still find all Pareto-optimal results (any suboptimal results are easily removed in a postprocessing step). However, it *is* critical that Q never be rejected if (1) does not hold; otherwise, potentially optimal solutions may not be found. In short, our method must add slack to the pruning process, retaining states that could potentially be part of optimal full journeys, while staying tractable.

To find all Pareto-optimal journeys using travel time and fare, a domination rule for travel time is also needed. Because travel time is additive—there is no way that arriving at an intermediate point earlier can make you any later to your destination—the domination rule for travel time is simply

$$t_P \leq t_Q. \tag{2}$$

If (2) holds, it is trivial to show that replacing Q with P cannot cause any journey to have a longer travel time; this concept is fundamental to the RAPTOR search process of Delling *et al.* (2015).

In order to produce a Pareto set of optimal results, we consider a journey prefix Q dominated by alternate journey prefix P iff both (1) and (2) hold. If both hold, Q can never lead to a shorter or lower-fare full journey than P, so Q can safely be pruned. While evaluating whether (2) holds is straightforward, for most real-world systems it is intractable to evaluate (1) directly. The remainder of this section discusses refined domination rules that can be used in place of (1) to achieve a tractable algorithm.

3.1 Defining transfer allowance

These additional domination rules rely on the concept of a *transfer allowance,* the amount of money one could save on a future ride by having taken a certain sequence of transit vehicles in the past. In the simple example above, the transfer allowance from Bus A to Bus B is \$2 (because a user stands to save \$2—the full fare—on Bus B thanks to the free transfer). The transfer allowance from the subway to Bus B is \$0, because there are no transfer discounts for subsequent rides. Another way to conceptualize the transfer allowance is as the additional cost a user could incur if they were to lose their ticket or transfer slip before boarding the subsequent service. Mathematically stated, the transfer allowance a_{PS} for journey prefix P and journey suffix S is

$$a_{PS} = f_P + f_S - f_{PS}, \tag{3}$$

with a_{QS} defined similarly. We then define the maximum transfer allowance for journey prefix P as

$$a_P = \max(a_{PS} \quad \forall\, S \in U), \tag{4}$$

again with a_Q defined similarly.

We assume that $a_{PS} \geq 0$ and $a_{QS} \geq 0$ for all S. That is,

$$f_{PS} \leq f_P + f_S \quad \forall\, S \in U \tag{5}$$

and

$$f_{QS} \leq f_Q + f_S \quad \forall\, S \in U. \tag{6}$$

If either of these conditions were not true, the user would be better off discarding their ticket at x, *i.e.* they have an *negative* discounted transfer. We assume that no reasonable transit system exhibits this situation.[2]

[2] If such systems do exist, users would likely be clever enough to discard their tickets when advantageous. While many airline contracts of carriage disallow throwaway ticketing ploys, we are unaware of any transit agencies that have codified such restrictions.

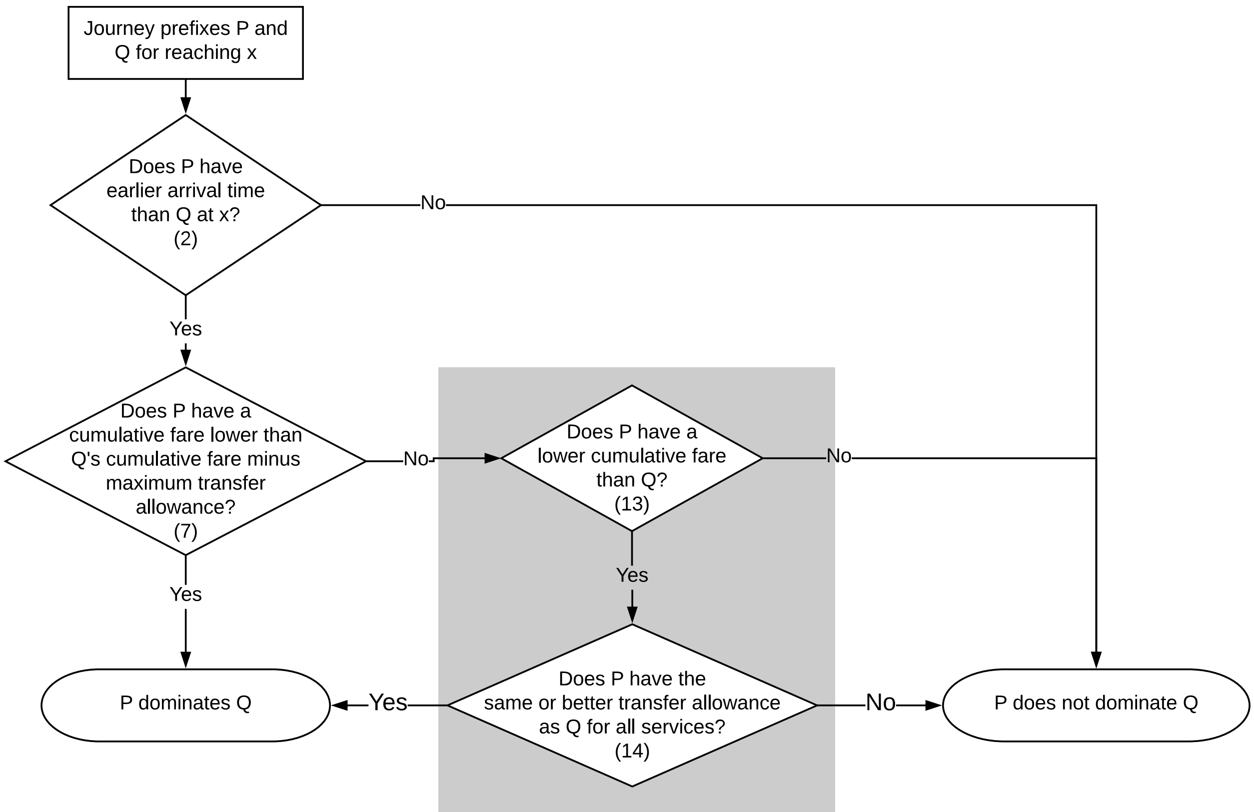

Figure 2: Domination rules for Pareto-optimal search on travel time and fares, with references to relevant equations

3.2 Fare domination rules using transfer allowances

We use the concept of the transfer allowance in two simplified domination rules to determine if (1) holds for a particular pair of journey prefixes P and Q. These fare domination rules are checked after the time domination rule, and are outlined in Figure 2 and explained in detail below.

The first fare domination rule checks whether the cumulative fare paid for P is less than the fare paid on Q *minus the maximum value of all possible transfer allowances from* Q. For instance, suppose a user stood to save up to $1 on future fares by taking journey prefix Q (*i.e.* $a_Q = 1$), but had already paid $6 for that journey prefix. If the alternate journey prefix P only costs $3, P is strictly better than Q, regardless of any future discounted transfers. Even if taking P rather than Q required the user to forego a future savings of $1, it is worth it because the user has saved $3 on getting to the intermediate point. Note that the maximum value of the transfer allowances from P is not considered; we do not know if they provide discounted transfers to the same services. Even if that maximum value of P's transfer allowance was equal to the value of Q's, the user might still be required to forego a future discount on a service that has a discounted transfer from Q but not from P.

If this condition holds, then (1) holds and P is as good as or better than Q in terms of fare for all possible journey suffixes, as proven in Theorem 3.1; Q can safely be rejected. Note that the opposite is not necessarily true; if this condition does not hold, (1) is not necessarily false.

Theorem 3.1. *If*

$$f_P \leq f_Q - a_Q \tag{7}$$

then (1) holds.

Proof. By substituting (4) and (3) into (7), we have

$$f_P \leq f_Q - \max(f_P + f_S - f_{PS} \quad \forall\, S \,\in\, U) \tag{8}$$

which, due to the use of the max operator, implies

$$f_P \leq f_Q - (f_Q + f_S - f_{QS}) \quad \forall\, S \,\in\, U \tag{9}$$

simplifying terms

$$f_P \leq f_{QS} - f_S \quad \forall\, S \,\in\, U \tag{10}$$

and adding f_S to both sides

$$f_P + f_S \leq f_{QS} \quad \forall\, S \,\in\, U \tag{11}$$

By combining (11) with (5), we see that

$$f_{PS} \leq f_P + f_S \leq f_{QS} \quad \forall\, S \,\in\, U \tag{12}$$

which is equivalent to (1). Theorem 3.1 is proved. □

Note that we are *not* assuming that the fare to reach the intermediate point is less than or equal to the full journey fare ($f_P \leq f_{PS}$); this assumption can be violated in systems where it costs more to exit the system at an intermediate point than to reach another station via a transfer at that point. For instance, in the San Francisco Bay Area Rapid Transit (BART) system, traveling from San Francisco International Airport (SFO) to Millbrae costs \$4.55. At some times of day, a transfer is required at San Bruno. The fare from SFO to San Bruno is \$7.85 (San Francisco Bay Area Rapid Transit District n.d.), but a passenger will only be charged this amount if they exit the paid area at San Bruno; if they board a Millbrae-bound train, they will be charged the lower fare when they exit the system. By setting the maximum transfer allowance to \$3.30, the path to Millbrae can be retained, even if the fare cutoff is less than \$7.85.

In some transit systems, this domination rule will not eliminate enough codominant states to yield a tractable algorithm. Consider a bus network where the first ride costs \$2 and all subsequent rides are free. Thus, after riding the first vehicle, the fare paid is \$2 and all transfer allowances are also \$2, because the user stands to save \$2 at all subsequent boardings.[3] No journey prefix will ever be better than or even equal to any other (because the fare paid f_P (\$2) will always be greater than the fare paid f_Q (\$2) minus the maximum transfer allowance a_Q (also \$2) on an alternate route; (7) will never hold), so all possible routes will be retained. The routing algorithm will simply ride every possible bus, doubling back on itself with great abandon.

We accordingly add a second domination rule (shown in the shaded area of Figure 2) for situations where two journey prefixes share the same fare characteristics, or one has strictly better fare characteristics. Consider again two journey prefixes P and Q. If P has the same or lower cumulative fare and the same or higher transfer allowance to every possible journey suffix as Q, rejecting Q cannot eliminate any journeys QS that are strictly better than a similar journey PS possible with journey prefix P, as proven in Theorem 3.2.

Theorem 3.2. *If both*

$$f_P \leq f_Q \tag{13}$$

and

$$a_{QS} \leq a_{PS} \quad \forall\, S \,\in\, U \tag{14}$$

hold, we can conclude that (1) holds.

[3] One might guess that the allowance is ∞ because one can save \$2 at *each* subsequent boarding, but this is actually false; the user can save \$2 at the next boarding over paying the full fare, but if they were to instead pay the full fare, they would receive the transfer allowance at the next boarding. Thus, regardless of how many vehicles have been ridden, after the first vehicle is ridden, the fare paid and the maximum transfer allowance are both \$2.

Proof. Once again, we wish to determine if

$$f_{PS} \leq f_{QS} \quad \forall\, S \,\in\, U \tag{1 repeated}$$

holds.

Adding (13) and (14), we have

$$f_P + a_{QS} \leq f_Q + a_{PS} \quad \forall\, S \,\in\, U \tag{15}$$

replacing a_{QS} and a_{PS} with (3)

$$f_P + f_Q + f_S - f_{QS} \leq f_Q + f_P + f_S - f_{PS} \quad \forall\, S \,\in\, U \tag{16}$$

subtracting $f_P + f_Q + f_S$ from each side

$$-f_{QS} \leq -f_{PS} \quad \forall\, S \,\in\, U \tag{17}$$

which is equivalent to (1). Theorem 3.2 is proved. □

There is generally a transfer limit imposed on such a search process. Initially, this might seem problematic; if a one-transfer journey prefix P is pruned and replaced with a two-transfer journey prefix Q, and the only way to get to a particular destination requires additional transfers, it is possible that a route with prefix Q would not be able to reach the destination at all, due to the transfer limit. However, because we use the McRAPTOR algorithm, this is unproblematic. Recall that in the McRAPTOR algorithm, each round k finds all optimal journey prefixes with $k - 1$ transfers (Delling *et al.* 2015). Thus, if a one-transfer journey prefix is pruned due to a better two-transfer journey prefix being found, the one-transfer journey prefix P will already have been used as the basis for a new prefix R with one additional transit ride, and thus the route to the destination will still be found.

The algorithm can return some journeys that are not Pareto-optimal. For example, consider a case where journey prefixes P and Q are codominant at a stop x near a particular destination (perhaps they both provide discounted transfers to different sets of services). Suppose that a bus A is used to reach the destination from x. Journeys using both P and Q to reach bus A will be found, even though one of them may not be Pareto-optimal; thus, a trivial postprocessing step removes any journeys that are not Pareto-optimal. When computing the accessibility results presented in the Case Study section, we further process this Pareto set in order to return the fastest journey that meets the given fare constraint, but of course other rules could be devised.

Both (1) and (2) use non-strict inequalities (*i.e.* $\leq$). Thus, two journey prefixes P and Q that are equivalent (same travel time, cumulative fare paid, and transfer allowance to all journey suffixes) will each dominate each other, and which one will be retained depends on the order in which they are checked. This still correctly identifies the Pareto frontier; the results of all journeys with P as prefix are identical in terms of time and cost to the same journeys with Q as prefix, and thus occupy the same location on the Pareto frontier. For any point on the Pareto frontier, however, not all paths that produce that point on the frontier are retained. While this is sufficient for calculating accessibility—we only need to know the time and fare to compute an accessibility metric, not any exact path—it may not be satisfactory for a customer-facing journey planner, a network-assignment algorithm, or as input to a route choice model. Some options for retaining more of the paths to extend the algorithm to these use cases are presented in Section 5.1.

3.3 Transfer Allowances in Practice

While we have shown above that our algorithm works in theory, the proofs above present two practical challenges, in the form of equations (4) and (14), reproduced here for convenience:

$$a_P = \max(a_{PS} \quad \forall\, S \,\in\, U) \tag{4 repeated}$$

$$a_{QS} \leq a_{PS} \quad \forall\, S \,\in\, U \tag{14 repeated}$$

Both of these conditions must be checked for U (all possible journey suffixes from the given location). The difficulty in implementation, of course, is that identifying all possible journey suffixes, let alone evaluating fares for all of them, is likely to be intractable. However, most transport systems have a relatively small number of classes of transfer allowance. If the classes are collectively exhaustive, defining transfers to *all* possible journey suffixes, (4) and (14) can be solved for only a small number of representative journey suffixes covered by these classes of transfer allowance, without loss of generality.

In some cases, including the Boston case study presented Section 4, there are a small number of possible 'bundles' of transfer allowances (for instance, a bus ride may afford a 'bundle' offering a free transfer to additional buses, and a discount on rail). There may be few enough distinct bundles of possible transfer allowances that tractability can be achieved by simply treating journey prefixes with different bundles as incomparable; rather than directly evaluating (14), we can show trivially that (14) holds when two journey prefixes have the same bundle of transfer allowances, and assume (14) does not hold when they do not (recall that failing to reject journey prefixes when there is another journey prefix that is better or equal does not result in incorrect results, it merely slows the algorithm somewhat).

This is conceptually similar to the aforementioned algorithm of Lo *et al.* (2003), who effectively consider two journey prefixes leading to the same location to be incomparable if they do not have the same sequence of modes leading up to them. There are, however, several key differences. We do not need to specify allowable sequences of modes *a priori*. We can also eliminate very costly journey prefixes using Theorem 3.1, even if the transfer allowances are incomparable. Perhaps most importantly for performance, we can compare journey prefixes that have not taken the same sequence of modes but afford the same classes of transfer allowance, even if they have different fares paid.[4]

Many systems have either a time limit or a transfer limit on their transfer allowances. For instance, passengers may be able to board additional buses for free within two hours of boarding the first bus, or be able to transfer to no more than one additional bus (or both). Thus, even if the potential cash discount on future services is the same for two transfer allowances resulting from riding journey prefixes P and Q, one or the other may be preferable due to a longer time remaining or more transfers remaining. In order to account for this while still allowing classes of transfer allowance to be defined, we add conditions that, in order for P to dominate Q by Theorem 3.2, it must have the same or more time remaining, and the same or more transfers remaining, for all classes of transfer allowance.

3.4 Speedup techniques

We start from the efficient algorithm of Delling *et al.* (2015), as implemented by Conway *et al.* (2017). Performing a multiobjective path search is considerably slower than a comparable single-objective path search, as many more paths are retained. To make our algorithm tractable and sufficiently fast for interactive sketch planning applications, returning accessibility results for a single origin within seconds, we must therefore undertake additional optimizations, described below.

First, the algorithm prunes any journey prefixes that have a travel time longer than the cutoff time used in the calculation of the cumulative accessibility metric. It also prunes journey prefixes that have a cumulative fare *minus maximum transfer allowance* greater than the fare cutoff. While in many systems it would be appropriate to prune when the cumulative fare exceeds the fare cutoff, we wish to support

[4]For instance, consider a system where at each bus boarding you get a transfer slip to board one additional bus, and after that you must pay full fare again. Thus, the transfer allowance afforded by a single bus journey prefix is the same as that afforded by a bus → bus → bus journey prefix, although the fare for the latter option is higher. While these options would be incomparable under Lo *et al.*'s framework, our framework can compare them (and will only retain the three-bus journey prefix if it is better in terms of time than the one-bus option).

situations where the fare to a transfer station exceeds the fare for the full journey, such as the BART example in Section 3.2.

Second, we use a random sample of the possible departure times. The algorithm in Conway *et al.* (2017) calculates accessibility for journeys starting at every minute of a departure window. This exhaustive approach allows for a robust consideration of frequency and network effects, but computation may be impractically slow for interactive sketch planning.

We do not use a simple random sample of departure times from the time window, but rather a 'constrained random walk,' as proposed by Owen and Jiang (2015). This approach yields times that are random but relatively evenly distributed over the journey starting time window, and performs better than other sampling strategies in terms of approximating the true value that would be achieved if every departure minute was sampled (Owen and Jiang 2015).[5] Starting from the beginning of the time window, we choose the first sampled departure time from a random uniform distribution over $[b, b + f]$ where b is the beginning of the time window and f is the desired sampling frequency. f is equal to l/n, where l is the length of the time window, and n is the desired size of the random sample. Subsequent departure times are chosen by adding a number sampled from a random uniform distribution over $[f/2, f + f/2]$ to the previously-sampled departure time.

Returned samples may not have size of exactly n, due to randomness. We desire a sample size of exactly n, as we will use a percentile of the returned travel times to calculate the accessibility, and if we set n appropriately we can compute an exact percentile without resorting to interpolation. Thus, as proposed by Owen and Jiang, we simply generate samples until obtaining one of size n.

Third, we cap the time remaining on transfer allowances (as described in Section 3.3) to be no longer than the remaining time given the travel time budget for the accessibility metric we use. For instance, if we wish to compute the number of opportunities available within 60 minutes travel time and a five-dollar fare, and the transit system has a two-hour free transfer window, we will set the expiration time of *all* transfers to be 60 minutes after the departure time from the origin. This way, if journey prefix P has a faster travel time than journey prefix Q, but journey prefix Q has a later boarding time but otherwise provides identical transfer allowances to all services, P will still be able to dominate Q. The longer time remaining in the transfer window from Q is immaterial, because any rides that took advantage of the longer window would exceed the travel time budget by definition.

4 Case Study

This case study evaluates cost-constrained accessibility for both a baseline scenario and an illustrative scenario with reduced prices for some commuter rail fares. While commuter rail in Greater Boston largely serves trips from outlying residential suburbs to downtown employment centers, a number of inner stations serve relatively short, local trips. To accommodate these trips, 2018 baseline fare policy assigns selected inner stations a special zone, within which the base fare is $2.25 (equivalent to the flat subway fare). Given the large increment in fares when crossing the boundary between this special Zone 1A and other zones, the re-assignment of additional stations to Zone 1A has been the subject of political pressure over the years (*e.g.* Vaccaro 2018). Additionally, the MBTA is currently developing a new fare-payment system (Massachusetts Bay Transportation Authority n.d.a), spurring conversation about fare policy.[6] While

[5] Owen and Jiang's method is slightly different from ours; they produce an accessibility figure for each minute, and average them, while we produce a specific percentile of travel time across all possible departure times for each origin-destination pair and use that to calculate accessibility. We believe Owen and Jiang's findings are still relevant. For more discussion on the tradeoffs between these two methods, see Conway *et al.* (2017, 2018).

[6] Boston also has a long history of activism surrounding transit fares more generally, as evidenced by the iconic protest song 'M.T.A.,' which bemoans the fate of a traveler, Charlie, who lacked the subway exit fare and 'may ride forever 'neath the streets of Boston.' Originally written to protest a subway fare increase, the song is now arguably part of the New England vernacular (Dreier

low-income individuals take the majority of transit trips in the United States, they take far fewer commuter rail trips (Pucher and Renne 2003, Manaugh and El-Geneidy 2012); lowering the fares for short commuter rail trips (rather than to wealthy, far-flung suburbs) may somewhat ameliorate this discrepancy.

The MBTA fare system includes a range of diverse fare structures and transfer policies, allowing a test of the proposed algorithm in a real-world case with a complex fare system. The 2018 fare structure includes flat fares for local bus and rapid transit routes, fares differentiated by distance for express bus and ferry routes, and zone-based fares for commuter rail (Table 2). In most cases, the transfer policy permits one 'pay-the-difference' transfer between and among local bus, express bus, and rapid transit routes, available for two hours after the first boarding. While our illustrative scenario only evaluates changes to certain commuter rail fares, baseline service and fares for all MBTA modes are included in the analysis to capture network effects.

Various pass products (*e.g.* weekly, monthly) and fare media (*e.g.* paper ticket, mobile phone application, contactless card) are available. We consider only pay-as-you-go fares using a single MBTA contactless card (CharlieCard) where accepted—which generally provides the lowest pay-as-you go costs and the most permissive transfer policies—and cash otherwise.[7] CharlieCards are the predominant method of fare payment at the MBTA; over 75% of all boardings, and a majority of pay-as-you-go boardings, are paid with CharlieCard (Stuntz *et al.* 2017). While a majority of passengers use passes (Stuntz *et al.* 2017), pay-as-you-go fares are still appropriate for understanding cost constraints on commuting. Low-income riders may be unable to pay up-front cost of monthly passes, driving them to weekly passes (Verbich and El-Geneidy 2017) or even pay-as-you-go fares. Additionally, low-income workers are more likely to work non-standard schedules (Enchautegui 2013), and may also work part-time, so their travel may not be sufficiently regular to warrant a pass.

Each commuter rail station is assigned a zone based loosely on distance from two downtown terminals. The existing zones are shown in Figure 3. As previously mentioned, fares within the central Zone 1A are \$2.25. The one-way fare between stations in the central zone (Zone 1A) and outlying stations ranges from \$6.25 (Zone 1) to \$12.50 (Zone 10). 'Interzone' fares between stations outside Zone 1A range from \$2.75 to \$7.00 depending on how many zones are traversed (Table 2).

As temporary mitigation for a subway station closure in 2018, one station (Quincy Center) was designated a special Zone 1A/1 boundary station. This arrangement makes \$2.25 fares available to/from other Zone 1A stations, including the downtown terminals, without raising prices for commuters who use interzone fares to/from outlying stations (*e.g.* a traveler from Brockton to Quincy Center who pays a \$4.00 Interzone 4 fare and who would have to pay \$8.25 if Quincy Center were assigned exclusively to Zone 1A).

Our illustrative scenario re-assigns all stations currently in Zones 1 and 2 (including Quincy Center) to a new Zone 1A/2. Trips within this zone, or between 1A/2 and 1A, are charged \$2.25, while trips between a Zone 1A/2 station and any outlying zone are charged the interzone fare they would be charged if the station were in Zone 2. Our scenario does not create any transfer incentives between commuter rail lines, or between commuter rail lines and any other service. We make no claims about the feasibility, ridership impacts, or fiscal prudence of this fare reduction scenario. It is meant only to demonstrate the characteristics of our algorithm and provide an example of calculating and interpreting cost-constrained accessibility metrics.

---

and Vrabel 2010).

[7] Our algorithm cannot fully model the CharlieCard transfer discount system, because it violates the nonnegativity of transfer allowances assumptions specified in equations (5) and (6). For instance, a user making a Local Bus → Inner Express Bus → Subway journey will pay \$6.25 (\$1.70 local bus, \$2.30 upgrade to express bus, and \$2.25 for the subway), but if they were to discard their transfer allowance after the first ride they would pay only \$5.70 (\$1.70 local bus, \$4 express bus, and a free transfer to subway). A user could also achieve this price if they rode a second local bus before boarding the express bus, expending their transfer allowance. The consequence of this violation is that a small number of lowest-fare trips that rely on such ploys may not be found. We suspect few users rely on such trips. This is a specific idiosyncracy of the MBTA fare system and does not undermine the correctness of the algorithm in more common situations where transfer allowances are nonnegative.

Full Fare Boarding	**Fare**	**Discounted Transfers**	**Transfer Cost**
Local Bus	$1.70	→ Local Bus	$0.00
		→ Subway†	$0.55
		→ Subway† → Local Bus	$0.55
		→ Inner Express Bus	$2.30
		→ Outer Express Bus	$3.55
Subway†	$2.25	→ Local Bus	$0.00
		→ Inner Express Bus	$1.75
		→ Outer Express Bus	$3.00
Inner Express Bus	$4.00	→ Local Bus	$0.00
		→ Subway†	$0.00
Outer Express Bus	$5.25	→ Local Bus	$0.00
		→ Subway†	$0.00
Inner Ferry	$3.50	None	
Outer Ferry	$9.25	None	
Commuter Rail		None	
Zone 1A ⇆ Zone 1A	$2.25		
Zone 1 ⇆ Zone 1A	$6.25		
Zone 2 ⇆ Zone 1A	$6.75		
Zone 3 ⇆ Zone 1A	$7.50		
Zone 4 ⇆ Zone 1A	$8.25		
Zone 5 ⇆ Zone 1A	$9.25		
Zone 6 ⇆ Zone 1A	$10.00		
Zone 7 ⇆ Zone 1A	$10.50		
Zone 8 ⇆ Zone 1A	$11.50		
Zone 9 ⇆ Zone 1A	$12.00		
Zone 10 ⇆ Zone 1A	$12.50		
Interzone 1	$2.75		
Interzone 2	$3.25		
Interzone 3	$3.50		
Interzone 4	$4.00		
Interzone 5	$4.50		
Interzone 6	$5.00		
Interzone 7	$5.50		
Interzone 8	$6.00		
Interzone 9	$6.50		
Interzone 10	$7.00		

All discounted transfers must be used within 2 hours of initial boarding and fare payment.
† Includes an unlimited number of free transfers between subway and light-rail lines, for an unlimited amount of time, as long as user does not leave the paid area of the subway system.

Table 2: MBTA CharlieCard Fares (adapted from Massachusetts Bay Transportation Authority n.d.b)

Figure 3: MBTA Commuter Rail fare zones, existing

4.1 Data Needs

MBTA schedule and network data are available in General Transit Feed Specification (GTFS) format. The MBTA does not, however, use GTFS fare rule or fare attribute tables. For this project, we augmented the GTFS with commuter rail zones and prices. We then programmed into R5 (open-source routing software with a McRAPTOR implementation) a custom fare calculator for the MBTA, which reads the augmented GTFS feed and handles some special cases created by the MBTA's fare policy and network configuration (*e.g.* stations where opposite-direction platforms are not connected behind fare gates, or stations connected by pedestrian tunnels). For brevity, we do not detail here the handling of these special cases; interested readers are referred to the annotated source code for R5, version `95d0768`, available online.[8] To account for walking time on the street network to access transit stops, we use data from OpenStreetMap.[9] Block-level data on job locations, used to compute accessibility metrics, are from the Longitudinal Employer-Household Dynamics program's Origin-Destination Employment Statistics.[10]

4.2 Analysis parameters

For the results below, the constrained random walk approach described in Subsection 3.4 was used to sample 40 departure times between 7:00 AM and 8:20 AM (*i.e.* average sampling frequency $f = 2$ minutes). Up to three transfers are allowed, which we believe encompasses the vast majority of practical trips in Boston. Travel time includes walking up to 20 minutes at 5 km/hr to/from transit stops, in-vehicle time, and waiting time according to scheduled services in the MBTA GTFS on July 18, 2018. We use the 5^{th} percentile of the distribution of travel times (which is composed of the fastest travel time possible given the fare constraint for each sampled departure time). If, for example, the only transit service for an origin-destination pair departed once in this 80-minute departure window, at 8:00, the 5th percentile total travel time would include 4 minutes of waiting time. While higher percentiles of travel time are recommended for assessing the impact of frequency on accessibility (Conway *et al.* 2018, p. 544), the 5^{th} percentile is appropriate for travelers who have accurate schedule information and flexibility about when they start their trip within a specified window, typical for commuter rail users.

4.3 Results

4.3.1 Baseline Results

Figure 4 shows the frontier of Pareto-optimal cost and travel time values, using 5^{th} percentile total travel times within the departure window, for the example origin-destination pair of Norwood and Copley Square. Norwood is a suburban town with bus service as well as Franklin Line commuter rail service at Zone 3 stations. Copley Square is a centrally-located employment cluster, within a short walk of commuter rail lines at Back Bay Station (see Figure 3). The fastest option, riding a direct commuter rail train, takes 40 minutes and requires the payment of a Zone 3 fare ($7.50). The cheapest transit option, riding two buses, takes 95 minutes and requires the payment of one bus fare ($1.70), which includes a free transfer to the second bus. Other Pareto-optimal solutions include transferring from bus to subway, or transferring at intermediate commuter rail stations (*e.g.* Readville) to take advantage of interzone fares. This example illustrates the shortcoming of sequentially obtaining shortest-time journeys, then calculating the fare for those journeys. If the sequential approach, rather than computing Pareto sets within the routing algorithm, were used to

[8] https://github.com/conveyal/r5/blob/95d076896e0a076207ced83742deb8fb620127df/src/main/java/com/conveyal/r5/analyst/fare/BostonInRoutingFareCalculator.java

[9] `https://openstreetmap.org`; data © OpenStreetMap contributors.

[10] `https://lehd.ces.census.gov/data/#lodes`

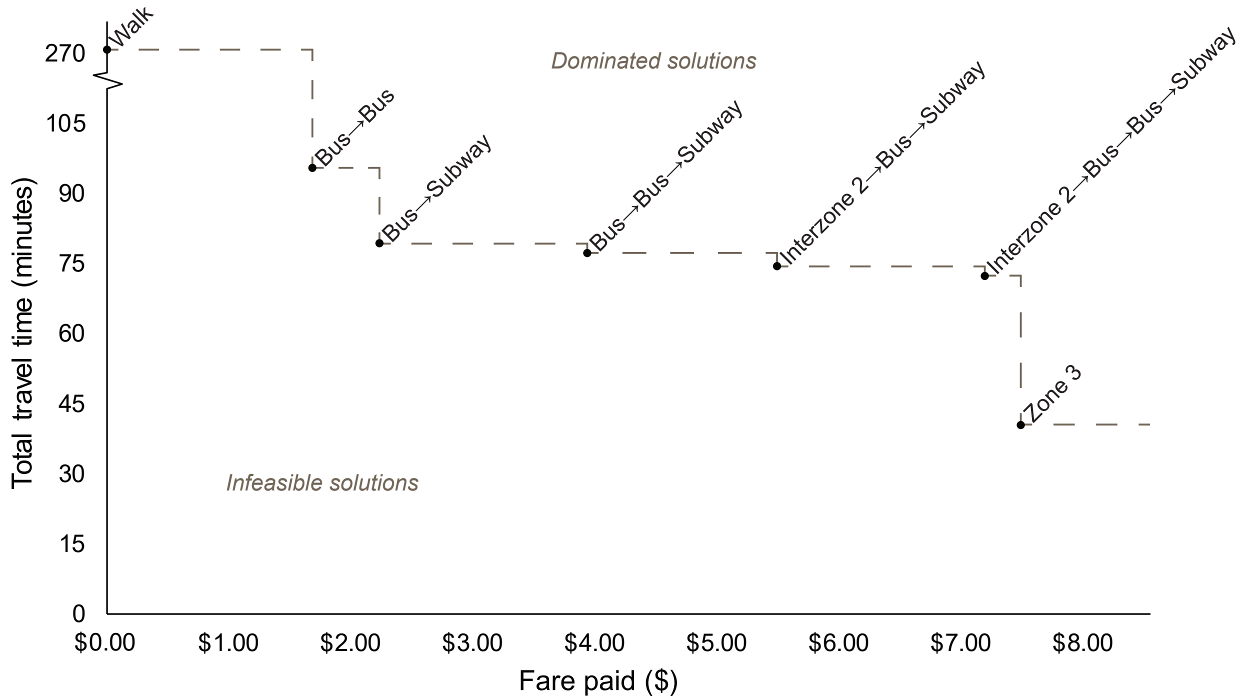

Figure 4: Pareto frontier, for travel from Norwood Central to Copley Square

assess accessibility with a monetary cutoff less than $7.50, Copley would be considered unreachable from Norwood, even though it can be reached for as little as $1.70 using slower alternatives.

Figure 5 shows cost 'contour maps' for trips originating from Copley Square, with travel time cutoffs of 60 and 120 minutes. For a travel time cutoff of 60 minutes, the cost to reach Norwood from Copley Square exceeds $7.00 (Figure 5(a)); but travelers willing to endure up to 120 minutes of travel time can pay less than $2.00 (Figure 5(b)), consistent with Figure 4.

A similar example can be seen for the stations north of Brockton, which are reachable from Copley within 60 minutes for someone paying $10.50 (Subway + Zone 4 fares). If one instead is willing to travel for up to 120 minutes (Figure 5(b)), a slower option costing only $5.75 is available—riding the Green Line to the Red Line to Braintree ($2.25), then transferring to commuter rail paying an Interzone 3 fare ($3.50) to Brockton. This example demonstrates why, at least in the Boston case, the Pareto frontier cannot be calculated by simply toggling more expensive modes such as commuter rail. Commuter rail is the only MBTA service at these stations, but the zone structure creates different options for time-cost tradeoffs.

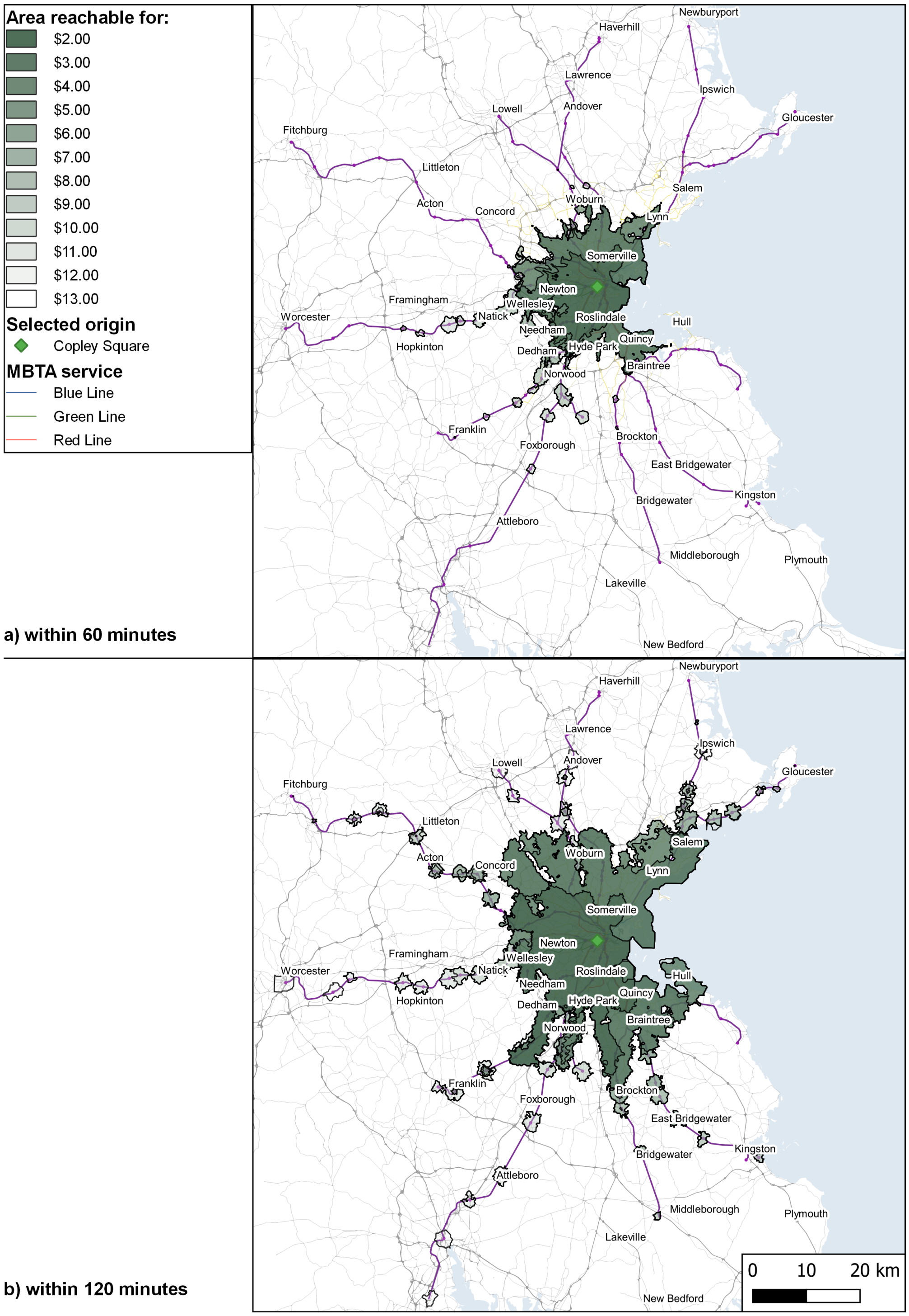

Figure 5: Cost contour maps showing fare to reach every location in the Boston region, from Copley Square, baseline

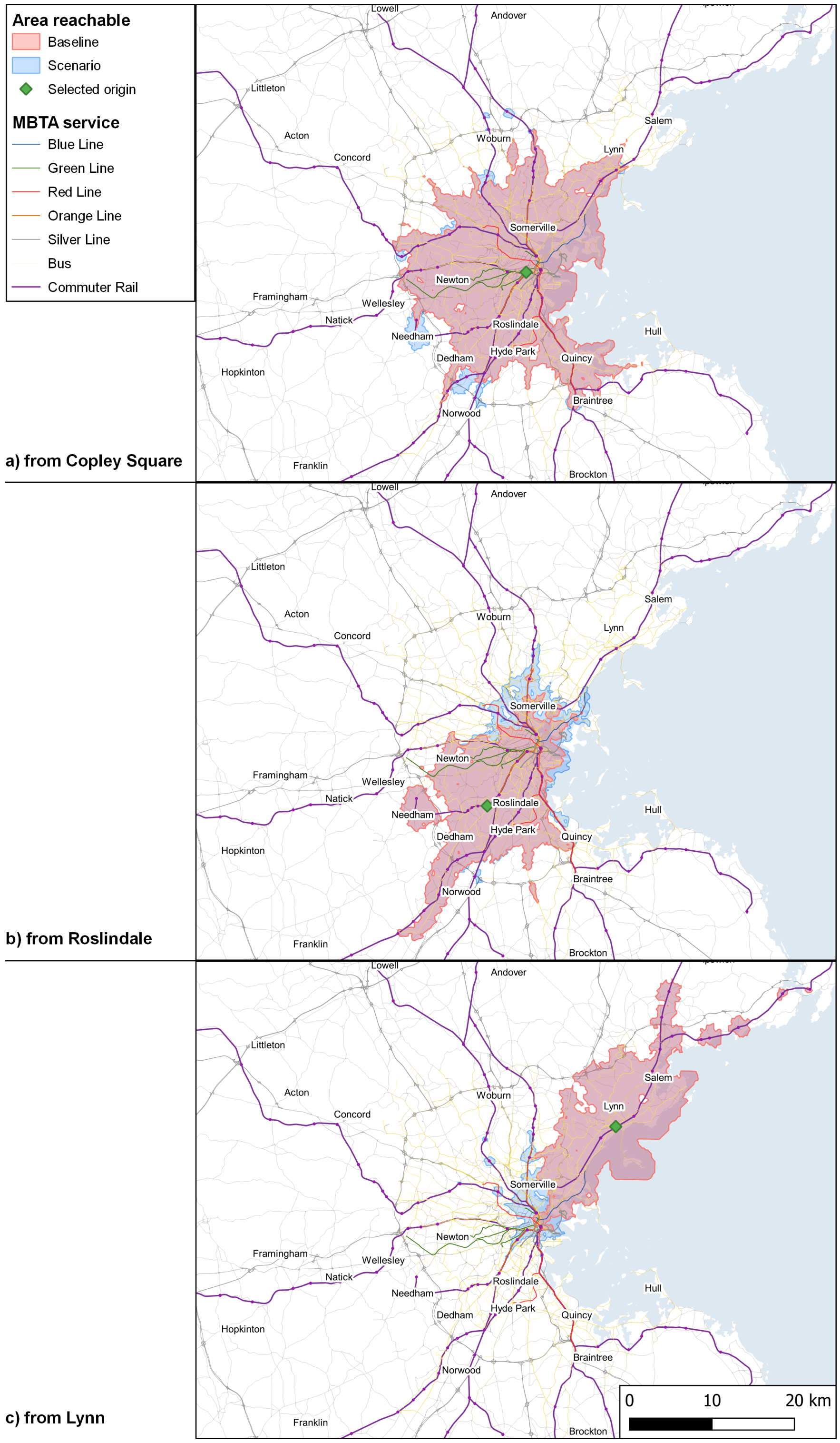

Figure 6: Reachable area within 60 minutes for less than $5, scenario vs. baseline

4.3.2 Scenario Impacts

The impact of the illustrative scenario, reassigning stations in Zones 1 and 2 to Zone 1A/2, is evaluated below using example cutoffs of 60 minutes and $5.00. In Figure 6, the reachable area under the baseline fare system is shown in red, while the reachable area under the illustrative scenario is shown in blue. As Figure 6(a) shows, the illustrative scenario enables access from Copley Square to places such as Braintree, Dedham, Needham, and Woburn within the example time and cost cutoffs.

The impact of the scenario is more substantial from Roslindale, a mixed-income residential neighborhood beyond the reach of the rapid transit system (Figure 6(b)). In the baseline, a $5.00 limit effectively limits people starting trips in Roslindale to taking a local bus to Forest Hills (the end of the rapid transit system, as well as a hub for buses), then transferring. The number of jobs reachable given these time and cost constraints (*i.e.* the number of jobs located within the red isochrone) is 724,869. The illustrative scenario's reduced fares would allow people to ride commuter rail then transfer, unlocking access to areas north, east, and south of Downtown, as well as south toward Norwood. With the scenario's fares, 788,318 jobs would be accessible with $5.00 and in 60 minutes—an increase of 63,449 over the baseline.

Lynn is a lower-income community north of Boston, also beyond the reach of the rapid transit system. From Lynn, the downtown core is barely reachable in the baseline given the 60 minute and $5.00 constraints (Figure 6(c)). The scenario makes the downtown core reachable more quickly, facilitating onward access by rapid transit and commuter rail to job centers around Back Bay and Somerville. Accessibility to jobs increases from 343,485 in the baseline to 579,021.

Changes in accessibility to jobs due to the scenario were similarly calculated for every point in the entire region, for time cutoffs of 30, 60, and 90 minutes. These results use a $5.00 cost cutoff. As a major component of transit travel time (and thus accessibility) is walking to and from transit stops, metrics should be calculated at a high resolution (Benenson *et al.* 2016, Boarnet *et al.* 2017). We use a regular grid with a resolution of approximately 300m by 300m, for a total of 217,602 origin points for each selected travel time cutoff.

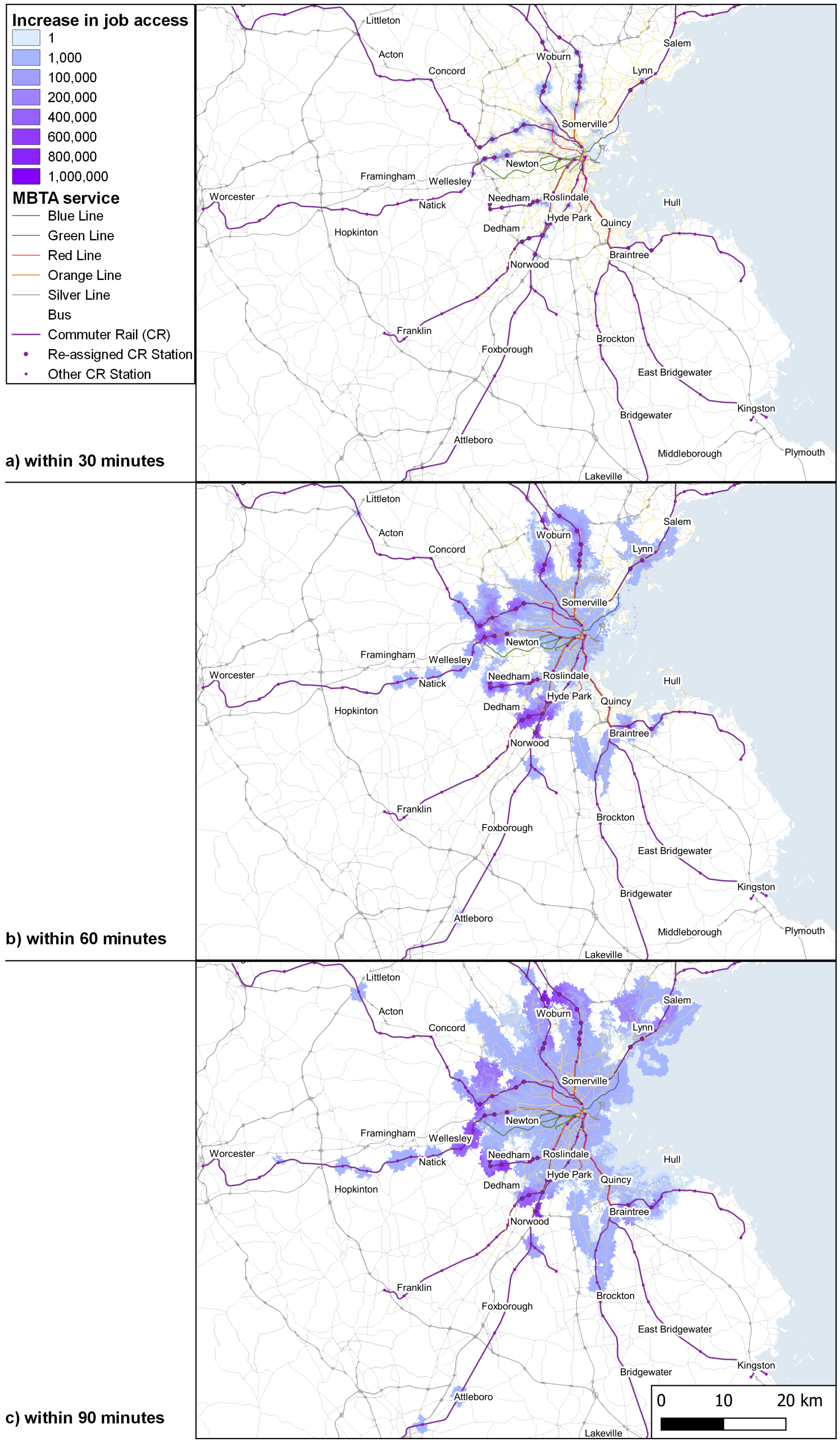

Figure 7: Increase in number of jobs accessible with $5, scenario vs. baseline

With a 30-minute cutoff (Figure 7(a)), relatively small increases in accessibility are concentrated around the re-assigned commuter rail stations. In a few instances, changes in access also extend out along routes that connect to these stations (*e.g.* the Red Line in western Somerville). With a 60-minute cutoff, changes are more widespread and network effects more pronounced (Figure 7(b)). Bus routes north of Lynn, for example, as well as routes south of Hyde Park, provide connectivity to commuter rail, enabling people who reside far from the commuter rail network to benefit from reduced commuter rail fares. With a 90-minute cutoff, the areas with the largest accessibility increases are further out (Figure 7(c)); from inner origins, most of the region's jobs can be accessed in 90 minutes even in the baseline. Small increases in access are widespread with the higher time cutoffs, especially for inner origins, likely reflecting more affordable reverse-commutes to jobs around commuter rail stations that were formerly assigned to Zone 2, and thus inaccessible from inner stations with a $5 cutoff. The sensitivity of these cumulative-opportunity metrics to the selected travel time cutoff could be addressed by instead constructing and using a metric based on a gravity-type or other decay function.

4.4 Algorithm performance

The algorithm is quite performant, with results for single origins available in seconds. Given the analysis settings in Section 4.2, from Copley Square, with a 60 minute time limit and a $5 fare limit, a modern consumer laptop with a 2.80 GHz Intel Core i7-7600U, 4MB CPU cache, and 16GB of RAM can compute the result in 4.5 seconds. During this process, fares are computed for approximately 5.5 million possible journey prefixes. A more complex task (departing from the Downtown Crossing area, with a 120-minute time limit and a $13 fare limit), took 17.3 seconds. Fares for 22.9 million journey prefixes were calculated in the process, or 1.3 million fares per second.

Regional accessibility results in Figures 7 were computed with a 50 machine cluster. We used virtual cloud servers with 4 cores and 30.5 GB RAM each. This cluster computed the baseline accessibility for 217,602 origins, a time cutoff of 60 minutes, and a fare cutoff of $5, in 4 minutes 52 seconds.

This performance approaches that of Conway *et al.*'s (2017) algorithm; their algorithm takes 3 minutes 32 seconds on the same problem without a fare limit. This is partly due to our algorithm using a sample of departure times, while Conway *et al.* exhaustively calculate travel times for each minute within the time window. Additionally, in a large regional accessibility analysis there are many areas that are only served by relatively expensive, long-distance services, which are not even explored when using a low fare limit such as $5.00. Accessibility values for these origins are likely to be computed much more quickly with our algorithm, partially compensating for relatively slower performance in central areas.

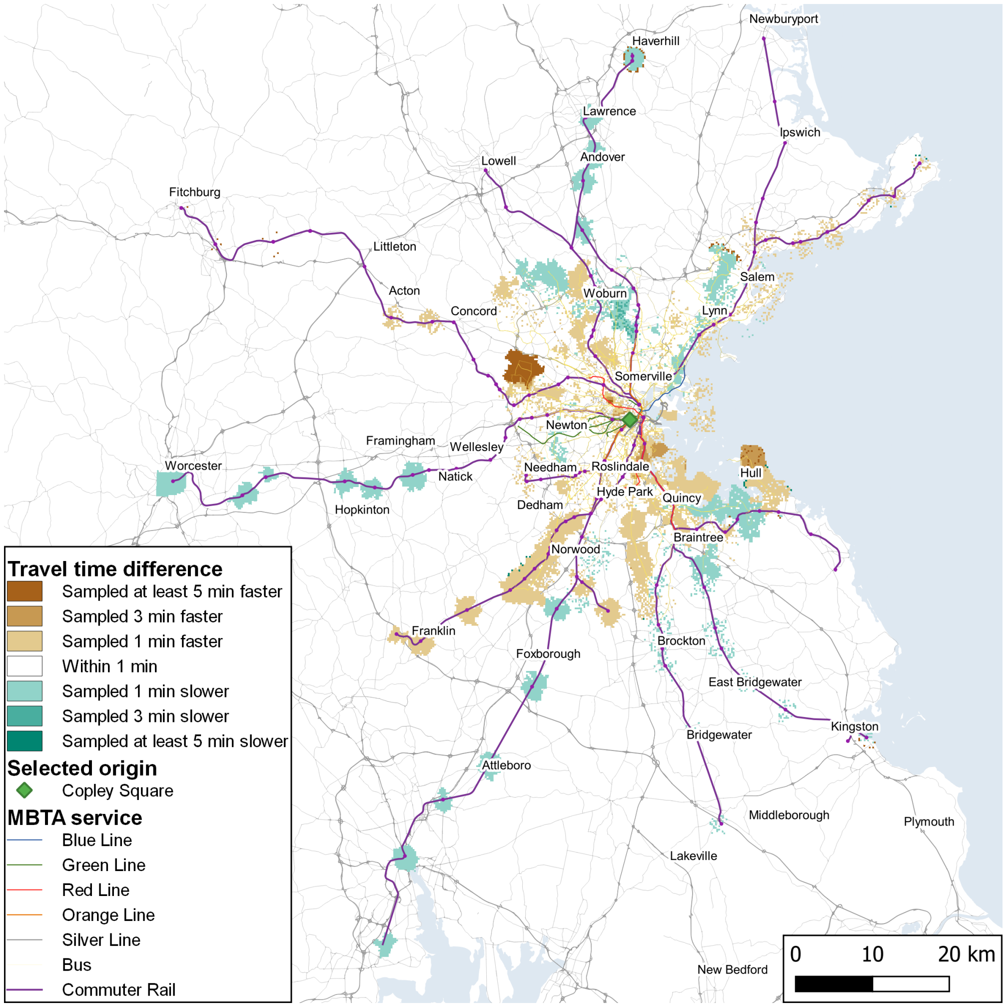

Figure 8: Comparison between results of this algorithm (multiobjective with sampled departure times) using a non-binding fare limit, versus the algorithm of Conway *et al.* (2017)

To assess the noise arising from the departure sampling procedure, we compared the travel times from Copley Square to every other location in Boston area, as computed by our algorithm and by Conway *et al.*'s, using a nonbinding fare cutoff of $1,000 (Figure 8). Reassuringly, the results are similar, with any travel time differences attributable to the random sampling of departure times. As a close look at this figure shows, travel times to locations along some infrequent transit lines increase, while areas along other lines show a decrease—because the randomly-drawn departure times made it more convenient to use some lines, and less convenient to use others. If the noise introduced were unacceptable for a particular application, the sampling frequency could be increased.

5 Further research

The key innovation of Conway *et al.* (2017) was to allow the analysis of sketch plans of transit systems that do not specify exact timetables for all lines, by generating many randomized timetables. While our method supports the use of randomized timetables, it uses much smaller numbers of shortest-path searches and thus randomized timetables than Conway *et al.* (2017). For our case study, where we are not analyzing a change to transit service, this is unproblematic. However, it may be an issue for scenario planning when there are many planned lines with uncertain timetables. An avenue for future research would be to evaluate the extent of this uncertainty by calculating the sampling distribution of our cost-constrained accessibility figures, for instance by using the bootstrapping techniques discussed in Conway *et al.* (2018). Conway *et al.* (2017) also used the rRAPTOR extension described by Delling *et al.* (2015), which allows more efficient queries over a time window; this optimization has not been implemented in our multicriteria router.

One limitation of the algorithm is that it assumes users always board the first vehicle that comes on a particular 'pattern.' In RAPTOR parlance, a pattern is a unique sequence of stops—conceptually similar to a route, except that it only represents a single direction, and routes that have detours or short-turns will be split into multiple patterns (Delling *et al.* 2015). However, in systems with peak and off-peak fares, boarding the first arriving vehicle on a given pattern might not produce the lowest cost trip that meets the time constraints; if a trip is happening near the end of a peak period, it may be advantageous for the user to wait to board until the peak period has expired. A slight modification to the algorithm could account for this. The RAPTOR algorithm considers each pattern as a separate option. If patterns were separated into on-peak and off-peak, the algorithm would thus consider them to be separate options. Trips that overlap both the peak and off-peak period could be assigned to unique patterns with their own fares.

Since each round k of the RAPTOR algorithm finds all (Pareto)-optimal paths with less than $k - 1$ transfers (Delling *et al.* 2015), it would be computationally inexpensive to extend the algorithm proposed herein to produce paths that are Pareto-optimal on travel time, fare, *and* number of transfers. All that would be needed would be to retain and tabulate the optimal paths that have been found after each round. Travelers often perceive transfers to be an additional cost of travel (Guo and Wilson 2011), and the perceived disutility of transfers may vary between individuals. For this reason it could be valuable to produce Pareto-optimal paths with differing number of transfers, to understand how an increase in the disutility of transferring may impact accessibility.

Additionally, the tractability of our algorithm depends on being able to eliminate journey prefixes that will not result in a unique Pareto-optimal combination of time and cost. In theory, a network with a very large number of bundles of transfer allowances could create a tractability issue, by preventing the domination rules presented in Theorem 3.2 from eliminating many journey prefixes which have different transfer allowances to various services. (The domination rule presented in Theorem 3.1 will still eliminate any particularly high cost journey prefixes.) It is not known to what extent fare systems which would cause these tractability issues exist in the real world; further research is needed to evaluate algorithm performance in these types of fare systems, and devise (possibly heuristic) solutions if necessary.

There are a number of relevant policy questions that this framework could shed light on. As new fare media become available, technical restrictions on the ability to provide free or discounted transfers lessen. In Boston, a pilot has been discussed to allow free transfers between some commuter rail lines and rapid transit, which would increase accessibility for budget-constrained travelers, allowing them to transfer from commuter rail to rapid transit and access much of the city without exceeding their budgets (Carvalho 2017). Additionally, many cities have implemented or are considering transit subsidies for low-income residents (*e.g.* Goodman 2018); this algorithm could be used to analyze the accessibility impacts of these programs, by treating them as an increase in the effective cost cutoff for the accessibility metric.

While Pareto-optimal shortest-path searches are not new, we have implemented a Pareto-optimal search process in a user-friendly transit scenario planning application. This opens avenues for computing accessibility metrics using other novel constraints, such as reliability or comfort. These constraints would not require the special computations needed to ensure correctness when using fares as an objective function, but could take advantage of the open-source multiobjective routing software we have created.

5.1 Retaining multiple codominant Pareto-optimal paths

The algorithm as described herein can compute the complete Pareto frontier for time and cost, but it is not guaranteed to retain every journey that can produce any particular point on the Pareto frontier. If two journeys are equivalent in time and fare, at some point the journey prefix of one of them may have been pruned. This is due to the use of non-strict inequalities in (1) and (2).

In theory, replacing the non-strict inequalities in (1) and (2) with strict inequalities would solve this problem. If we were to do this, *every* journey that can yield a Pareto-optimal result would be returned (at least, every journey that involved boarding the first vehicle to arrive on a particular 'pattern,' as described above). However, this would quickly become intractable, due to the combinatoric number of slight variations possible on each journey. For instance, if one option to get from an origin to a destination is to take a bus to another bus, one may be able to choose from a number of board stops, a number of transfer points, and a number of egress stops. Retaining all of these options could clearly create a tractability issue.

However, it is unlikely that retaining all of these options is actually desirable—most users will transfer at the point where two lines pass closest, for example, unless they can reduce their travel time or fare by transferring elsewhere. In order to retain the desirable paths, we propose a combination of additional objective functions and heuristics to improve the quality of the paths returned from our algorithm. Adding walking time as an objective function would ensure that the returned paths for each point on the Pareto frontier would not unnecessarily walk to further away stops or take long on-street transfers, unless it provided a lower fare or earlier arrival. A heuristic could retain both journey prefixes P and Q if they have identical travel times and fares, but use different routes (for instance, multiple feeder bus options to a train).

If the results are going to be used as choice sets in a discrete choice model, the dimensions of the choices that are used as independent variables in the model should be entered as optimization criteria, if possible. If path characteristics that are not optimization criteria are used as choice attributes in the model, there is no guarantee that the utility-maximizing choice will be in the choice set. In contrast, if all choice attributes used in the model are included as optimization criteria, it is guaranteed that the utility-maximizing choice will be returned, assuming coefficient signs are as expected, because the utility-maximizing choice is guaranteed to be a location on the Pareto surface, and our algorithm identifies all feasible locations on the Pareto surface. This neatly sidesteps the issue of not returning all possible paths that can yield a particular point on the Pareto surface; if all the variables of interest are optimization criteria in the algorithm, the choice set of all possibly utility-maximizing points is exactly the set of feasible points on the Pareto surface.

If results are instead intended for a human-facing journey planner, a combination of optimization criteria and heuristics can be used to present a qualitatively satisfactory set of choices to the user interface. This may involve not presenting all of the Pareto-optimal points; some may be considered too extreme—for example, a route that avoids a \$0.25 transfer surcharge by walking a long distance to the destination rather than taking the direct bus. It may involve adding objectives (*i.e.* Pareto dimensions), for example, to eliminate journeys with unnecessary walking. It may also involve heuristically differentiating routes that are identical in terms of the objective functions; for instance, it may be desirable to retain two journeys with the same fare and arrival time, that use different feeder bus routes.

6 Conclusion

In this article, we introduced a method for computing Pareto sets of shortest-path results in public transit systems, optimizing on travel time and fare. The method is general and applicable to a broad range of public transit fare systems. We applied the method to an illustrative fare change scenario in the Greater Boston region. The algorithm was able to compute results in a satisfactory amount of time, even given the region's dense transit network and complex fare system. These results were used to compute cumulative accessibility measures constrained both in time and cost.

Researchers and practitioners are engaged in ongoing debates about equity in transport provision. Fare policy should be a major part of this debate. Our method allows accessibility and other metrics to be calculated in transit systems with complex fare policies, without needing to assume a value of time. Planners can use the results of the method to analyze the combined accessibility impacts of proposed changes to fare policy and/or service, and understand the travel time and travel cost tradeoffs faced by budget-constrained travelers.

Governments often require transit agencies to analyze the equity impacts of changes to their service and fare policy (*e.g.* Federal Transit Administration 2012). Standard catchment-type analyses of fare policy do not account for network effects, but it is clear from our study that network effects exist; changing the fare on a few rail lines affected not only the accessibility in station areas around those lines, but also the accessibility in areas around lines that connect to those rail lines. Our method considers these important network effects.

Our methodological innovation is the ability to compute Pareto sets of transit journeys, optimizing on both time and fare. Our algorithm is able to find the lowest cost path that meets a travel time constraint, even if there is a faster, more expensive option. Our algorithm can exactly handle complex, non-additive fare systems with varying transfer privileges between routes, a goal that has previously been elusive. We applied the algorithm to an accessibility analysis, in which we demonstrated it to be highly performant. Similar algorithms could be used in customer-facing journey planners, or to develop choice sets for route-choice models, by slightly adjusting the objective functions or adding heuristics to the search process.

7 Acknowledgements

Road basemap © OpenStreetMap contributors.

References

Anderson, P., Levinson, D., and Parthasarathi, P., 2013. Accessibility futures. *Transactions in GIS*, 75 (2), 683–705. Available from: http://doi.wiley.com/10.1111/tgis.12024.

Antsfeld, L. and Walsh, T., 2012. Finding multi-criteria optimal paths in multi-modal public transportation networks using the transit algorithm. *In*: *Proceedings of the 19th ITS World Congress*, Vienna. Available from: https://pdfs.semanticscholar.org/8158/276fdc49b27293187e54fc7a835e7e62abee.pdf [Accessed 2018-09-21].

Benenson, I., *et al.*, 2016. The benefits of a high-resolution analysis of transit accessibility. *International Journal of Geographical Information Science*, 31 (2), 213–236. Available from: http://dx.doi.org/10.1080/13658816.2016.1191637.

Bertaud, A., 2018. *Order without design: how markets shape cities.* Cambridge, MA: MIT Press.

Boarnet, M.G., *et al.*, 2017. First/last mile transit access as an equity planning issue. *Transportation Research Part A*, 103, 296–310.

Börjesson, M. and Eliasson, J., 2019. Should values of time be differentiated? *Transport Reviews*, 39 (3), 357–375. Available from: https://www.tandfonline.com/doi/full/10.1080/01441647.2018.1480543.

Carvalho, E.C., 2017. H.2623: An act establishing a rapid transportation pilot for the Fairmount corridor. Available from: https://malegislature.gov/Bills/190/H2723 [Accessed 12 March 2019].

Conway, M.W., Byrd, A., and van der Linden, M., 2017. Evidence-based transit and land use sketch planning using interactive accessibility methods on combined schedule and headway-based networks. *Transportation Research Record*, 2653, 45–53.

Conway, M.W., Byrd, A., and van Eggermond, M., 2018. Accounting for uncertainty and variation in accessibility metrics for public transport sketch planning. *Journal of Transport and Land Use*, 11 (1), 541–558.

Cui, M. and Levinson, D., 2018. Full cost accessibility. *Journal of Transport and Land Use*, 11 (1), 661–679. Available from: https://www.jtlu.org/index.php/jtlu/article/view/1042.

Delling, D., Pajor, T., and Werneck, R.F., 2015. Round-based public transit routing. *Transportation Science*, 49 (3), 591–604.

Dreier, P. and Vrabel, J., 2010. Did he ever return?: The forgotten story of Charlie and the M.T.A. *American Music*, 28 (1), 3–42.

El-Geneidy, A., *et al.*, 2016. The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transportation Research Part A: Policy and Practice*, 91, 302–316.

Enchautegui, M.E., 2013. *Nonstandard work schedules and the well-being of low-income families*. Washington, DC. Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.375.8404 [Accessed 31 July 2018].

Federal Transit Administration, 2012. *FTA circular 4702.1B: Title 6 requirements and guidelines for Federal Transit Administration recipients*. Available from: https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/FTA_Title_VI_FINAL.pdf [Accessed 25 July 2018].

Foth, N., Manaugh, K., and El-Geneidy, A.M., 2013. Towards equitable transit: examining transit accessibility and social need in Toronto, Canada, 1996-2006. *Journal of Transport Geography*, 29 (C), 1–10.

Geurs, K.T. and van Wee, B., 2004. Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography*, 12 (2), 127–140.

Goodman, J.D., 2018. Leaders poised to provide discounted MetroCards to low-income New Yorkers. *The New York Times*. Available from: https://www.nytimes.com/2018/06/07/nyregion/nyc-subway-metrocard-fair-fares-subsidize.html.

Grengs, J., 2012. Equity and the social distribution of job accessibility in Detroit. *Environment and Planning B: Planning and Design*, 39 (5), 785–800. Available from: http://epb.sagepub.com/lookup/doi/10.1068/b36097.

Guo, Z. and Wilson, N.H.M., 2011. Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. *Transportation Research Part A*, 45 (2), 91–104. Available from: http://dx.doi.org/10.1016/j.tra.2010.11.002.

Hess, S., *et al.*, 2017. A framework for capturing heterogeneity, heteroskedasticity, non-linearity, reference dependence and design artefacts in value of time research. *Transportation Research Part B*, 96, 126–149.

Karner, A., 2018. Assessing public transit service equity using route-level accessibility measures and public data. *Journal of Transport Geography*, 67, 24–32.

Kujala, R., *et al.*, 2018. Travel times and transfers in public transport: Comprehensive accessibility analysis based on Pareto-optimal journeys. *Computers, Environment and Urban Systems*, 67 (C), 41–54. Available from: https://doi.org/10.1016/j.compenvurbsys.2017.08.012.

Lo, H.K., Yip, C.W., and Wan, K.H., 2003. Modeling transfer and non-linear fare structure in multi-modal network. *Transportation Research Part B: Methodological*, 37 (2), 149–170.

Mahmoudi, M., *et al.*, 2019. Accessibility with time and resource constraints: Computing hyper-prisms for sustainable transportation planning. *Computers, Environment and Urban Systems*, 73, 171–183. Available from: https://doi.org/10.1016/j.compenvurbsys.2018.10.002.

Manaugh, K. and El-Geneidy, A.M., 2012. Who benefits from new transportation infrastructure? Using accessibility measures to evaluate social equity in public transport provision. *In*: K. Geurs, K. Krizek and A. Reggiani, eds. *Accessibility and transport planning challenges for Europe and North America*. London: Edward Elgar Publishing, 211–227. Available from: http://tram.mcgill.ca/Research/Publications/Transit_plan_equity.pdf [Accessed 31 July 2018].

Martens, K., 2017. *Transport justice*. New York: Routledge.

Martens, K. and Golub, A., 2018. A fair distribution of accessibility: Interpreting civil rights regulations for regional transportation plans. *Journal of Planning Education and Research*, 78 (1), 0739456X1879101–20.

Massachusetts Bay Transportation Authority, n.d.a. Automated fare collection 2.0. Available from: http://web.archive.org/web/20180801213143/https://www.mbta.com/projects/automated-fare-collection-20-afc-20 [Accessed 1 Aug 2018].

Massachusetts Bay Transportation Authority, n.d.b. Fares. Available from: http://web.archive.org/web/20180802013933/https://mbta.com/fares [Accessed 1 Aug 2018].

Miller, E.J., 2018. Accessibility: measurement and application in transportation planning. *Transport Reviews*, 38 (5), 551–555.

Miller, H.J., 1991. Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems*, 5 (3), 287–301. Available from: http://www.tandfonline.com/doi/abs/10.1080/02693799108927856.

Müller-Hannemann, M., *et al.*, 2007. Timetable information: Models and algorithms. *In*: *Algorithmic methods for railway optimization*. Berlin, Heidelberg: Springer, Berlin, Heidelberg, 67–90.

O'Sullivan, D., Morrison, A., and Shearer, J., 2000. Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach. *International Journal of Geographical Information Science*, 14 (1), 85–104.

Owen, A. and Jiang, H., 2015. Temporal sampling intervals and service frequency harmonics in transit accessibility evaluation. Unpublished manuscript.

Pereira, R.H.M., Schwanen, T., and Banister, D., 2017. Distributive justice and equity in transportation. *Transport Reviews*, 37 (2), 170–191.

Pucher, J. and Renne, J.L., 2003. Socioeconomics of urban travel: evidence from the 2001 NHTS. *Transportation Quarterly*, 57 (3), 49–77.

Rodriguez, C., *et al.*, 2017. Accessibility, affordability, and addressing informal services in bus reform. *Transportation Research Record*, 2634, 35–42.

San Francisco Bay Area Rapid Transit District, n.d. Fare calculator. Available from: http://www.bart.gov/tickets/calculator.

Stewart, A.F., 2017. Mapping transit accessibility: Possibilities for public participation. *Transportation Research Part A*, 104, 150–166.

Stuntz, A., Atanucci, J., and Salvucci, F.P., 2017. Process for transit fare structure and fare level analysis: Case study at Massachusetts Bay Transportation Authority. *In*: *Transportation Research Board Annual Meeting*, Washington, DC. 1–16.

Vaccaro, A., 2018. City councilor wants lower prices on commuter rail in Boston neighborhoods. *The Boston Globe*. Available from: https://www.bostonglobe.com/metro/2018/02/09/city-councilor-wants-lower-prices-commuter-rail-boston-neighborhoods/Z5LUfD7vHbqzgtXFuu6CcL/story.html.

Verbich, D. and El-Geneidy, A.M., 2017. Public transit fare structure and social vulnerability in Montreal, Canada. *Transportation Research Part A*, 96, 43–53.

Xie, C. and Waller, S.T., 2012. Parametric search and problem decomposition for approximating Pareto-optimal paths. *Transportation Research Part B*, 46 (8), 1043–1067.