Time-Dependent Capacitated Transit Routing with Real-Time Demand and Supply Data

Ömer Verbas · Vadim Sokolov · Joshua Auld · Hubert Ley

Abstract This paper proposes an online A* transit routing algorithm that incorporates real-time information both on the supply side in terms of travel times and on the demand side in terms of transit vehicle loads.

Keywords: Real-time · Routing · Shortest path · A* · Bayesian statistics · Weibull

1 Introduction

Many transit agencies in the world face the recurring problem of severe weather events which cause disruptions and emergency situations at various levels including schedule delays, deviations from the planned route, route shut-downs, evacuation of exposed populations etc. Although rerouting of buses and trains real-time is a very complex and important problem, another challenge is providing the travellers with fast and reliable routing information under different levels of disruption. This paper proposes an online intermodal routing algorithm that incorporates real-time information both on the supply side in terms of travel times and on the demand side in terms of transit vehicle loads.
Shortest path algorithms have evolved in many different directions since the introduction of the label setting (Dijkstra 1959) and label correcting (Ford Jr 1956; Bellman 1958) algorithms in 1950’s. The first time-dependent algorithm was introduced by Cooke and Halsey in (1966), whereas the first point-to-point (A*) algorithm was introduced by Hart et al. in (1968).

The presented algorithm here extends on the previous work by Khani et al. (2015) and Verbas et al. (2018). This algorithm differs from Khani et al. (2015)’s as follows: the walking links are derived from the street network as opposed to limiting the walking links to designated transfer points; the presented algorithm does not follow a trip’s further downstream stops; and the labelling is performed on the links as opposed to the nodes to account for node switching penalties in an easier fashion. Moreover, the proposed algorithm extends on Verbas et al. (2018) by introducing real-time updates on the supply side and on the demand side. The demand-side information allows for incorporating capacity penalties into the algorithm.

2 Problem description and model formulation

2.1 Problem description

Our focus is the development of an algorithm that provides a real-time online least cost transit path to travellers that updates both link travel times and transit vehicle loads at an acceptable frequency. First, the travel times and vehicle loads have to be calculated using both real-time and historical data. Moreover, the least cost algorithm must be fast enough to provide the traveller with a new path in an acceptable time window and has to reflect traveller heterogeneity.

2.2 Model formulation

The objective is to find the least cost path from a given origin link \( r \in A \) to the destination link \( s \in A \), where \( A \) is the set of all links in the network. Along the least cost path the following is true for every link:

\[ F_j^* = G_j^* + H_j \]  

(1)

\( G_j^* \) is the cost of the least cost path from the origin link \( r \in A \) to the end of the given link \( j \in A \). \( H_j \) is an admissible estimated (heuristic) cost from the end of the given link \( j \in A \) to the destination link \( s \in A \). The following is true for \( G_j^* \) if \( j \) is a transit link:

\[ G_j^* = G_i^* + w_\omega \omega_{jq}^* + w_\mu \mu_{jq}^* + w_\chi \chi_{jq}^* + w_C C(m_{jq}) \]  

(2)

\( G_i^* \) is the cost at the end of the optimal predecessor link. \( \omega_{jq}^*, \mu_{jq}^*, \chi_{jq}^* \), and \( C(m_{jq}) \) are the waiting time (can be zero if not boarding), the in-vehicle travel time, the transfer penalty, and the load penalty respectively. \( q^* \) is the optimal transit trip to be taken that passes through link \( j \) at/after the arrival time at the end of \( j \). The transfer penalty and load penalty calculations are omitted in this abstract. Interested readers are referred to Verbas and Mahmassani (2015) and Verbas (2014). \( m_{jq}^* \) is the load on trip \( q^* \) at the time it arrives at the beginning of link \( j \). \( w_\omega, w_\mu, w_\chi, w_C \) are weights associated with each cost component to reflect traveller heterogeneity.

The following is true if \( j \) is a walking link, where \( \tau_j \) is the walking time and \( w_\kappa \) is the weight that reflects the walk propensity of a traveller:
\[ G_j^* = G_j + w_n \tau_j \] (3)

2.2.1 Estimation of link travel times and waiting times

On the supply side, the link travel times \( \mu_{jq} \) are updated using real-time information mixed with historical data using the following Weibull survival function (Klein and Moeschberger 2005), where \( \Delta t \) is the difference between the look-up time and current time, \( \alpha \) is the scale parameter, \( \beta \) is the shape parameter, and \( \theta \) is the selected weight between 0 and 1:

\[
\Theta = \exp \left[ - \left( \frac{\Delta t}{\alpha} \right)^\beta \right] \theta
\] (4)

The link travel time \( \mu_{jq} \) is then a weighted mix of the prevailing travel time \( \mu_{jq}^1 \) and the historical travel time \( \mu_{jq}^2 \):

\[
\mu_{jq} = \Theta \mu_{jq}^1 + (1 - \Theta) \mu_{jq}^2
\] (5)

2.2.2 Estimation of demand and vehicle loads

Given a model of a transit network \( \{V, C, D\} \), where the vector \( V = v_r; r = 1, \ldots, n \) is the set of origin and destination locations. Spatial resolution of locations depends on the problem and can range from individual transit stations and stops to traffic analysis zones. \( C = c_{rs}; r, s = 1, \ldots, n \) is the accessibility matrix, with each entry being the travel cost from location \( r \) to location \( s \) and infinity if there is no transit connection between locations. Further, \( D = d_{rs}; r, s = 1, \ldots, n \) is the demand matrix, with each entry being the number of trips from location \( r \) to location \( s \).

The problem of real-time monitoring it to dynamically estimate \( C_t \) and \( D_t \) at time \( t \) by assimilating data from automated vehicle location (AVL) data and transactions data generated by the fair card users. AVL data contains the position of the vehicles and thus allows to estimate travel times in real time as well as the automated passenger counts that allow to update the assumptions about the current demand matrix. In other words we would like to calculate new demand matrix \( D_t \) that is consistent with currently observed boardings and alighting as measured from automated passenger counts (APC) and transactions data.

The traditional approach to the problem is to assume that the measured data contains no errors and use it in deterministic way. Given boarding data \( b_r, r = 1, \ldots, n \) and alighting data \( a_s, i = 1, \ldots, n \) which is some weighted average of APC data and transactions data we would like to find a demand matrix which is a close as possible to historically observe demand (as estimated by a travel demand model) \( D_H \) and satisfies the boarding/alighting constraints. The resulting problem is as follows:

\[
\min_{D} ||D - D_H||_F^2, \quad D \in R^{n \times n}
\] (6)

s. t \[ \sum_{r=1}^{n} d_{rs} = a_s, \quad \forall s = 1, \ldots, n \] (7)

\[ \sum_{s=1}^{n} d_{rs} = b_r, \quad \forall r = 1, \ldots, n \] (8)
This problem formulation has a few drawbacks. First, it assumes no measurement errors in the right hand side of the constraints $a$ and $b$. Second, it simply tries to find the closest matrix to demand matrix and implicitly assumes that the travel costs $C$ hasn’t changed. This assumption might be valid when a small deviation from historical demands is observed. However, during a hazard event when large portion of demand is induced or as a result of major change in schedules/routes the assumption of constant $G$ and small change in $D$ won’t hold.

Another way to model the problem is via a hierarchical statistical model. Instead of assuming a deterministic relation between current demand $d$, historic travel patterns $d_H$ and induced demand $d_I$ we model the relation probabilistically using a hierarchical model. The dependency structure is shown below in Fig. 1:

![Dependency structure](image)

**Fig. 1** A Bayesian estimation method for demand

The lowest layer represents the variable that we observe (alighting, boarding and card data), the middle layer is the variable to be estimated and the top layer are the variable that are not observed, but known from the historic data with some uncertainty.

The problem at hand is to estimate $p(d \mid a, b, c)$. Bayes (1970) rule allows to efficiently estimate such a distribution:

$$p(d \mid a, b, c) = \frac{p(a, b, c \mid d) p(d)}{p(a) p(b) p(c)}$$

(9)

### 2.2.3 Real-time TDIMA* algorithm

See **Fig. 2** for the pseudo-code of the real-time TDIMA* algorithm.
Fig. 2 The real-time TDIMA* algorithm

3 Experiments and results

3.1 Problem case study

The problem will be applied to the Chicago metropolitan network with the properties below. See Fig. 3 for further details:

- 54,028 nodes
  - 35,077 transit
  - 18,951 walking
- 160,642 links
  - 37,642 transit
  - 123,000 walking
- 173,236 activity locations (serving as origin and destination points)
- 1,961 traffic analysis zones
- 344 transit routes (three agencies)
- 2,098 transit patterns
- 28,138 transit trips
Fig. 3 Chicago metropolitan network

3.2 Preliminary Results

The offline TDIMA* algorithm has been validated in Verbas et al. (2018) and a single query is returned in less than 10 milliseconds on average. The Bayesian estimation method for demand is under development and preliminary work has been presented by Le et al. (2017).

4 Conclusion and future work

This paper presented the framework of an online A* transit routing algorithm that incorporates real-time information both on the supply side in terms of travel times and on the demand side in terms of transit vehicle loads. The offline TDIMA* algorithm will be interacted with the Bayesian estimation method in order to finalize the development of the online algorithm.

Acknowledgements: Argonne National Laboratory’s work was supported by the U.S. Department of Transportation, Federal transit Administration under interagency agreement, through U.S. Department of Energy contract DE-AC02-06CH121357. The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (“Argonne”). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up, nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.

References


Ford Jr LR (1956) Network flow theory. DTIC Document


