

An optimization problem for dynamic OD trip matrix estimation on transit networks with different types of data collection units.

ESTEVE CODINA, LIDIA MONTERO[†]

Universitat Politècnica de Catalunya
Statistics and Operational Research Dept.
C/ Jordi Girona, 1-3.

Campus Nord. C5 Building. Office 216

[†]Campus Nord. C5 Building. Office 215
08034 Barcelona, Spain

esteve.codina@upc.edu, lidia.montero@upc.edu

Abstract: Dynamic O-D trip matrices for public transportation systems provide a valuable source of information of the usage of public transportation system that may be used either by planners for a better design of the transportation facilities or by the administrations in order to characterize the efficiency of the transport system both in peak hours and off-peak hours. Since all these evaluations are intended to be done off-line as a part of an evaluation/planning process, the estimation models have also been aimed at processing off-line large amounts of data comprising one or several journeys. Also, initially, on-board equipments have been oriented at sending the collected information after the operational journey of the transportation units. In this context there are currently several models and methods for estimating dynamic O/D trip matrices using observed volumes. In operational environments, however, such trip tables should be dynamic. The literature on methods for estimating dynamic O/D matrices for public transport systems is much smaller than in the static case. Recently, the use of dynamic O-D trip matrices has been suggested for other contexts, such as for instance the management of alleviation strategies in disrupted systems. Typically, these recovery systems have been planned off-line using static O-D information from historical data. This paper presents a dynamic O/D trip matrix estimation model for transit networks which may be aimed at an on-line usage, that incorporates emerging information and communication technologies (ICT), especially those based on the electronic signature detection of devices used by passengers on board that provide a rich source of data of higher quality than simple counts. The procedure presented is based on the statement of a mathematical programming-based model for the adjustment of passenger counts on a diachronic network model. The distinctive approach presented in the paper takes advantage of a detection system, capable of measuring not only the passenger's load on segments, but also to track, up to some extent, the passenger trajectories on the network, comprising but not limited to end-to-end trajectories. The estimation model is not limited to a single line but is a formulation for a general public transportation network. The formulation of the estimation model permits to include requirements on passenger's route choice which may be estimated either by historical or survey data or may come from static passenger transit assignment models. As a first paper of a series of three, this one only presents and justifies the model's formulation showing how it is oriented for being solved effectively, leaving for a second one the advantages and the enhancements in the quality of the estimation of combining all the types of passenger counts. Finally, a third one would discuss algorithmic solution to the proposed model.

Key-Words: Dynamic OD Matrix Estimation from passenger counts, end-to-end passenger counters, Public Transportation Networks, Network flow models.

1 Introduction

When designing efficient public transport systems it is of utmost relevance to obtain data about the demand flows from each origin to each destination (Origin/Destination (O-D) trip matrix). While aggregated O-D flows usually are sufficient at the strategic planning level, each time more applications covering tac-

tical and operational levels demand a disaggregated time-varying O-D trip matrix. Because of that, the estimation of time-dependent origin-destination (O-D) matrices for transit networks based on observed passenger counts has been the object of a good number of works since some years ago. The technology for obtaining these counts has had also an evolution that has influenced the O-D matrix estimation models.

Usually, obtaining these passenger counts has relied on the use of Automatic Passenger Counting (APC) systems which are capable to register total boardings and/or alightings at some points in the network. The early work of [18] applies for schedule-based transit networks and assumes fixed proportions in the usage of the transit network which can be estimated using shortest path techniques initially developed in [14]. Their model follows an entropy maximizing approach which permits the use of a-priori information and they use as test networks the Mass Transit Railway in Hong Kong.

However APC systems were initially designed for fulfilling transit systems usage and mileage as part of the statistics to evaluate them (see, [3], [15]). These systems can be complemented using Automatic Data Collection (ADC), which can be classified as AVL (Automatic Vehicle Location) and tracking systems and also AFC (Automatic Fare Collection) systems although in the case of buses on-board ticketing systems do not record where and when passengers get off a bus. In [11], other types of information collected by these systems are GPS Coordinates, Time and Date, Vehicle Stops / Starts, door opening and closing, usage of wheelchairs. This makes possible other types of analysis such as route analysis and adherence to prespecified schedules. Also, in [16] the use of archived data coming from ADC systems combined with smart card data is explored in order to obtain trip chaining and to infer O-D matrices and to have information about the changes in patterns during weekdays and weekends and also connecting bus routes headways. The information provided by dynamic O-D tables it is known to contribute to design better transit schedules since it is then possible to determine where the effects of the congestion due to peak periods appear in terms of longer waiting times at stations and stops and occupancy of units. In [13] it is illustrated how the use of smart card-based automated fare collection systems makes possible to enhance the quantification of changes in the demand at the spatial-temporal level.

Another approach for estimating dynamic O-D trip matrices in transit systems has been using new and modern models for dynamic passenger transit assignment (see, for instance [4]) which may reproduce congestion effects. These models are embedded into a bilevel programming framework, thus paralleling schemas used for O-D trip matrix estimation in static traffic and transit networks. This is the approach followed in [6] where an heuristic method is developed. The method is applied only to small transit networks. Recently, new systems that use off-the-shelf Bluetooth hardware systems for detecting end-to-end passenger journeys have arisen (see [8]). These systems

have had a similar usage for the case of dynamic O-D trip matrices estimation in traffic networks [] These systems are used along with the previous systems for AVL for complementing the data collection as well as to match trips to stops along the journey. These systems have already been considered in traffic networks also for the purpose of estimating dynamic O-D trip matrices, as in [1]. These dynamic O-D trip matrix estimation procedures have been considered for solving advantageously several additional applications at the operational/tactical level such as transit recovery approaches based on bus-bridging, as in [19] and [17], whereas traditionally, as in [7], the travel demand has been assumed derived from historical fare card data.

In this paper a formulation for the problem of estimation of dynamic O-D trip matrices using several sources of on-line information coming from passenger counts is presented. The estimation model combines traditional countings for boardings and alightings, first boarding validating devices and counts coming from detectors of electronic addresses of Bluetooth/WiFi devices used by passengers in, for instance, mobile phones. In the paper it is assumed that these detection systems may operate jointly under the proper technological or engineering developments or that they can be developed in a near future and that existing deficiencies (if any) may be overcome ([2]). Although previous works based in this new source of information exist, (as in [8]) the estimation of O-D flows of passengers is limited to movements of users of in a single line equipped with Bluetooth detectors. The distinctive approach of this paper is to present a model based in mathematical programming for the estimation of dynamic O-D trip matrices for a set of lines making up a (sub)network of public transport, taking into account the previously mentioned passenger counting methods assuming that a proper on-board software system operates on the units, that the detection technology Bluetooth/WiFi-based has enough reliability and also, that the appropriate communications system permits to send this information during the units' operation.

2 Structure of the model

The core of the model is a diachronic network where links are associated to possible passenger movements and movements of units in a service during a period of time of length T . Throughout the text, double indexed elements such as links in a graph will be denoted either by a single letter a or by the labels of the tail and head nodes, (i, j) at convenience. For a graphical description of this diachronic network see Figure 1. In order to capture the effects of the flows' evolution out-

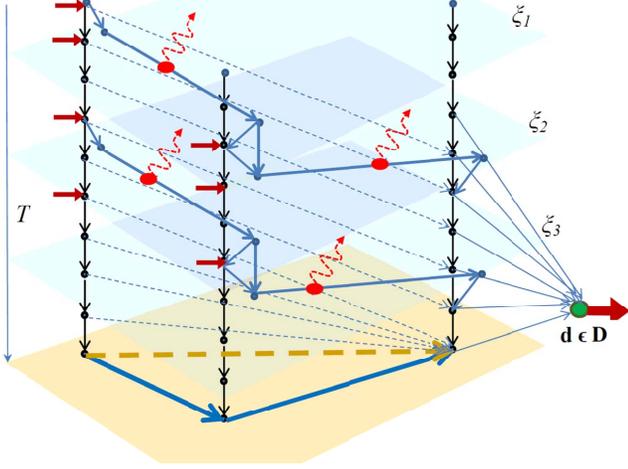


Figure 1: Discrete diachronic network $G = (N, A)$.

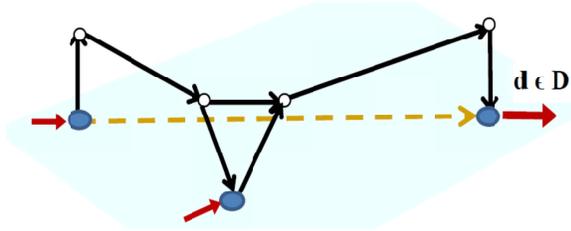


Figure 2: The corresponding static transit network $\mathbf{G} = (\mathbf{N}, \mathbf{A})$ to the diachronic network in Figure 1.

side the time horizon, the diachronic network "ends" in a layer on which nodes and links reproduce the planar structure of the transportation network. This part of the diachronic network will be referred to the "planar subnetwork".

The diachronic network has a structure very similar to the ones in [10]. It will be associated to a time granularity given by a small time subinterval Δ , typically of ≤ 1 minute. The graph for the diachronic network will be denoted by $G = (N, A)$. Nodes $i \in N$ are assigned a time label t_i . The set of links A is divided into $A = A_x \cup A_e \cup A_T \cup A_a \cup A_y \cup A_P \cup A_d$. Links in each of these subsets play different roles as shown in Figure 3 below (A_P are links for modelling transfers). Nodes i, j in a link $a \in A_T$ make the subset of nodes N_T . The forward and backward stars of links emerging/incoming from/at node $i \in N$ will be denoted as $E(i)$ and $I(i)$ respectively.

Let also L denote the set of transport lines. The sequence of stops in line $\ell \in L$ will be denoted by $\mathbf{S}(\ell)$. Let n_ℓ the number of services initially expected for line $\ell \in L$ during the time horizon.

Also, for services s on line ℓ , $1 \leq s \leq n_\ell$, the double index $\pi = (\ell, s)$ will be used. At a stop $\sigma \in \mathbf{S}(\ell)$, the links in the diachronic network corresponding to a service $\pi = (\ell, s)$ will be denoted by $a_e(\pi, \sigma)$, $a_a(\pi, \sigma)$ and $a_y(\pi, \sigma)$ and appear in next figure 3.

Passenger flows on links $a \in A_x$ are dwelling at a stop during boarding/alighting operations; flows on links $a \in A_e$ are moving through a line segment from station to station; flows on links $a \in A_a$ are boarding from a stop whereas flows on links $a \in A_y$ are descending to a stop. Transfers and other external movements are captured by links $a \in A_P$. Passenger flows waiting at stops are captured by links $a \in A_T$. $A_x(\ell)$, $A_y(\ell)$, $A_e(\ell)$, $A_a(\ell)$ will denote subsets associated to line $\ell \in L$, $A_x(\ell, s)$, ... the subsets associated to service s for line $\ell \in L$.

For any node $i \in N$ in the diachronic network, the corresponding node in the static network will be denoted by $\mathbf{i}(i)$. A node $i \in N$ which is head or tail of a link in A_x , A_y or A_a , has associated a transit stop in the static network. Given a service π and a stop σ it will be convenient to identify the corresponding nodes in the diachronic network. The nodes in the diachronic network within N_T will be denoted by $i(\pi, \sigma)$ (see figure 3). The set of all such nodes for a service π will be denoted by $\mathbf{i}(\pi)$.

In a static representation, the transportation network could be modeled by a graph $\mathbf{G} = (\mathbf{N}, \mathbf{A})$ for the spatial distribution of the transportation system. As a convenient notation, all elements associated to the planar graph will be written in bold. In the static network there are some nodes that play the role of cen-

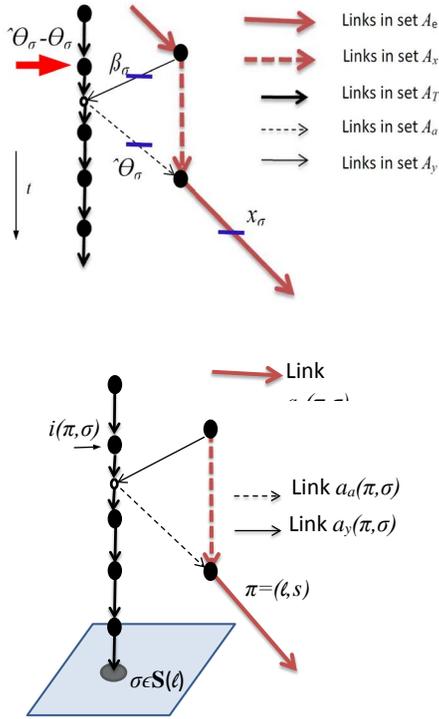


Figure 3: Types of links in the diachronic network $G = (N, A)$ and volumes captured by the detection system (left). For a given service s in a line ℓ , if $\pi = (\ell, s)$, the associated links $a_e(\pi, \sigma)$, $a_y(\pi, \sigma)$ and $a_o(\pi, \sigma)$ and the nearest node $i(\pi, \sigma)$ in the diachronic graph are shown, for a stop $\sigma \in \mathbf{S}(\ell)$.

troids or nodes where static O-D flows may originate or arrive. The subset of centroids that are destinations will be denoted by \mathbf{D} and the set of origin-destination pairs will be denoted by \mathbf{W} .

The origin-destination structure of the diachronic network is as follows. Passenger flows may originate at nodes $i \in N_T$, so that $\mathbf{i}(i)$ is a centroid in the static network. These nodes $i \in N_T$ are associated to a time subinterval. These passenger flows will end at a given destination $\mathbf{d} \in \mathbf{D}$. The arrival time to the destination will not in general be known. Because of this, the destination nodes in the diachronic network can be considered to lay in the planar subnetwork. Also, the label for destination nodes in the diachronic network will be $\mathbf{d} \in \mathbf{D}$, whereas for other nodes and links not in the planar subnetwork, no boldface characters will be used. Notice that destination nodes in the diachronic network are not associated to a time subinterval. Thus, the origin-destination pair set, denoted by W , will be a subset of $N_T \times \mathbf{D}$ and their elements $\omega \in W$ will be of the form $\omega = (i, \mathbf{d})$.

For a node $\mathbf{i} \in \mathbf{N}$ in the planar network, there exists a set of nodes in the diachronic network $N(\mathbf{i})$. Likewise, for a link $\mathbf{a} = (\mathbf{i}, \mathbf{j}) \in \mathbf{A}$ in the planar network there will exist the corresponding set of links $A(\mathbf{i}, \mathbf{j})$ in the diachronic network, where $A(\mathbf{i}, \mathbf{j})$ is defined as:

$$A(\mathbf{i}, \mathbf{j}) = \{ (i, j) \in A \mid i \in N(\mathbf{i}) \ \& \ j \in N(\mathbf{j}) \} \quad (1)$$

For each node $\mathbf{i} \in \mathbf{S}$ in the planar graph, the diachronic graph G contains a linear sequence of links $a \in A_T(\mathbf{i})$, each of them with a travel cost Δ . Their head and tail nodes j_a and i_a , respectively, are assigned a time label that is a multiple of Δ . For a node $\mathbf{i} \in \mathbf{S}$, in the planar graph, the set of nodes that are head and tail for links in $A_T(\mathbf{i})$ will be denoted by $N_T(\mathbf{i}) = \{j_a, i_a \mid (i_a, j_a) \in A_T(\mathbf{i})\}$. Passenger flows on these links correspond to waiting at node $\mathbf{i} \in \mathbf{S}$.

During the time horizon considered, the journey of a bus or transport unit along a transportation line will be referred to as a service, and each service s , $1 \leq s \leq n_\ell$, in a line $\ell \in L$, will be indexed by a double index $\pi = (\ell, s)$. In the diachronic network a service π is represented by a sequence of links $a_1, a_2, \dots, a_\nu \in A$, so that $a_{j-1} \in A_e$, $a_j \in A_x$, $a_{j+1} \in A_e \dots$ and $a_1, a_\nu \in A_e$.

3 Network flows and feasibility constraints

In the following a notation will be described for the various types of vector flows and origin-destination volumes. In previous section the structure of the diachronic network has been described and as it will be shown in the model, it is necessary that the flows during the time horizon on this diachronic network should keep consistency with the aggregated flows on a static representation of the network (the static network). Thus, whereas in figure 1 the diachronic network is shown, in figure 2 appears the corresponding static transit network.

The routing and behaviour of passengers on urban transit networks has been extensively modelled and is a key component of transit assignment models. For the case of static transit networks these models differ very much from static traffic networks. Whereas in the latter case the routing is associated directly to decisions based on which is the set of paths that are chosen by travellers, in the case of transit networks containing transport lines ruled by a frequency based schedule, user decisions are associated with the concept of strategies which are represented by hyperpaths (see, [12], [9]).

In subsection 3.1 the relationship between flows on the diachronic network and on the static network is shown and also, how the flows to the static network are forced to obey to routes or hyperpaths which may be registered from surveys or that may come from external planning models of transit assignment.

In general bold vectors will apply for the static network, whereas non-bold will apply for the diachronic network. Also, superscripts will apply for vectors whereas subscripts will denote components of a (possibly superscripted) vector. In the formulations of section 6, flows on the diachronic and static network will be structured per O-D pair. Thus, the following notation will be used:

- $v^\omega = (\dots, v_a^\omega, \dots; a \in A) \in \mathbb{R}_+^{|A|}$, where, as stated earlier, $\omega = (i, \mathbf{d}) \in W$, is the vector of O-D link flows on the diachronic network, being i a node in N_T so that $\mathbf{i}(i)$ is a centroid in the static network and \mathbf{d} is a centroid for which, in the static network, there exists a destination also with label \mathbf{d} .
- $v = \sum_{\omega \in W} v^\omega \in \mathbb{R}_+^{|A|}$ is the vector of total flows on links of the diachronic network and a component of it will be $v_a = \sum_{\omega \in W} v_a^\omega$, $a \in A$.
- g^ω is the O-D flow on the diachronic network model. Since $\omega = (i, \mathbf{d})$ and a time instant is associated to node $i \in N_T$ the set of flows g^ω is a dynamic transit O-D matrix. An alternative notation will be $g^{i,\mathbf{d}}$.
- $e^\omega \in \mathbb{R}^{|N|}$. Associated to an O-D pair $\omega \in W$ on the diachronic network, the vector e^ω is defined as:

$$(e^\omega)_{i'} = \begin{cases} 1 & \text{if } i' = i \\ -1 & \text{if } i' = \mathbf{d} \\ 0 & \text{if none of the above} \end{cases} \quad (2)$$

- $\mathbf{v}^\mathbf{w} = (\dots, \mathbf{v}_\mathbf{a}^\mathbf{w}, \dots; \mathbf{a} \in \mathbf{A}) \in \mathbb{R}_+^{|\mathbf{A}|}$, $\mathbf{w} \in \mathbf{W}$, is the vector of O-D link flows on the static transit network.
- $\mathbf{g}^\mathbf{w}$ is the O-D flow for the O-D pair $\mathbf{w} \in \mathbf{W}$ on the static transit network.

The diachronic network model and also the static transit network model can be both stated as a multi-destination network flow. This fact will be used advantageously when algorithmic solutions to the problem will be devised.

Notice that as index a is shorthand for an ordered pair of indexes (i, j) , both in N , then $v_a, v_a^\omega, \dots, a \in$

A are equivalent to $v_{i,j}, v_{i,j}^\omega, \dots, (i, j) \in A$. The same applies for $\mathbf{v}_\mathbf{a}, \mathbf{a} \in \mathbf{A}$, and $\mathbf{v}_{\mathbf{i},\mathbf{j}}$, since $\mathbf{a} = (\mathbf{i}, \mathbf{j})$.

3.1 Relationship between flows on the diachronic network and the strategy-based behaviour of transit flows.

Line capacity constraints should be taken into account at links in A_e corresponding to movements on line segments of the diachronic network. Line capacities c^π will directly associated to a service π and thus c^π is the maximum number of passengers that be on board of the unit in service π . Then, ideally:

$$v_a \leq c^\pi, \forall a \in A_e(\pi), \quad 1 \leq s \leq n_\ell, \ell \in L, \pi = (\ell, s) \quad (3)$$

Let $\mathcal{H}^\mathbf{w}$ be a selected set of strategies for O-D pair $\mathbf{w} = (\mathbf{o}, \mathbf{d})$ $\mathbf{d} \in \mathbf{D}$, which are known to be likely used by trip makers of the transportation network. This set of strategies should be determined off-line using historical O/D information for the public transport system or by using a static transit assignment model of the network. Let $s_h^\mathbf{w}$ a vector corresponding to strategy $h \in \mathcal{H}^\mathbf{w}$ such that if $\gamma_h^\mathbf{w}$ is the O-D flow of passenger following strategy h for O-D pair $\mathbf{w} = (\mathbf{o}, \mathbf{d})$, then the total O-D flow $\mathbf{g}^\mathbf{w}$ will be decomposed into the flows $\gamma_h^\mathbf{w}$ for each of the strategies:

$$\mathbf{g}^\mathbf{w} = \sum_{h \in \mathcal{H}^\mathbf{w}} \gamma_h^\mathbf{w}, \quad \mathbf{w} = (\mathbf{o}, \mathbf{d}) \in \mathbf{W} \quad (4)$$

Also, the O-D flow $\mathbf{g}^\mathbf{w}$ can be expressed in terms of the dynamic O-D grip matrix g^ω for those O-D pairs on the diachronic network $\omega = (i, \mathbf{d})$, with a common destination \mathbf{d} and origin i whose projection on the static network is the origin \mathbf{o} of O-D pair $\mathbf{w} \in \mathbf{W}$.

$$\mathbf{g}^\mathbf{w} = \sum_{i \in N(\mathbf{o})} g^{i,\mathbf{d}}, \quad \mathbf{w} = (\mathbf{o}, \mathbf{d}) \in \mathbf{W} \quad (5)$$

Likewise, the components of the vector of link flows per O-D pair $\mathbf{v}^\mathbf{w}$ on the static network must be expressed in terms of the flows on the diachronic network:

$$\mathbf{v}_\mathbf{a}^\mathbf{w} = \sum_{i \in N(\mathbf{o})} \sum_{a' \in A(\mathbf{a})} v_{a'}^{i,\mathbf{d}}, \quad \mathbf{w} = (\mathbf{o}, \mathbf{d}) \in \mathbf{W}, \mathbf{a} \in \mathbf{A} \quad (6)$$

and also, it will be written in terms of the vectors $s_h^\mathbf{w}$ corresponding to the strategies $h \in \mathcal{H}^\mathbf{w}$ as:

$$\mathbf{v}^\mathbf{w} = \sum_{h \in \mathcal{H}^\mathbf{w}} \gamma_h^\mathbf{w} s_h^\mathbf{w}, \quad \mathbf{w} = (\mathbf{i}, \mathbf{d}) \in \mathbf{W} \quad (7)$$

Notice that both v^w and g^w can be considered as the projection of flows in the diachronic network onto the static transit network.

4 Description of the data collection system

Four types of units for data collection will be assumed to coexist in the public transportation network:

1. Not equipped at all with any passenger detection system. For these type of units the only information known is the planned time schedule. They will be referred to as type-D units.
2. Units equipped with Automatic Passenger Counting (APC) systems, which permit to count total boardings and total alightings at a stop. It will be assumed that this information can be sent with a given periodicity while the unit is in operation. They will be referred to as type-C units and the detections system will be referred to as a type C detection system.
3. Type-B units, additionally to be equipped with type-C detection system, also count the passengers that validate their tickets by first time when boarding the unit. This detection system will be referred to as a type B detection system. It will be assumed that this information can be sent with a given periodicity while the unit is in operation.
4. Type-A units are those that in addition to have the type-B equipment are able to identify, store and process the labels or electronical mark/address of some electronical devices used by passengers (mobile/smart phones, tablets, ...). This specific detection equipment will be referred to them as type A detection system. They also send periodically all their available information.

The characteristics of the equipment of type A vehicles are the following:

The data detection system is assumed to be composed of the following types of devices: on board of the units it is assumed a) that an Automatic Passenger Counting (APC) system is installed and permits to know total boardings and alightings, b) along its journey passengers are identified by some electronic address that allows to know which is the sequence of line segments carried out on the unit (this identification, conveniently encrypted, is lost after the passenger exits the unit) and c) if the system comprises metro lines, then there is also an APC system that permits total number of entering passengers. It is assumed



Figure 4: A sensor for detecting passengers boarding and alighting.

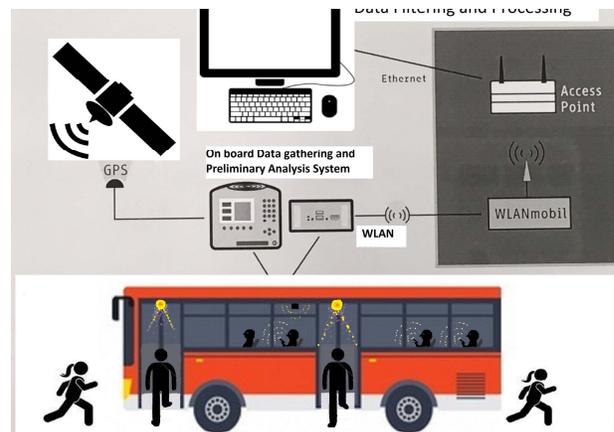


Figure 5: A schematic picture showing the different components of the passenger's detection system. Presence detectors based on electronic addresses, boardings and alightings counters and ticketing system.

also that the system is well connected with the ticketing system and detects whether a passenger boards by first time or it comes from a transfer from another line.

Figure 5 provides a schematic description of the A and B types of detection systems that may be operating simultaneously at a type A transport unit, as well as the communication networks. It is assumed that type A equipment comprises an on-board data gathering and preliminary analysis system (ODGPAS) capable to provide the information of type A detected passengers in the format specified in next subsection 4.1 and combine it with the automatic ticketing system (type B detection equipment).

The error sources may be originated by several factors: passengers may switch off their mobile phones while they are boarded on the vehicles or they may switch on the mobiles on board after their 2nd stop. Detectors devices may count external mobiles and also, there can be passengers not paying the trans-

port fee ... amongst others.

It is assumed that at time instants $\zeta_1, \zeta_2, \dots, \zeta_k \in [0, T]$ the data gathered by the collection system on units are sent to an external data processing system. We will refer to this amount of information as a *burst*. Because of the different types of on board equipments described at the beginning of this section, services $\pi = (\ell, s)$ operating in the time period may be comprised in the corresponding different subsets: $\Pi^A, \Pi^B, \Pi^C, \Pi^D$.

At each time instant ζ_ν in the previous sequence, let $\Pi(\zeta_\nu)$ be the set of services that have sent its data to the data processing center,

$$\Pi(\zeta_\nu) = \{ \pi = (\ell, s), \ell \in L, 1 \leq s \leq n_\ell \mid \exists (i, j) \in A_e(\pi), t_i \leq \zeta_\nu < t_j \} \quad (8)$$

and let $\sigma(\pi, \zeta_\nu)$ be the nearest stop that service $\pi = (\ell, s)$ has just left behind at time monitoring ζ_ν . The set of all stops that have been just monitored at time instant ζ_ν , $\mathbf{S}(\zeta_\nu)$ by service $\pi = (\ell, s)$, will be given by

$$\mathbf{S}^-(\zeta_\nu) = \{ \sigma_1, \sigma_2, \dots, \sigma_{N_\nu} \} \quad (9)$$

Thus, the first stop monitored is σ_1 (the stop at the beginning of the line), whereas at a time instant ζ_k , a vehicle unit on $\pi = (\ell, s) \in \Pi(\zeta_\nu)$, has just visited stop $\sigma_{N_\nu} \equiv \sigma(\pi, \zeta_k)$ and the sequence of previously visited stops will be in $\mathbf{S}^-(\ell, \sigma(\pi, \zeta_k))$.

For simplicity of exposition it will be assumed the number of stops of any public transportation line covered by the system is a multiple of some integer $r > 1$. It will be also assumed that transport units of type A send their bursts of information right after visiting r stops and that no interruptions occur of this process of detection during the operational period.

Also, the models that are to be described assume that the passenger flows can be divided accordingly to the following two concepts:

1. whether they have or not an electronic mark/address that can be detected in type A detection systems and. (type A passenger flows and type \bar{A} passenger flows)
2. at the local level of the type A detection systems, whether they have boarded a unit as result of transfer or by first time. It must be noticed that type B detection systems discriminate clearly transfers from first boardings, although of course, there may be first boardings of type A passengers and also first boardings of type \bar{A} passengers.

4.1 Description of the informaton sent in a burst accordingly to the transportation unit type.

4.1.1 Type C

- $\hat{\beta}_j, \hat{\alpha}_j$ total boardings and alightings respectively at stop j .
- At the beginning of a burst, total passengers on board \hat{X}_1 are known and then the remaining total of passengers \hat{X}_j may be known, $2 \leq j \leq r$, since $\hat{\beta}_j, \hat{\alpha}_j$ are known:

$$\begin{aligned} \hat{X}_{r-1} &= \hat{X}_r - \hat{\alpha}_r + \hat{\beta}_r \\ \hat{X}_{r-2} &= \hat{X}_{r-1} - \hat{\alpha}_{r-1} + \hat{\beta}_{r-1} \\ &\dots \\ \hat{X}_1 &= \hat{X}_2 - \hat{\alpha}_2 + \hat{\beta}_2 \end{aligned} \quad (10)$$

4.1.2 Type B

- θ_j total boardings at stop j by first time.

4.1.3 Type A

- β_j, α_j total boardings and alightings of type A passengers respectively at stop j .

A transportation unit having visited N_ν stops will have registered at the corresponding time instants the following information:

- from stop 1, the following amounts of type A passengers: $x_{11}, x_{21}, \dots, x_{N_\nu 1}$. The total cumulative amount of passengers up to that moment, $X_1 = \sum_{\ell=1}^{N_\nu} x_{\ell 1}$, will be recorded.
- \vdots
- from stop j , the following amounts of type A passengers: $x_{jj}, x_{j+1,j}, \dots, x_{N_\nu j}$. The total cumulative amount of passengers up to that moment, $X_j = \sum_{\ell=1}^{N_\nu} x_{\ell j}$, will be part of the information in the burst. Notice that X_j is the measured occupancy of type A passengers in the transport unit after leaving stop j -th in the sequence of r stops comprised in the burst.

Thus, x_{ij} is, from the number of passengers of type A that boarded at stop i , those that are detected right after stop j ($i \geq j$). Obviously $x_{ij} \geq x_{ij'}$ if $j' > j$.

The previous information, x_{ij} , will be conveniently aggregated by the ODPAS, which will be assumed to have a buffer memory of size M^- recording

occupations of the, at most, previous M^- stops and taking into consideration passenger movements which will be ahead at most M^+ stops. Usually for a burst in the middle of the transport line, $M^+ = M^-$, but for a burst at the beginning of the line $M^- = 1$, whereas $M^+ > 1$ and at the final burst $M^- > 1$, whereas $M^+ = 1$. With this in mind the following magnitudes will be calculated for each burst of size r by the ODGPAS:

$$\begin{aligned} \beta_j &= x_{jj}, & 1 \leq j \leq r \\ \alpha_j &= X_{j+1} - X_j + \beta_j, & 1 \leq j \leq r \end{aligned} \quad (11)$$

$$X_{r+1} = \sum_{\ell=r+1}^{r+M^-} x_{\ell,r+1}$$

$$X_2 = \sum_{k \geq 2}^{r+M^-} x_{k,2}, \quad X_3 = \sum_{k \geq 3}^{r+M^-} x_{k,3}, \dots \quad (12)$$

$$X_j = \sum_{k \geq j}^{r+M^-} x_{k,j}, \quad 1 \leq j \leq r$$

Although not relevant in the details of the model formulation, for type A transport units it will be possible to know which are the entry and exit times at a stop j through some global positioning system.

4.2 Fitting the flow of type A passengers within a burst

Because of detecting errors that may happen in equipments of type A the flows involved in a burst may differ from the measured flows sent to the data filtering and processing system (see figure 5). The theoretical flows involved in a burst appear depicted in figure 7. By a simple rearrangement of nodes in figure 5, the bipartite structure of the graph can be easily shown. Assume that a burst has a buffer for M^- previous stops and for M^+ subsequent stops. By η_j , $1 \leq j \leq M^- - 1$, it is designated the flow variable corresponding to the total number of type A passengers that boarded at the j -th stop, whereas η_{M^-} is the flows corresponding to the M^- -th and previous stops. Likewise, by φ_j , $1 \leq j \leq M^+ - 1$, it is designated the total flow that will alight at the next j -th stop boarding either at a stop in the burst or previous to the burst. By $u_{y(j)}$ and $u_{b(j)}$ are designated the alighting flows and the boarding flows of type A passengers in the stops $1 \leq j \leq r$.

The remaining flow variables f, h along with variables φ and η appear in form of a table shown below for the case $r = 4$, $M^+ = 4$, $M^- = 5$ (the same case than the one depicted in figure 5).

	1	2	3	4	1	2	3	4	
1	0	...		0	h_{11}	0	...	0	$u_{b(1)}$
2	f_{21}	0	...	0	0	h_{22}	0	0	$u_{b(2)}$
3	f_{31}	f_{32}	0	...	0	0	h_{33}	0	$u_{b(3)}$
4	f_{41}	f_{42}	f_{44}	0	0	...	0	h_{44}	$u_{b(4)}$
5	f_{51}	f_{52}	f_{53}	f_{54}	0	...	0	h_{54}	η_1
6	f_{61}	f_{62}	f_{63}	f_{64}	0	...	0	h_{64}	η_2
7	f_{71}	f_{72}	f_{73}	f_{74}	0	...	0	h_{74}	η_3
8	f_{81}	f_{82}	f_{83}	f_{84}	0	...	0	h_{84}	η_4
9	f_{91}	f_{92}	f_{93}	f_{94}	0	...	0	h_{94}	η_5
	$u_{y(1)}$	$u_{y(2)}$	$u_{y(3)}$	$u_{y(4)}$	φ_1	φ_2	φ_3	φ_4	

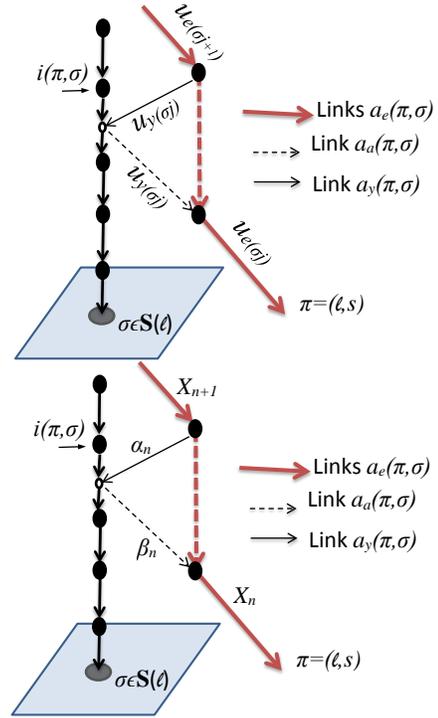


Figure 6: Top: flows on the time expanded network. Bottom: the detected flows; X_n total detected occupancy, α_n alightings and β_n boardings.

The following relationships (13) to (17) implement the balance equations of flows at nodes of the network corresponding to a burst. Notice that at the left side of the colon appears the measured flow or observed volume corresponding to the flow on the left hand side of the equality relationship.

$$2 \leq j \leq r : \quad (13)$$

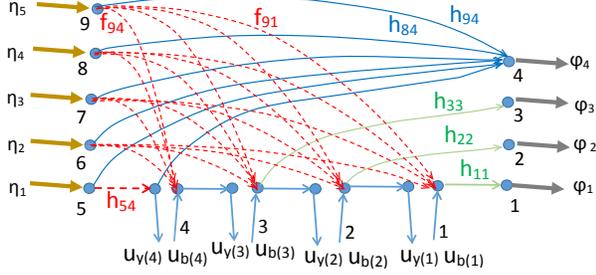


Figure 7: Network for the passenger flows captured in a burst. $r = 4$, $M^- = 5$, $M^+ = 4$.

$$\begin{aligned}
 X_1 : \quad & u_{e(\sigma_1)} = \sum_{k=1}^{M_+} \varphi_k \\
 X_2 : \quad & u_{e(\sigma_2)} = \sum_{\ell>1}^{r+M_-} f_{\ell 1} + \sum_{k=2}^{M_+} h_{kk} + \sum_{k=1+M_+}^{r+M_-} h_{kM_+} \\
 & \vdots \\
 X_j : \quad & u_{e(\sigma_j)} = \sum_{\ell>j}^{r+M_-} f_{\ell 1} + \sum_{k=j}^{M_+} h_{kk} + \sum_{k=1+M_+}^{r+M_-} h_{kM_+} \\
 2 \leq j < r : \quad & \quad \quad \quad (14)
 \end{aligned}$$

$$\begin{aligned}
 \alpha_j : \quad & u_{y(\sigma_j)} = \sum_{k>j}^{r+M_-} f_{kj} \quad (= u_{M(j)X(j)}) \\
 \beta_j : \quad & u_{b(\sigma_j)} = \sum_{\ell=1}^{j-1} f_{k\ell} + h_{jj} \quad (= u_{L(j)G(j)})
 \end{aligned}$$

$$\begin{aligned}
 1 \leq j \leq M_-, \text{ let } \bar{M}(j) = \min\{M_+, r+j\} : \quad & (15) \\
 x_{j+r,r+1} : \quad & \eta_j = \sum_{\ell=1}^r f_{j+r,\ell} + h_{j,\bar{M}(j)},
 \end{aligned}$$

$$\begin{aligned}
 1 \leq j \leq M_+ - 1 : \quad & (16) \\
 x_{j,1} : \quad & \varphi_j = h_{jj} \\
 \sum_{j=r}^{r+M_-} x_{j,1} : \quad & \varphi_{M_+} = \sum_{\ell=M_+}^{r+M_-} h_{\ell,M_+}
 \end{aligned}$$

$$\begin{aligned}
 x_{j,r} : \quad & h_{jj}, \quad 1 \leq j \leq r-1 \\
 x_{k,j} - x_{k,j-1} : \quad & f_{k,j-1}, \quad r+M_- \geq k \geq j, \\
 & k \geq 2 \\
 x_{j,1} : \quad & h_{j,M_-}, \quad r+1 \leq j \leq r+M_- \\
 & (17)
 \end{aligned}$$

5 Filtering the flow variables in a burst

For flows and measures of type A passengers, a quadratic optimization problem may be defined in order to make flows close to measures. This problem assumes that the measured flows X, α, β, x , in a burst, are known. The objective function ϕ^2 depends on the flow variables of this quadratic problem f, h, η, φ, u , i.e., $\phi = \phi(f, h, \eta, \varphi, u)$. It consists of the terms $\phi_e, \phi_y, \phi_b, \phi_\eta, \phi_\varphi, \phi_f, \phi_h$, so that $\phi^2 = \phi_e^2 + \dots + \phi_f^2 + \phi_h^2$. These terms are defined as follows in (18).

$$\begin{aligned}
 \phi_e^2 &= \sum_{j=2}^r (u_{e(j)} - X_j)^2, \quad \phi_\varphi^2 = \sum_{j=1}^{M_+-1} (h_{jj} - \varphi_j)^2 \\
 \phi_y^2 &= \sum_{j=1}^r (u_{y(j)} - \alpha_j)^2, \quad \phi_\eta^2 = \sum_{j=1}^{M_-} (\eta_j - x_{j+r,r+1})^2 \\
 \phi_b^2 &= \sum_{j=2}^{r+M_-} (u_{b(j)} - \beta_j)^2 \\
 \phi_f^2 &= \sum_{k \geq 2} \sum_{j \leq k} (x_{kj} - x_{k,j-1} - f_{k,j-1})^2 \\
 \phi_h^2 &= \sum_{j=1}^{r-1} (x_{jr} - h_{jj})^2 + \sum_{j=r+1}^{r+M_+} (x_{j1} - h_{jM_-})^2 \\
 & (18)
 \end{aligned}$$

Then, for a generic burst of size r and buffers M_-, M_+ consider the following optimization problem with the optimization variables $z^\top \triangleq (f^\top, h^\top, \eta^\top, \varphi^\top, u^\top)^\top$ and some arbitrary parameter vector λ :

$$\begin{aligned}
 \text{Min}_z \quad & \frac{1}{2} \phi^2 + \lambda^\top z \\
 \text{s.t.} : \quad & \mathcal{C} \begin{pmatrix} f \\ h \end{pmatrix} = \begin{pmatrix} \eta \\ \varphi \\ u \end{pmatrix}, \quad f, h \geq 0 \\
 & \eta, \varphi \geq 0, \quad (19)
 \end{aligned}$$

$$\begin{aligned}
 \# \text{variables} : \quad & \frac{1}{2} r(r-1) + rM_- + M_+ \\
 \# \text{constraints} : \quad & 2r + M_+ + M_- \\
 \# \text{observed flows} : \quad & \frac{1}{2} r(r-1) + (r-1)M_- + M_+ \\
 \# \text{var} - \# \text{obs.flows} - \# \text{constr.} : \quad & -3r < 0
 \end{aligned}$$

6 A quadratic problem for determining a dynamic O-D trip matrix

6.1 The quadratic problem

In previous section 5 a quadratic problem has been presented in order to obtain consistent

flow variables internal to a burst (i.e., variables $(f^\top, h^\top, \eta^\top, \varphi^\top, u^\top)^\top$) with detections of flows of type A passengers. It must be noticed that variables u are flows on links in the time expanded network described in section 3.

In the problem presented in this section two types of flows may appear: a) those flows corresponding to passengers that don't have devices with detectable electronic address. In order to distinguish them the associated magnitudes, flows on links, O-D flows, will be topped with an overbar, i.e. \bar{u} , \bar{v} , \bar{g} . Notice that this notation has not been adopted in previous sections since it was not still necessary and b) those flows corresponding to passengers with devices with detectable electronic address. The corresponding variables will not have any specific mark.

In this subsection a large quadratic problem is formulated in order to enforce consistency between internal flows in the set of bursts comprised in a given time period and flows on the time expanded network for that same time period.

6.2 Taking into account strategies

In subsection (3.1) the macroscopic flows \mathbf{v}^w resulting from the used strategies for each O-D pair $\mathbf{w} \in \mathbf{W}$ are formulated using (7). It must be noticed that, for any O-D pair $\mathbf{w} \in \mathbf{W}$, the amount of flow γ_h^w following each strategy h is left to be adjusted by the model itself and it will be considered unknown a priori. On the other hand, these macroscopic flows should be the result of aggregating flows of the corresponding links from the dichronic network, as it is expressed in relationship (6). Then, in order to include the previous relationships as penalties in the model, it will be convenient to define the following functions:

$$\psi_{\mathbf{a},\mathbf{w}} = \sum_{h \in \mathcal{H}^w} \gamma_h^w (s_h^w)_{\mathbf{a}} - \sum_{i \in N(\mathbf{o})} \sum_{a' \in A(\mathbf{a})} v_{a'}^{i,\mathbf{d}}, \quad \begin{array}{l} \mathbf{w} \in \mathbf{W} \\ \mathbf{w} = (\mathbf{o}, \mathbf{d}) \\ \mathbf{a} \in \mathbf{A} \end{array} \quad (20)$$

$$\bar{\psi}_{\mathbf{a},\mathbf{w}} = \sum_{h \in \mathcal{H}^w} \bar{\gamma}_h^w (s_h^w)_{\mathbf{a}} - \sum_{i \in N(\mathbf{o})} \sum_{a' \in A(\mathbf{a})} \bar{v}_{a'}^{i,\mathbf{d}}, \quad \begin{array}{l} \mathbf{w} \in \mathbf{W} \\ \mathbf{w} = (\mathbf{o}, \mathbf{d}) \\ \mathbf{a} \in \mathbf{A} \end{array} \quad (21)$$

Likewise, as regards to relationships (4) and (5), it will be convenient to define the following functions:

$$Q_{\mathbf{w}} = \sum_{h \in \mathcal{H}^w} \gamma_h^w - \sum_{i \in N(\mathbf{o})} g^{i,\mathbf{d}}, \quad \begin{array}{l} \mathbf{w} \in \mathbf{W} \\ \mathbf{w} = (\mathbf{o}, \mathbf{d}) \end{array} \quad (22)$$

$$\bar{Q}_{\mathbf{w}} = \sum_{h \in \mathcal{H}^w} \bar{\gamma}_h^w - \sum_{i \in N(\mathbf{o})} \bar{g}^{i,\mathbf{d}}, \quad \begin{array}{l} \mathbf{w} \in \mathbf{W} \\ \mathbf{w} = (\mathbf{o}, \mathbf{d}) \end{array} \quad (23)$$

Then, let ψ^2 , $\bar{\psi}^2$, Q^2 and \bar{Q}^2 be defined as:

$$\begin{aligned} \psi^2 &= \sum_{\mathbf{w} \in \mathbf{W}} \sum_{\mathbf{a} \in \mathbf{A}} \psi_{\mathbf{a},\mathbf{w}}^2, & \bar{\psi}^2 &= \sum_{\mathbf{w} \in \mathbf{W}} \sum_{\mathbf{a} \in \mathbf{A}} \bar{\psi}_{\mathbf{a},\mathbf{w}}^2 \\ Q^2 &= \sum_{\mathbf{w} \in \mathbf{W}} Q_{\mathbf{w}}^2, & \bar{Q}^2 &= \sum_{\mathbf{w} \in \mathbf{W}} \bar{Q}_{\mathbf{w}}^2 \end{aligned} \quad (24)$$

6.3 The aggregated model

The aggregated problem will take into account the measured volumes of passengers captured by all the detection systems of type C, type B and type A.

To this end let the following squared error terms be defined as:

$$E_X^2 = \sum_{\pi=(\ell,s)} \sum_{\sigma \in S(\ell)} \sum_{a \in a_e(\pi,\sigma)} (X_a - v_a)^2 \quad (25)$$

$$\hat{E}_X^2 = \sum_{\pi=(\ell,s)} \sum_{\sigma \in S(\ell)} \sum_{a \in a_e(\pi,\sigma)} (\hat{X}_a - (v_a + \bar{v}_a))^2 \quad (26)$$

$$E_\alpha^2 = \sum_{\pi=(\ell,s)} \sum_{\sigma \in S(\ell)} \sum_{a \in a_y(\pi,\sigma)} (\alpha_a - v_a)^2 \quad (27)$$

$$\hat{E}_\alpha^2 = \sum_{\pi=(\ell,s)} \sum_{\sigma \in S(\ell)} \sum_{a \in a_y(\pi,\sigma)} (\hat{\alpha}_a - (v_a + \bar{v}_a))^2 \quad (28)$$

$$E_\beta^2 = \sum_{\pi=(\ell,s)} \sum_{\sigma \in S(\ell)} \sum_{a \in a_a(\pi,\sigma)} (\beta_a - v_a)^2 \quad (29)$$

$$\hat{E}_\beta^2 = \sum_{\pi=(\ell,s)} \sum_{\sigma \in S(\ell)} \sum_{a \in a_a(\pi,\sigma)} (\hat{\beta}_a - (v_a + \bar{v}_a))^2 \quad (30)$$

$$E_\theta^2 = \sum_{\pi \in \Pi^B} \sum_{i \in i(\pi)} \sum_{\mathbf{d} \in \mathbf{D}} (\theta_i - (g^{i,\mathbf{d}} + \bar{g}^{i,\mathbf{d}}))^2 \quad (31)$$

The quadratic error terms (25), (27) and (29) are due to observed volumes which result from aggregated measures made by the type A detection units. The quadratic error terms (26), (28) and (30) are due to observed volumes by type C detection units and the quadratic error term (31) is originated by errors made by type B detection units (first boarding counters). The term (31) also express that at nodes $i \in N_T$ corresponding to services carried out by type B units the total volume of first boardings θ_i comprises flows of type A and also flows type B, C and D (i.e., flows of passengers carrying devices with detectable electronic addresses and flows of passengers without these

devices). These nodes i will be origins of O-D flows $\omega = (i, \mathbf{d})$.

The previous quadratic error terms must be weighted accordingly to the different reliabilities of the detection systems. Let $\tau_X^2, \tau_\alpha^2, \tau_\beta^2$ the variances of the errors made by type A units in counting and calculating occupancies, alightings and boardings. Also, let $\hat{\tau}_X^2, \hat{\tau}_\alpha^2, \hat{\tau}_\beta^2$ the variances in counting occupancies, alightings and boardings for type C units. Let τ_ϕ^2 be the variance of the errors in detecting type A passengers by type A units and finally, τ_θ^2 the variance of the errors in counting the first boardings (type B units).

Also consider the following differentiable penalty function P in order to take into account line capacities as expressed in (3):

$$P = \sum_{\ell \in L} \sum_{s=1}^{n_\ell} \sum_{\substack{a \in A_e(\pi) \\ \pi = (\ell, s)}} \min^2\{0, c^\pi - (v_a + \bar{v}_a)\} \quad (32)$$

Consider the following quadratic problem (33) to (40):

$$\begin{aligned} \text{Min}_{\gamma, v, u, g} \quad & \tau_X^{-2} E_X^2 + \hat{\tau}_X^{-2} \hat{E}_X^2 + \tau_X^{-2} E_\alpha^2 + \hat{\tau}_X^{-2} \hat{E}_\alpha^2 + \\ & + \tau_X^{-2} E_\beta^2 + \hat{\tau}_X^{-2} \hat{E}_\beta^2 + \tau_\theta^{-2} E_\theta^2 + \\ & + \tau_\phi^{-2} \sum \phi_n^2 + \rho_p P + \\ & + \rho_v \psi^2 + \rho_v \bar{\psi}^2 + \rho_q Q^2 + \rho_q \bar{Q}^2 \\ \text{s.t. :} \end{aligned}$$

$$B\bar{v}^\omega = \bar{g}^\omega e^\omega, \bar{v}^\omega \geq, \bar{g}^\omega \geq 0, \omega = (i, \mathbf{d}) \in W \quad (33)$$

$$Bv^\omega = g^\omega e^\omega, v^\omega \geq, g^\omega \geq 0, \omega = (i, \mathbf{d}) \in W \quad (34)$$

$$v = \sum_{\omega \in W} v^\omega, \quad \bar{v} = \sum_{\omega \in W} \bar{v}^\omega \quad (35)$$

$$\gamma_h^{\mathbf{w}} \geq 0, \bar{\gamma}_h^{\mathbf{w}} \geq 0, h \in \mathcal{H}^{\mathbf{w}}, \mathbf{w} = (\mathbf{o}, \mathbf{d}) \in \mathbf{W} \quad (36)$$

$$C^n \begin{pmatrix} f^n \\ h^n \end{pmatrix} = \begin{pmatrix} \eta^n \\ \varphi^n \\ u^n \end{pmatrix}, \quad \begin{aligned} f^n, h^n &\geq 0 \\ \eta^n, \varphi^n &\geq 0, \\ \forall n \text{ burst} \end{aligned} \quad (37)$$

$$\varphi^{p(n)} = \eta^n, \quad \text{if } p(n) \neq \emptyset \quad (38)$$

$$\eta_0^n = \eta^n, \quad \text{if } p(n) = \emptyset \quad (39)$$

$$v[n] = u^n, \quad (40)$$

where $p(n)$ is the burst preceeding to burst n in the same service π . Constraints (33) link the type A flow vector v with the type A O-D flows matrix using the

node-link incidence matrix B for the diachronic network depicted in figure 1. Constraints (34) apply for flows different from type A flows (flows type B,C,D). Constraints (35) relate total link flows with flows per O-D (both, for type A flows and non-type A flows). Constraints (??) simply express the (obvious) relationship between flows per O-D pair and per destination flows. Constraints (32) are capacity limitations at links in A_e in the diachronic network (γ^π is the vehicle capacity of service π expressed in pax.). The penalty terms (20) and (21) appear affected by a penalty parameter ρ_v in the objective function and are the result of combining relationships (6) and (7) for flows of passengers with detectable electronic addresses and flows of passengers without them. Also, the penalty terms (22) and (23) in the objective function appear affected by a penalty parameter ρ_q and are the result from the combination of relationships (4) and (5), for O-D flows of passengers with detectable electronic devices and O-D flows of passengers without them. Constraints (37) express the relationships between internal flows within a burst n and the contour constraints between consecutive bursts during the operational period. Finally, in (40), vector flows on the time expanded network are linked to boundary flows u^n for a given burst n , where it is assumed that all bursts have the same size fixed by parameter r . Notice that a term $\phi_n^2 (= \phi_n^2(f^n, h^n, \eta^n, \varphi^n, u^n))$, appears in the objective function for each burst n in the time horizon.

$$u^n \triangleq (u_{b(r)}^n, \dots, u_{b(1)}^n, u_{y(r)}^n, \dots, u_{y(1)}^n) \quad (41)$$

6.4 A Lagrangian decomposition algorithm

Notice that if in previous problem (33) to (40), constraints (40) are dropped, then the problem splits into a) a large quadratic network flow problem in the variables v, g, \bar{v}, \bar{g} and, b) a series of small quadratic problems of the type (19), one for each burst n . Thus, the structure of the problem makes that an approximate solution can be seeked applying a Lagrangian Decomposition scheme based on the dualization of constraints (40). Another aspect that must be remarked is that solving the set of small quadratic subproblems (19), one for each burst, can be parallelized since they are mutually independent. Even for a large public transportation network, which would result in a relatively high number of bursts, carrying out sequentially the resolution of these problems would not be an issue, even taking into account that these computations should be done at each iteration of the dual lagrangian decomposition method. However, the problem's struc-

ture by itself would make to solve it in a parallel computing environment a natural choice.

On the other hand, the large network flow quadratic problem has the advantage that, could be conveniently reformulated in terms of per-destination flows at some stages of the solution algorithm (i.e., shortest paths computations). Moreover capacity constraints can be dealt with as penalties since that, due to its diachronic character, there would be no cycles in the network. Thus, known efficient vertex generating algorithms, such as for instance Hearn et al. (1987) could be applied to obtain a solution with acceptable quality and reasonable computational burden.

7 Conclusion

This paper presents a dynamic O/D matrix estimation model which may be aimed at an on-line usage, that incorporates emerging information and communication technologies (ICT), especially those based on the electronic signature detection of devices used by passengers on board that provide a rich source of data of higher quality than simple counts. The procedure presented is based on the statement of a mathematical programming-based model for the adjustment of passenger counts on a diachronic network model. The distinctive approach presented in the paper takes advantage of a detection system, capable of measuring not only the passenger's load on segments, but also to track, up to some extent, the passenger trajectories on the network, comprising but not limited to end-to-end trajectories. The estimation model is not limited to a single line but is a formulation for a general public transportation network. Although the details in this paper are specific for a system of bus lines, it can be extended without any loss of generality to rapid transit systems, metro systems or any public transportation system in which a passenger tracking system may be implemented. The formulation of the estimation model permits to include requirements on passenger's route choice which may be estimated either by historical or survey data or may come from static passenger transit assignment models. Without stating specifically a solution algorithm for the formulated problem, its characteristics are examined in order to apply advantageously a decomposition algorithm based on Lagrange relaxation, so that some of the resulting subproblems can be solved efficiently using specific non-linear network flow methods such as for instance vertex generating methods.

Acknowledgments. This research was supported by project grant TRA2016-76914-C3-1-P by the Spanish Ministerio de Economía, Industria y Competitividad.

References:

- [1] Barceló, J., Montero, L., Bullejos, M., Serch, O., Carmona, C., 2013. A Kalman filter approach for the estimation of time dependent od matrices exploiting bluetooth traffic data collection. *GITS Journal of Intelligent Transportation Systems* 17, 2, 123-141.
- [2] The Transport Data Revolution. Investigation into the data required to support and drive intelligent mobility. Catapult Transport Systems. March 2015.
- [3] Chu. X. A Guidebook for Using Automatic Passenger Counter Data for National Transit Database (NTD) Reporting. NCTR778-03, FDOT BDK85 977-04, Final Report 2010 NTCR University of South Florida.
- [4] Haupt, T., Nökel, K., Reiter, U., 2009. Schedule based dynamic transit assignment, in: Wilson, N., Nuzzolo, A. (Eds.), *Schedule-based modeling of transportation networks*. Springer.
- [5] Hearn, D. W., Lawphonpanich, S. and Ventura, J. A. (1987), *Restricted simplicial decomposition: computation and extensions*, *Mathematical Programming Study*, 31, 99-118.
- [6] Hua-ling, R. (2007) *Origin-Destination Demands Estimation in Congested Dynamic Transit Networks*. 2007 14th International Conference on Management Science & Engineering. August 20-22, 2007 Harbin, P.R.China
- [7] Jin, J.G., Teo, K.M., Odoni, A.R., 2016. Optimizing Bus Bridging Services in Response to Disruptions of Urban Transit Rail Networks. *Transportation Science* 50, 3, 790-804.
- [8] Kostakos, V., Camacho, T., Mantero, C., 2010. Wireless detection of end-to-end passenger trips on public transport buses, *Proceedings of the IEEE Conference on Intelligent Transportation Systems*, Funchal, Portugal. pp. 1795-1800.
- [9] Nguyen, S. and Pallotino, S.: *Equilibrium traffic assignment in large scale transit networks*, *European Journal of Operational Research* 37 (2) 176-186 (1988).
- [10] Nuzzolo, A., Crisalli, U. The schedule based approach in Dynamic Transit modelling: a general overview. In *Schedule-Based Dynamic Transit Modeling: Theory and Applications* Nigel H. M. Wilson, Agostino Nuzzolo, Eds. (2004).

- [11] Rucker, D. Automatic Passenger Counting Systems Expanding Beyond Just Counting People. 2003 Bus and Paratransit Conference Proceedings. Milwaukee, Wisconsin USA
- [12] Spiess, H.: Contribution à la théorie et aux outils de planification des réseaux de transport urbains. Ph D thesis, Département d'Informatique et Recherche Opérationnelle, Publication 382, CRT, U. de Montréal (1984)
- [13] Sun, L., Jin, J.G., Lee, D.H., Axhausen, K.W., Erath, A., 2015. Demand-driven timetable design for metro services. *Transportation Research Part C* 46, 284-299.
- [14] Tong, C. O. and Richardson, A. J. (1984) A computer model for finding the time-dependent minimum path in transit systems with fixed schedules. *Journal of Advanced Transportation* 18, 145-161.
- [15] TCRP synthesis 77. Passenger counting systems. Asynthesis of transit practice. Transportation Research Board of the National Academies. TRB 2008.
- [16] Wang, W., 2011. Bus Passenger Origin-Destination Estimation and Related Analyses Using Automated Data Collection Systems. *Journal of Public Transportation* 14, 4, 131-150.
- [17] Wang, W., 2015. Optimizing bus bridging services in response to disruptions of urban transit rail networks. *Transportation Science Articles in Advance*, doi:10.1287/trsc.2014.0577.
- [18] Wong, S., Tong, C., 1998. Optimizing bus bridging services in response to disruptions of urban transit rail networks. *Transportation Research-B* 32, 2, 35-48.
- [19] Xuesong, Z., Mahmassani, H.S., 2006. Optimizing bus bridging services in response to disruptions of urban transit rail networks. *IEEE Transactions on Intelligent Transportation Systems* 7, No 1, 105-114.