A MARKOVIAN DYNAMIC TRANSIT ASSIGNMENT APPROACH FOR LARGE-SCALE MULTIMODAL NETWORKS

Ricardo de la Paz Guala, Transport Division, Civil Eng. Dept., FCFM, U. de Chile – ricardo.delapaz@uchile.cl
Marcela Munizaga, Transport Division, Civil Eng. Dept., FCFM, U. de Chile – mamuniza@ing.uchile.cl
Jaqueline Arriagada, Doctorate in Engineering Systems, FCFM, U. de Chile – afjg.7440@gmail.com

Keywords: Dynamic transit assignment , Transit networks , Uncertainty , Smartcard data

ABSTRACT

We develop the Markovian dynamic transit assignment (MDTrA) modelling framework for large multi-modal networks to incorporate into the transit assignment analysis the dynamic aspects that come from the time-dependent relationship between demand and supply and the stochasticity that comes from differences in passenger's costs perception, measurement errors, and other sources of uncertainty. The intuition is that passengers choose their routes by a recursive arc-choice process, according to the expected minimum costs from their current node to their destinations. Our approach presents an important oportunity to use smartcard data, as demand profiles and users's choices over time are obtained from it.

1. INTRODUCTION

Transit assignment models play a fundamental role in the planning process of public transport systems. They allow transport planners and transport researchers to estimate the flow of passengers for each alternative route to reach a certain destination from a given origin. Developing a transit assignment model is a complex task since it involves emulating the passengers' route decision making process and identifying the attributes of different routes. This challenge is even bigger in a dense multimodal transit network, which usually contains hundreds of routes. Currently, the available data from different types of sources (such as smartcards), the inherent uncertainty in users' choices, and the natural time-dependence of demand and supply of transit transport networks, pushes researchers to deal with dynamic and stochastic dimensions if they expect to contribute with approaches that are able to interpretate our current transport context.

Studies about transit assignment are mostly based on *deterministic passenger equilibrium* (*UE*) (Nguyen & Pallottino, 1988; Spiess & Florian, 1989; de Cea & Fernández, 1993; Cominetti & Correa, 2001; Cepeda et al., 2006), which assumes that passengers have perfect knowledge about route costs. Therefore, algorithms find the shortest hyperpath (Nguyen & Pallottino, 1988) and all demand is assigned among these routes, assuming that the passenger takes the first line that arrives at the stop (Spiess & Florian, 1989; Nguyen & Pallottino, 1988). There is also literature based on *stochastic passenger equilibrium* (*SUE*) (Lam et al., 1999; Kiencke & Nielsen, 2000; Nielsen & Frederiksen, 2006; Yang & Lam, 2006), which assumes that passengers have imperfect information about route costs and, therefore, algorithms use *random utility maximization* (*RUM*) models for assigning demand on each route, assuming that passengers choose routes with higher perceived utility.

In assignment literature, we find the Markovian modelling framework approach to address uncertainty in users' choices. Its main motivation comes from the intuition that users of transport networks construct the route they travel through to go from their origins to their destinations by making recursive arc choices, rather than choose a fixed route. It is formally introduced by Baillon & Cominetti (2008), by presenting the concept of *Markovian traffic equilibrium* (*MTE*) for the static case in private transport networks. The concept is later applied to generate the first transit assigment model with a Markovian approach (*STE*) in Cortés et al. (2013), a work that is later extended in Pineda et al. (2016), by integrating it with private transportation (*STP*). Later on, in de la Paz Guala (2020), the Markovian approach is extended, again for the case of private transport nerworks, to address the dynamic dimension that comes from the time dependence of the demand. All of these contributions have in common how the arc-choice model in their respective assignment models is built, as they all consider as the users' choice criterion the expected minimum costs from their current node to their destinations, generating nested cost operators that are able to avoid path enumeration and allow working with overlapping routes.

When it comes to the use of real data, model calibration and validation process is fundamental to set the model's parameters and to evaluate if it is generating passenger flows according to real passengers' behaviour. Although emerging technologies, such as *automatic fare collection* (*AFC*) and *automatic vehicle location* (*AVL*) systems, offer an opportunity to deal with these problems, there are relatively few studies for calibration and validation purposes. In this line, Tavassoli et al. (2020) proposed a framework to calibrate and validate existing transit assignment models using smartcard transaction data and AVL data from public transport systems.

Motivated by the contributions in transit assignment models, the referred Markovian approaches, and the use of smartcard data, the main goal of this paper is to propose a Markovian modelling framework to address the dynamic aspects that come from time-dependent demand and supply and the uncertainty that comes from users's choices, where smartcard data is used to estimate fundamental information for the model. To accomplish this, we present: the *Markovian dynamic transit assignment* (*MDTrA*) model, an illustrative example of how it works, an example of how to use real smartcard data to estimate dynamic demand profiles, and a solution method for the proposed model.

2. MARKOVIAN DYNAMIC TRANSIT ASSIGNMENT MODEL

The core contribution of our paper is the Markovian dynamic transit assignment (MDTrA) model, an integration of the MDTA modelling framewok by de la Paz Guala (2020), conceived as a DTA approach of the MTE (Baillon & Cominetti, 2008), both for private transportation, the transit networks' hypergraph approach by Nguyen & Pallottino (1988), first presented as a search of optimal strategies in Spiess & Florian (1989), and the Markovian approach in public transport in Cortés et al. (2013) and Pineda et al. (2016). The intuition is the following: given exogonous time-dependent demand and supply profiles, passengers, at each stop/station they are currently at, will choose the next convenient arc to move forward to their destinations according to the expected minimum costs from their current node to their destination.

The MDTrA model has three parts: *demand and supply profiles*, *cost and time model*; and an *arc-choice model*. Before developing each part, we first address how to construct the digraph that the model is formulated on and then we introduce the *reasonability* concept.

2.1. Transit network's underlying digraph

To formulate the MDTrA model, we first generate a digraph from the original transit network's hypergraph. This digraph works as the base structure for the formulation.

Considering a directed hypergraph $H(N_H, E_H)$, with N_H and E_H the set of nodes and links, respectively, we recall that, given a pair of nodes, there can be several links conecting them directly, from where links have the form (i, j, s), meaning that line s serves (i, j) by taking passengers directly from node i to node j. Now, given a node $i \in N_H$, we denote: the sets of lines that serve from/to node i (lines that can be boarded from/alighted to i) as $S_i^+ = \{s : (i, j, s) \in E_H\}$ and $S_i^- = \{s : (j, i, s) \in E_H\}$, respectively; and the set of links that serve from/to i as $E_i^+ = \left\{(i, j, s) \in E_H : j \in N \wedge s \in S_i^+\right\}$ and $E_i^- = \left\{(j, i, s) \in E_H : j \in N \wedge s \in S_i^-\right\}$, respectively.

A defining feature of our approach is that it takes into account that passengers could choose to walk from one stop/station to another. A way to decide if walking is an option from i is to set a consideration ratio around it and set that nodes that are within this ratio are walkable from i.

Considering this, we construct a digraph $G = (N, A)$ by the following criteria:

- **Set of nodes**: N is the union of: the set of *original nodes* from the original hypergraph (stops/stations), N_H; the set of *replicated nodes* N_S, that contains, for each node $i \in N_H$ and for each line $s \in S_i^+ \bigcup S_i^-$, a node i_s that represents that line s serves from/to i;
- **Set of arcs**: A is the union of: the set of *boarding arcs* A_b, that contains, for each $i \in N_H$ and for each one of its replicated nodes i_s, the arc (i, i_s) that represents boarding line s at i; the set of *alighting arcs* A_a, that contains, for each node $i \in N_H$ and for each one of its i_s, the arc (i_s, i) that represents alighting from line s at i; the set of *in-vehicle arcs* A_V, that contains, for each $(i, j, s) \in E_H$, the arc (i_s, j_s), that represents moving from i to j using line s; the set of *walking arcs*, A_W, that contains, for each j walkable from i, the arc (i, j).

To show the results of applying these steps, let us consider the hypergraph H in Figure 1 and its resulting digraph G in Figure 2 (node 4 is walkable from node 3).

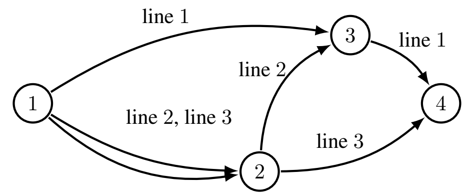

Figure 1: Original Hypergraph H.

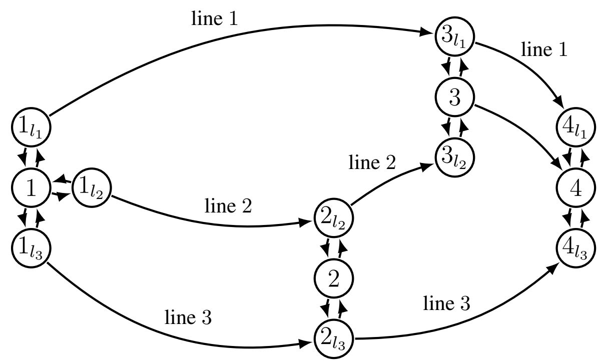

Figure 2: Underlying digraph G to the hypergraph H.

2.2. Reasonable arcs towards destinations

When passengers have to choose between options to travel, from a pure stochastic point of view, each options has a positive probability of being chosen. In reality, this does not generally hold, as some options may not be eligible for passengers, depending on their criteria. Following this intuition, Dial (1971) defined, in the context of static stochastic traffic assignment in private transport networks, that, given an O-D pair (o, d), a route from node i to node j is a reasonable route for (o, d) if the minimum cost of going from o to i is less than the minimum cost of going from o to j and, simultaneously, the minimum cost of going from j to d is less than the minimum cost of going from i to d. In de la Paz Guala (2020) this concept is adapted, defining it over arcs (instead of routes) and according to destinations (instead of O-D pairs), stating that, given a destination node d, an arc (i, j) is a reasonable arc towards d if the minimum cost of going from node j to d is less or equal to the minimum cost of going from node i to d.

On other line of work, Nuzzolo & Comi (2017) introduce the concept of *master hyperpaths*, where, among all the strategies represented in the transport network's hypergraph, the master hyperpath is formed by those that are actually considered by passengers.

In this paper, we introduce a *reasonability* concept, as in de la Paz Guala (2020), to reduce the route options to those that are actually considered by passengers, as in Nuzzolo & Comi (2017). Given a destination node d, a line service s, and two consecutive stops i and j (served by s), we say that (i_s, j_s) is a *reasonable arc towards* d if the minimum cost of going from j to d is less or equal to the minimum cost of going from i to d. Also, the boarding and alighting arcs associated to a reasonable arc are also reasonable (and vice versa). The set of all arcs reasonable towards a destination d is denoted as R^d. Later, in subsection 2.3.2, analytical cost formulations are further developed. Even though different ways of defining convenient options for passengers can be adopted, such as lines of frequent use, they will be addressed in later stages of our work.

2.3. Building the MDTrA model

Now, according to the transit network's digraph, we define the three main parts of the model, considering a period of time given by the interval $[0, T]$.

2.3.1 Demand and supply profiles

Let us consider a set of O-D pairs OD. The demand profile is represented by the exgoneous time-dependent functions $\mathcal{D}_{(o,d)}(\cdot) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$, for all $(o, d) \in OD$. Here, given $t \geq 0$, $\mathcal{D}_{(o,d)}(t)$ represents the demand rate that, at time t, is generated at node o and goes to destination node d. In section 3, we address how to estimate a demand profile based on real smartcard data.

The supply profile is understood as the exgoneous time-dependent nominal frecquency of the different lines, either bus or metro, that serve transit network. Given a node $i_s \in N_S$ and its underlying line $s \in S_i$, the time dependent function $\phi_{i_s}(\cdot) : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ represent the time-dependent nominal frequency of line s at node i.

2.3.2 Time and cost models

This part of the model defines the time and cost structure of the different types of arcs of the transit network's digraph. Time functions are fundamental, as they allow us to locate passengers in time, thus, representing the dynamic aspect of the problem. Also, they are an important component of the cost functions used to compute the expected minimum costs, the core of the choice criterion. It is worth remarking that, given our arc-based approach, times and costs are defined for each arc and, thus, we do not require them to be defined according to routes. This allows us to work with overlapping routes, not needing the condition of route's times and costs independence, as it is widely needed in the literature related to transit assignment.

For each arc $a \in A$, depending on its type, we define the time function $T_a(t)$ and, according to each destination node $d \in D$, a cost function $C_{ad}(t)$, $t \in [0, T]$. Both functions are heavily related as, at each arc, part of its cost is defined by its underlying times. Then, for $t \in [0, T]$, we have:

- For boarding arcs $a = (i, i_s) \in A_b$: $T_a(t)$ is given by the time that takes the passenger to access station/stop i and reach the physical point where line s arrives, b_a, plus the average time that he/she waits for line s, $1/\phi_a(b_a + t)$ ($\phi_a(b_a + t)$ is the mean frequency of service s at node s at time $b_a + t$). Now, independently of the destination $d \in D$, $C_{ad}(t)$ is the direct transformation of times into costs (using the factors α and β).

- For alighting arcs $a = (i_s, i) \in A_a$: $T_a(t)$ is given by the time that takes to alight service s and exit station/stop i, l_a. Now, the first component of $C_{ad}(t)$ is given by transfroming l_a into cost (using a factor δ). Then, given a destination node $d \in D$, the second component of $C_{ad}(t)$: if $i \neq d$, is a transfer penalty p_a for alighting at a node that is not the destination; or if $i = d$, is 0, as destination is reached and the trip ends (no transfer penalty is added).
- For in-vehicle arcs $a = (i_s, j_s) \in A_V$: $T_a(t)$ is given by the time that takes to travel from i to j using service s, V_a. Now, given a destination node $d \in D$, $C_{ad}(t)$ is the transformation of V_a into cost (using the factor γ), plus a cost associated to overcrowding, that depends proportionally (with a factor H_a) on the total inflow assigned to a at instant t.
- For walking arcs $a = (i, j) \in A_w$: $T_a(t)$ is the time w_a and, regardless the destination $d \in D$, $C_{ad}(t)$ is the travel time transformed into cost (using the factor ρ).

Then, for each arc $a = (i, j) \in A$ and at a time $t \in [0, T]$, the time that takes to travel through arc a having entered it at t is given by:

$$T_a(t) = \begin{cases} b_a + \frac{1}{\phi_a(b_a+t)}, & \text{if } a \in A_b, \\ l_a, & \text{if } a \in A_a, \\ V_a, & \text{if } a \in A_V, \\ w_a, & \text{if } a \in A_c, \end{cases} \tag{1}$$

while, for each $d \in D$, the cost of travelling through arc a while heading to destination node d, having entered the arc at t, is given by:

$$C_{ad}(t) = \begin{cases} \alpha b_a + \frac{\beta}{\phi_a(b_a+t)}, & \text{if } a \in A_b, \\ \delta l_a, & \text{if } a \in A_a \text{ and } j \neq d, \\ \delta l_a + p_a, & \text{if } a \in A_a \text{ and } j = d, \\ \gamma V_a + H_a \sum_{d' \in D} E_a^{d'}(t), & \text{if } a \in A_V, \\ \rho w_a, & \text{if } a \in A_c. \end{cases} \tag{2}$$

2.3.3 The choice model

This submodel of the MDTrA model captures the Markovian aspect of our approach thus the way we address stochasticity. The stochastic dynamic transit assignment that results from applying this choice model gives our approach the defining property of being arc-based rather than route-based, as the paths that users follow are formed by the recursive choices that they make while moving towards their destinations. This arc-choice model, according to the expected minimum costs to each destination, assigns inflow rate from each node among its outgoing reasonable arcs. Although it may seem straight forward to apply, it actually has delicate considerations when formulating the model as, for example, not all arcs are necessary reasonable and, at a given time, if inflow rate has been assigned to an arc (i, j), it can not be immediatly assigned back through its inverse arc (j, i) (if it exists).

Expected minumum costs are funtamental functions, as they define the criterion of the logit rule that represents the arc-choice process of passengers to move forward to their destinations. Given a destination d, an arc $a = (i, j)$ and a time t, the exptected minimum cost of using a to go to d, having entered at t, denoted as $Z_{ad}(t)$, is given by the cost of a going to d at t, $C_{ad}(t)$, plus the expected minimum cost from node j to destination d (arriving at j at $t + T_a(t)$). The latter is computed considering all the outgoing reasonable arcs to d from node j (except for (j, i), if it exists (j, i), as flow rate must not immediatly return). Given this, the construction of the expected minimum costs is a recursive process where, even though route enumeration is not needed, its second term keeps the information of the exptected minimum cost of what is left of the trip to d.

To formulate the model, we first need to define some sets. Let us consider a node $i \in N$, a destination node $d \in D$ and the set of reasonable arcs towards d, R^d. Let us denote two subsets of arcs in R^d, the ones that are outgoing from i and the ones that are incoming to i, ${R^d}_i^+ = \left\{(i, j) \in A_i^+ : (i, j) \in R^d\right\}$ and ${R^d}_i^- = \left\{(j, i) \in A_i^- : (j, i) \in R^d\right\}$, respectively. Now, we denote the subsets of arcs of ${R^d}_i^+$ and ${R^d}_i^-$ whose inverse is also reasonable arc towards d as ${B^d}_i^+ = \left\{(i, j) \in {R^d}_i^+ : (j, i) \in {R^d}_i^-\right\}$ and ${B^d}_i^- = \left\{(j, i) \in {R^d}_i^- : (i, j) \in {R^d}_i^+\right\}$, respectively.

Then, extending to a dynamic transit context the equations for the MTE (Baillon & Cominetti, 2008), for each destination node $d \in D$, for each arc $a = (i, j) \in A$ and at each time $t \in [0, T]$, the expected minimum cost of using a going to d, entering it at t, is computed as:

$$Z_{ad}(t) = \begin{cases} C_{ad}(t), & \text{if } j = d, \\ C_{ad}(t) - \frac{1}{\theta} \ln \left(\sum_{b \in R_j^d \setminus \{(j,i)\}} \exp\left(-\theta Z_{bd}\left(t + C_{ad}\left(t\right)\right)\right) \right), & \text{otherwise,} \end{cases} \tag{3}$$

where θ is a known dispersion parameter for the logit model.

Now, according to these expected minimum costs, the assignment is performed. Given a destination node d, a node i and an instant t, the outflow rate of all incoming arcs b arriving to i, each one denoted as $G_{bd}(t)$ (equation 6), and the demand rate generated at i at time t, both with destination d, are aggregated as a single flow rate and then assigned as inflow rates among the outgoing reasonable arcs from node i. Again, even though the idea seems simple, the formulation of the assignment has to be carefully constructed, as the already presented considerations still hold.

Given a destination node $d \in D$, for each node $i \in N$ and at each instant $t \in [0, T]$ we have two cases. (1) If $i = d$ (the node the inflow rate is assigned from is already the destination), then the ouflow rates arriving to i have already reached destination, thus, $\forall a \in A_i^+$ we have that the inflow rate of arc a going to destination d at time t is given by:

$$E_{ad}(t) = 0. \tag{4}$$

On the other hand, (2) if $i \neq d$, then the outflow rates going to d arriving to i have to keep traveling through the transit network to reach d. Then, we have that, for all $a = (i, j) \in A_i^+$:

$$E_{ad}(t) = \begin{cases} \dfrac{e^{-\theta Z_{ad}(t)}}{\sum\limits_{b \in R^{d}{}_i^+} e^{-\theta Z_{bd}(t)}} \left(\sum\limits_{b \in A_i^- \setminus B^{d}{}_i^-} G_{bd}(t) + \mathcal{D}_{(i,d)}(t) \right) + \sum\limits_{(i,m) \in B^{d}{}_i^+} \left(\dfrac{e^{-\theta Z_{ad}(t)}}{\sum\limits_{b \in R^{d}{}_i^+ \setminus \{(i,m)\}} e^{-\theta Z_{bd}(t)}} G_{(m,i),d}(t) \right), \\ \dfrac{e^{-\theta Z_{ad}(t)}}{\sum\limits_{b \in R^{d}{}_i^+} e^{-\theta Z_{bd}(t)}} \left(\sum\limits_{b \in A_i^- \setminus B^{d}{}_i^-} G_{bd}(t) + \mathcal{D}_{(i,d)}(t) \right) + \sum\limits_{(i,m) \in B^{d}{}_i^+ \setminus \{a\}} \left(\dfrac{e^{-\theta Z_{ad}(t)}}{\sum\limits_{b \in R^{d}{}_i^+ \setminus \{(i,m)\}} e^{-\theta Z_{bd}(t)}} G_{(m,i),d}(t) \right), \\ 0. \end{cases} \tag{5}$$

if $a \in R^{d}{}_i^+ \setminus B_i^+$, or if $a \in B_i^+$, or otherwise, respectively.

Now, when inflow rate enters an arc $a = (i, j)$ through i, it reaches j and exits as outflow rate after the travel time of the arc, $T_a(t)$. Thus, given a destination $d \in D$, for each arc $a \in A$ and at each instant $t \in [0, T]$, the inflow and the outflow rates of arc a going to d are related as follows:

$$G_{ad}(t + T_a(t)) = E_{ad}(t). \tag{6}$$

3. SMARTCARD DATA USE

In this section, we show how to use smartcard data in our approach, particularly, to obtain a time-dependent demand profile. The analysis for this study uses the Santiago's (Chile) multimodal public transport network, operated by headway scheduling. Data from the fare system is fully integrated, with flat fares between urban buses, metro, and some rail services. In a typical week, 3 million passengers use the system, making 25.5 million trips. The network includes 7 metro lines, more than 300 bidirectional transit lines, and one rail service. The available data includes very detailed demand and supply information from three sources: the Automatic Fare Collection (AFC) database, the Automatic Vehicle Location (AVL) database, and the *General Transit Feed Specification* (*GTFS*) database.

The AFC system is implemented through a smartcard, which is the only accepted payment method, where passengers must validate when boarding a bus or entering a metro station and no alighting validation is required. However, the data is processed to estimate boarding and alighting positions for all validations and trips (stages) associated with an O-D journey using Munizaga & Palma's (2012) methodology. AFC combined with AVL data allows identifying the chosen paths between each O-D pair, and to estimate information about a trip, such as in-vehicle travel time, out-of-vehicle time (waiting, transfer, and walking times), and the number of transfers. Additionally, AVL data from buses, all equipped with GPS devices that record time and location every 30 $[sec]$, allows obtaining the observed frequency of transit lines to estimate time intervals bewteen buses and waiting times. The GTFS data provides the geographic information of stops and lines structures and the scheduled frequency information for all transit lines.

Before estimate a demand profile to give an example of how the available data can be used, we remark that, for strategic analysis, the Santiago's public transport authorities have successfully implemented the static approach for the public transit assignment models, but it does not take into account the within-day dynamic demand that is evidenced in the smartcard data. To show evidence of this temporal dependence and to highlight why it should be treated from a dynamic apporach, we have collected smartcard data to obtain the numer of transactions per minute made at a metro station in a workday, presented in Figure 3. Note that the number of trips per minute is highly variable along the day, even within periods such as morning peak, afternoon peak and off peak. This variation, in addition to the variability of observed frequencies, produces passenger congestion at specific times. It is this type of dynamic dimension that our proposed MDTrA model allows to address, while also addressing uncertainty in passengers choices from a Markovian point of view.

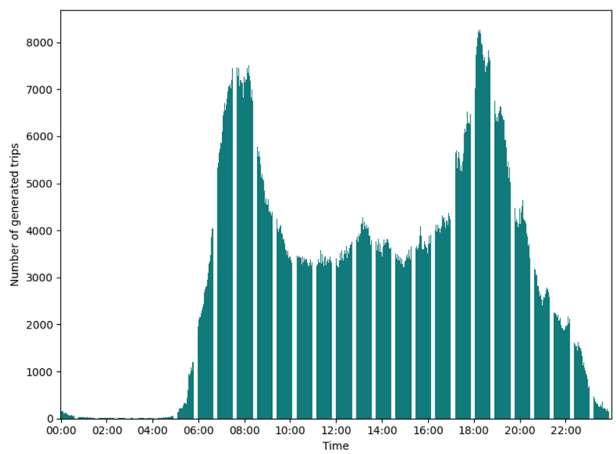

Figure 3: Number of trips generated in each minute of a workday (08 May, 2019).

4. ILLUSTRATIVE EXAMPLE

To show how the MDTrA model works, let us consider digraph G from Figure 2, where nodes represent stations/bus stops and: metro line 1 serves node 1, 3 and 4; metro line 2 serves nodes 1, 2 and 3; and bus line 3 serves nodes 1, 2 and 4. Also, node 4 is walkable from node 3.

The demand rate of passengers with origin at node 1 and destination at node 4, over 200 minutes, is shown in Figure 4. The suppply is defined by the constant mean frequencies of lines 1, 2 and 3, given by $\phi_1(t) = 1\ \frac{metro}{min}$, $\phi_2(t)$ =0.5 $\frac{metro}{min}$, and $\phi_3(t)$ =0.2 $\frac{bus}{min}$, respectively.

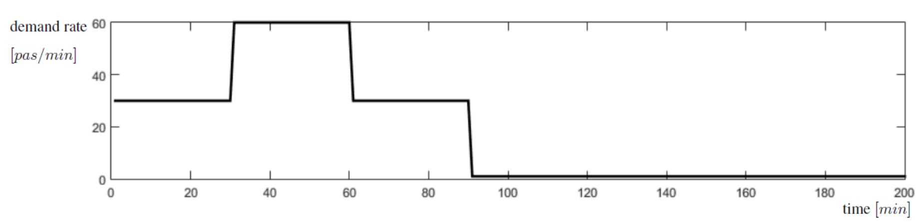

Figure 4: Demand rate from node 1 to node 4.

In this example, crowding does not affect and the cost of an arc a is directly defined by its times, according to equations 1 and 2, thus, the only case in which cost and travel time differ is when a is an alighting arc that does not arrive to node 4. Table 1 shows arc costs associated to node-line interactions: access times; exit times; transfer penalties for not alighting at node 4; and mean waiting times of lines. Table 2 shows costs associated to in-vehicle and walking arcs.

	access	exit	transfer penalty	waiting
node 1 - line 1	3	3	5	1
node 1 - line 2	3	3	5	2
node 1 - line 3	1	1	10	5
node 2 - line 2	2	2	5	2
node 2 - line 3	1	1	10	5
node 3 - line 1	2	2	5	1
node 3 - line 2	2	2	5	2
node 4 - line 1	2	2	0	1
node 4 - line 3	1	1	0	5

Table 1: Cost of node-line interactions $[min]$.

	travel time
node 1 to 2 - line 2	18
node 1 to 2 - line 3	30
node 1 to 3 - line 1	20
node 2 to 3 - line 2	10
node 2 to 4 - line 3	25
node 3 to 4 - line 1	25
node 3 to 4 - walking	35

Table 2: Travel times between nodes $[min]$.

Considering the given information, the constant time and cost $[min]$ of using each arc (going to destination node 4) are shown in Figure 5:

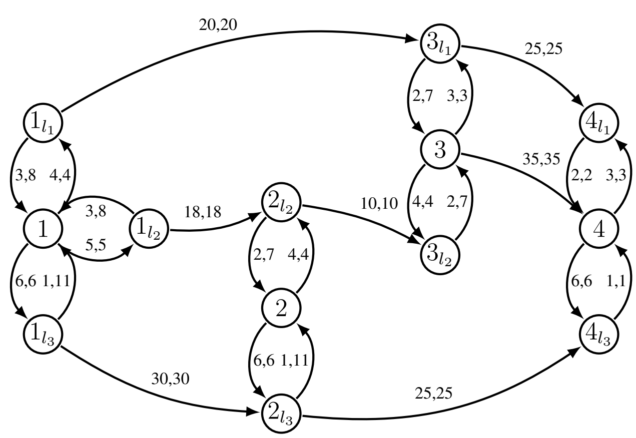

Figure 5: Times and costs of each arc in $[min]$.

Now, as people does not board a line to inmediatly alight it (and vice versa), there are arcs that won't be assigned with flow rate, as choosing them represents this unrealistic action. This arcs are $(1_{l_1}, 1)$, $(1_{l_2}, 1)$, $(1_{l_3}, 1)$, $(3_{l_2}, 3)$, $(4, 4_{l_1})$ and $(4, 4_{l_3})$. This is important to note as the assignment requires to properly identify the arcs that are actual options to move forward to.

Considering the times and costs configuration presented in Figure 5, we get the expecetd minimum costs to destination node 4, from arcs and from nodes, presented in Table 3.

(i,j)	$Z_{(i,j),4}(t)$	(i,j)	$Z_{(i,j),4}(t)$	(i,j)	$Z_{(i,j),4}(t)$
$(1, 1_{l_1})$	43.26	$(2, 2_{l_2})$	39.84	$(3, 4)$	35
$(1, 1_{l_2})$	46.30	$(2, 2_{l_3})$	32	$(3, 3_{l_1})$	30
$(1, 1_{l_3})$	56.84	$(2_{l_2}, 2)$	39	$(3_{l_1}, 3)$	42
$(1_{l1}, 3_{l_1})$	39.26	$(2_{l_2}, 3_{l_2})$	35.48	$(3_{l_1}, 4_{l_1})$	27
$(1_{l_2}, 2_{l_2})$	41.30	$(2_{l_3}, 2)$	50.48	$(3_{l_2}, 3)$	25.48
$(1_{l_3}, 2_{l_3})$	50.84	$(2_{l_3}, 4_{l_3})$	26	$(4_{l_1}, 4)$	2
				$(4_{l_3}, 4)$	1

Table 3: Expected minimum costs, from arcs and from nodes, to node 4 in $[min]$.

Figure 6 shows the evolution of the inflow rates that are assigned to each line from each node. It depicts how the flow of passengers is distributed, at each station, among the different options to move toward station 4. It is important to highlight how Figure 6 shows the dynamic nature of the problem. Even though some of the considered functions are constant, the dynamics of the different parts of MDTrA model can be seen through the assignments evolution. As flow rate originates dynamically at station 1, it travels through the arcs arriving at different times to each station. The evolution of the inflow rates from the stations transforms the curve of the demand rate generated at station 1 from Figure 4 into the curve of flow rate arriving destination, station 4, in Figure 7.

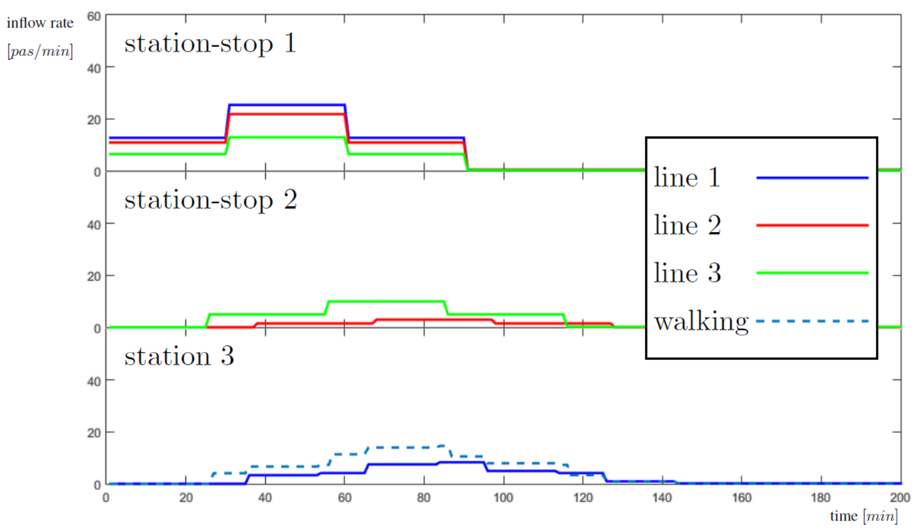

Figure 6: Evolution of inflow rates from each node.

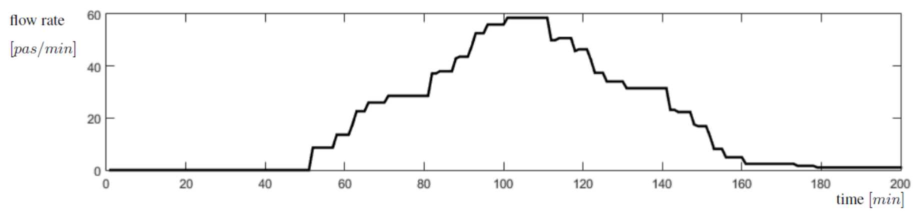

Figure 7: Evolution of flow rates arriving to station 4.

5. MDTA ALGORITHM FOR TRANSIT NETWORKS

To search for a solution for the general case of transit networks, we propose the *Markovian dynamic transit assignment* (*MDTrA*) algorithm, a dynamic programming method that works over a discretization of the analysed period of time. It is an adaptation of the MDTA algorithm (de la Paz Guala, 2020), based in a repeated and reversed-step version of *Dial's algorithm* (Dial, 1971) and constructed to perform a DTA under a Markovian approach.

The MDTrA algorithm has as inputs: (1) the digraph (N, A); (2) the set of destinations, $D \subseteq N$; (3) the time-dependent demand-supply profile $(\mathcal{D}(\cdot), \phi(\cdot))$; (4) an array P of access times, exit times and transfer penalties; (5) an array of the travel times between consecutive nodes for all lines, V; (6) the length of the time period to be analysed, T; (7) the timestep size of the discretization, Δt; and (8) the dispersion parameter for the logit model, θ. Then, being $K = T/\Delta t$ the number of time increments resulting from the discretization, the outputs of the algorithm are the arrays $E = \left(E_{ad}^k\right)_{a\in A, d\in D, k=1,\ldots,K}$ and $C = \left(C_{ad}^k\right)_{a\in A, d\in D, k=1,\ldots,K}$, where, for an arc $a \in A$, a destination $d \in D$ and a time increment $k \in \{1, ..., K\}$, E_{ad}^k is the inflow rate going to destination d entering arc a at time k and C_{ad}^k is the cost of using arc c to go to destination d entering at k.

As a pseudocode, the *MDTrA* algorithm can be expressed as follows:

Algorithm 1 (E, C) = MDTrA$((N, A), D, (\mathcal{D}(\cdot), \phi(\cdot)), P, V, T, \Delta t, \theta)$

STEP 0: INITIALIZATION Technical settings
for k=1,...,K **do**
  STEP 1: BACKWARD
  for all $d \in D$ **do**
    for all $i \in N$, in the order given π_d **do**
      for all $a = (i, j) \in A_i^-$ incoming arcs to i, **do**
        Compute expected minimum costs of using a to go to d
      end for
      Compute expected minimum costs from i to d
    end for
  end for
  STEP2: FORWARD
  for all $i \in N$ **do**
    for all $d \in D$ **do**
      for all $a = (i, j) \in A_i^+$ outgoing arcs from i, **do**
        Assign the aggregation of outflow rates of incoming arcs to i, except by (j, i), and the flow rate generated at i as inflow rate

```
            end for
        end for
    end for
    STEP 3: COST UPDATES
    for all a ∈ A do
        Update the cost of a
    end for
    STEP 4: STOP CONDITION
    if k = K or there are no more flow rates to assign then
        End MDTrA algorithm
    end if
end for
```

6. FINAL COMMENTS AND CONCLUSIONS

In this paper, we present a new approach to address the problem of dynamic transit assignment with uncertainty in passenger's choices. We do so by presenting the Markovian dynamic transit assignment (MDTrA) model, showing an example of how the MDTrA model works, giving an insight in how to use smartcard data to obtain the required exogenous demand profile, and proposing the MDTrA algorithm, a solution method for a discrete version of the original problem.

The MDTrA model, intuitivelly, seeks to represent that, according to the transit network's underlying digraph and given demand and supply profiles, at each node (station/stop) passengers travel to their destinations by choosing the next reasonable arc (line/walk) to move forward to, given their perceived costs of using said arc to go to their destinations. The demand is given by a time-dependent function that defines the flow rate of passengers from origins to destinations, while the supply is given by the time-dependent frequencies of the transit network services. The arc travel time functions are fundamental, as they allow us to represent the dynamic aspect of the problem, locating in time the different variables of the model, while the arc cost functions represent the different factors that influence the passengers travelling experience. The arc-choice model, after computing the expected minimum costs of arcs, performs the assignment according to a logit model of known dispersion parameter, and, even though paths are not directly chosen, the routes that passenger follow end up being constructed by a recursive arc-choice process.

We also present how smartcard data can be used to generate a demand profile that evolves over time. Transactions when boarding at each station/stop can be collected, recording the time and estimating the destination, following Munizaga & Palma's (2012) methodology.

The proposed MDTrA algorithm allows obtaining a solution for a version of the model considering a discretization of the time period to be analysed. In brief, after an initialization, it runs over each time increment that results from the discretization a backward step to compute expected minimum costs and then a forward step to perform the assignment.

Some of the defining features of our contribution are that: it reduces the set of options through the reasonability concept; in addition to usual modes (metro and buses), it aknowledges that passengers may choose to walk between stations, adding a mode option not usually considered in the literature; given its arc-based construction, routes costs independence and route enumeration are not needed to formulate the model or to construct the algorithm; and smartcard data, among other type of data, allows constructing the exogenous information that the model and the algorithm requires.

For later stages of this research, among other ideas, we want to: apply the effective frequency theory in the dynamic context (as in Cortés et al. (2013); Pineda et al. (2016)); study diferent instances, fictional and real, to test the implementation of the algorithm; address congestion of passengers in station/stops; represent the effect of consecutive transbords.

ACKNOWLEDGEMENTS

We thank Bastián Henríquez, student of Universidad de Chile's Doctorate in Engineering Systems program, for his fundamental assistance in the process of literature revision.

REFERENCES

Baillon, J.-B., & Cominetti, R. (2008, Jan 01). Markovian traffic equilibrium. **Mathematical Programming**, 111 (1), 33–56. Retrieved from `https://doi.org/10.1007/s10107-006-0076-2` doi: 10.1007/s10107-006-0076-2

Cepeda, M., Cominetti, R., & Florian, M. (2006). A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria. **Transportation Research Part B: Methodological**, 40 (6), 437-459. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0191261505000743` doi: https://doi.org/10.1016/j.trb.2005.05.006

Cominetti, R., & Correa, J. (2001). Common-lines and passenger assignment in congested transit networks. **Transportation Science**, 35 (3), 250-267. Retrieved from `https://doi.org/10.1287/trsc.35.3.250.10154` doi: 10.1287/trsc.35.3.250.10154

Cortés, C. E., Jara-Moroni, P., Moreno, E., & Pineda, C. (2013). Stochastic transit equilibrium. **Transportation Research Part B: Methodological**, 51, 29-44. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0191261513000210` doi: https://doi.org/10.1016/j.trb.2013.02.001

de Cea, J., & Fernández, E. (1993). Transit assignment for congested public transport systems: An equilibrium model. **Transportation Science**, 27 (2), 133-147. Retrieved from `https://doi.org/10.1287/trsc.27.2.133` doi: 10.1287/trsc.27.2.133

de la Paz Guala, R. (2020). **Mdta: Markovian dynamic traffic assignment, a new approach for stochastic dta** (Unpublished doctoral dissertation). Universidad de Chile.

Dial, R. B. (1971). A probabilistic multipath traffic assignment model which obviates path enumeration. **Transportation research**, 5 (2), 83–111.

Kiencke, U., & Nielsen, L. (2000, nov). Automotive control systems: For engine, driveline, and vehicle. **Measurement Science and Technology**, 11 (12), 1828–1828. Retrieved from `https://doi.org/10.1088/0957-0233/11/12/708` doi: 10.1088/0957-0233/11/12/708

Lam, W., Gao, Z., Chan, K., & Yang, H. (1999). A stochastic user equilibrium assignment model for congested transit networks. **Transportation Research Part B: Methodological**, 33 (5), 351-368. Retrieved from `https://www.sciencedirect.com/science/article/pii/S019126159800040X` doi: https://doi.org/10.1016/S0191-2615(98)00040-X

Munizaga, M. A., & Palma, C. (2012). Estimation of a disaggregate multimodal public transport origin–destination matrix from passive smartcard data from santiago, chile. **Transportation Research Part C: Emerging Technologies**, 24, 9-18. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0968090X12000095` doi: https://doi.org/10.1016/j.trc.2012.01.007

Nguyen, S., & Pallottino, S. (1988). Equilibrium traffic assignment for large scale transit networks. **European Journal of Operational Research**, 37 (2), 176-186. Retrieved from `https://www.sciencedirect.com/science/article/pii/037722178890327X` doi: https://doi.org/10.1016/0377-2217(88)90327-X

Nielsen, O., & Frederiksen, R. (2006, 04). Optimisation of timetable-based, stochastic transit assignment models based on msa. **Annals OR**, 144, 263-285. doi: 10.1007/s10479-006-0012-0

Nuzzolo, A., & Comi, A. (2017). Transit travel strategy as solution of a markov decision problem: Theory and applications. **2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)**, 850-855.

Pineda, C., Cortés, C. E., Jara-Moroni, P., & Moreno, E. (2016). Integrated traffic-transit stochastic equilibrium model with park-and-ride facilities. **Transportation Research Part C: Emerging Technologies**, 71, 86-107. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0968090X16300961` doi: https://doi.org/10.1016/j.trc.2016.06.021

Spiess, H., & Florian, M. (1989). Optimal strategies: A new assignment model for transit networks. **Transportation Research Part B: Methodological**, 23 (2), 83-102. Retrieved from `https://www.sciencedirect.com/science/article/pii/0191261589900349` doi: https://doi.org/10.1016/0191-2615(89)90034-9

Tavassoli, A., Mesbah, M., & Hickman, M. (2020, 10). Calibrating a transit assignment model using smart card data in a large-scale multi-modal transit network. **Transportation**, 47. doi: 10.1007/s11116-019-10004-y

Yang, L., & Lam, W. H. K. (2006). Probit-type reliability-based transit network assignment. **Transportation Research Record**, 1977 (1), 154-163. Retrieved from `https://doi.org/10.1177/0361198106197700118` doi: 10.1177/0361198106197700118