

Observability quantification of public transportation systems with heterogeneous data sources: An information-space projection approach based on discretized space-time network flow models

Jiangtao Liu

School of Sustainable Engineering and the Built Environment,
Arizona State University, Tempe, AZ, 85281, USA

Email: jliu215@asu.edu

Xuesong Zhou

School of Sustainable Engineering and the Built Environment,
Arizona State University, Tempe, AZ, 85281, USA

Email: xzhou74@asu.edu

Tel.: +1 480 9655827

(Corresponding Author)

Submitted for publication in Transportation Research Part B

Abstract

Focusing on how to quantify system observability in terms of different interested states, this paper proposes a modeling framework to systemically account for the multi-source sensor information in public transportation systems. By developing a system of linear equations and inequalities, an information space is generated based on the available data from heterogeneous sensor sources. Then, a number of projection functions are introduced to match the relation between the unique information space and different system states of interest, such as, the passenger flow/density on the platform or in the vehicle at specific time intervals, the path flow of each origin-destination pair, the earning collected from the tickets to different operation companies etc., in urban rail transit systems as our study object. Their corresponding observability represented by state estimate uncertainties is further quantified by calculating its maximum feasible state range in proposed space-time network flow models. All of proposed models are solved as linear programming models by Dantzig-Wolfe decomposition, and a k-shortest-path-based approximation approach is also proposed to solve our models in large-scale networks. Finally, numerical experiments are conducted to demonstrate our proposed methodology and algorithms.

Keywords

System observability quantification; Information space; Heterogeneous data sources; Public transportation system; Dantzig-Wolfe decomposition;

1. Introduction

The currently rapid innovations and developments of transportation system intelligence in multi-source sensing and information sharing continuously generate huge volumes of various data and information for planners and managers to better observe time-varying traffic conditions and accordingly propose adaptive travel demand management and supply (capacity) control strategies. However, the data sparsity problem still exists, because it is impossible to install fixed sensors on each link or to cover all links by point-to-point moving sensors. As a result, it surely requires new methodologies to recover the system-wide transportation conditions based on the limited observations (Zheng et al., 2014). It leads to a fundamental question; that is, how well the time-dependent transportation system states can be estimated or observed based on currently available heterogeneous source data. Further, it can be used to evaluate the value of current sensor networks and provides feedbacks to the future sensor network design, especially when current sensor data are incapable of offering enough key information for accurate system state estimates and also there are a number of different selections in sensor types, sensor amount and locations. In a broader extent, as shown in Fig.1, the system observability problem is the key to incorporate continuous multi-source sensor information to estimate different level of system states (flow, density, travel time), covering from link level, route level, OD pair level to even the whole network level, as the fundamental inputs for any active travel demand management and system supply resource allocation. Meanwhile, the feedback loop between system states and optimal design/control keeps moving forward along the time horizon with any external new disturbances, such as, weather, incidence, special events, new land use, population changes, etc.

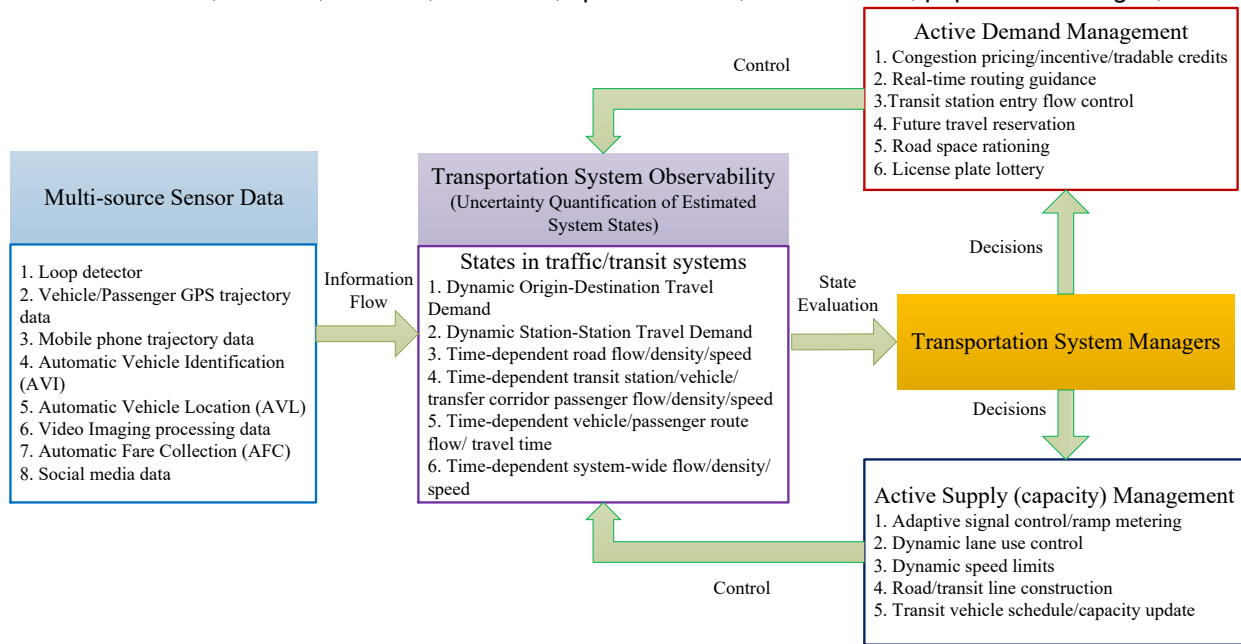


Fig.1 Transportation system management under multi-source sensor data

Daganzo (2007) pointed out some common challenges in deploying transportation network models. Particularly, requiring information on too many input parameters, such as dynamic origin-destination (OD) matrix and a wide range of dynamic traveler choice behavioral factors, could be major barriers for applying such models. However, the increasingly available multi-source big data and powerful computation capability are creating great opportunities to continuously provide accurate model inputs, capture individually complicated travel behaviors, and efficiently calculate the interested outputs. For example, the AFC data, mobile phone data and travel request mobile phone apps greatly improve the accuracy of observed OD travel demand, and the GPS-enabled devices enable us to record each passenger/vehicle's high-resolution trajectory in real-time ways.

Meanwhile, it should not be neglected that the overwhelming volume of data are also incurring new challenges on data use and transportation modeling. In the data use side: (i) Is the big data useful enough and what is the marginal value of the data? (ii) Under what goals, one kind of data is more useful than others'? (iii) How to fuse multi-source data to keep observation result consistency? On the other hand, in the transportation modeling side: (i) how to mathematically represent the available multi-source information

so that different system states can be estimated in a unified modeling framework; (ii) how to model the exact inner relation between the information and some interested system states, (iii) how to quantify the system observability (uncertainty of estimated states) for further optimal control and management, and (iv) how to design efficient and scalable algorithms for solving those models. Motivated by those general challenges stated above, this paper aims to explore the theoretical relation among sensor data, system states, and system observability in public transportation systems, especially taking the urban rail transit system as a starting point.

1.1 Observability in traffic systems

Observability is a concept introduced by Kalman for linear dynamic systems in control theory. It is a measure for how well internal states of a system can be inferred by knowledge of its external outputs. In other words, it aims to quantify or measure the uncertainty of estimated internal states based on the available external observations under a given sensor environment with sampling errors, sensor error and model errors. A comprehensive literature review can be found at the paper by [Castillo et al. \(2015\)](#). In the general transportation observability problems, a number of studies ([Castillo et al., 2007](#); [Castillo et al., 2008](#); [Gentili and Mirchandani, 2012](#)) modeled the problems as a system of linear equations and/or inequalities and then determine whether the system or one unknown variable is observable or not by analyzing the properties of its coefficient matrix. Meanwhile, in the system of linear inequalities, a general bound of unknown variables can be derived through the dual cone approach. However, those previous observability problems more cares about the list of variables to be observed rather than the specific system states uncertainty ranges.

As for evaluating the estimation uncertainty or accuracy, origin-destination (OD) trip matrix estimation is a widely studied classical problem due to its under-determination attribute, which means that there is an infinite number of OD trips that can generate link flows consistent with the observations. [Yang et al. \(1991\)](#) first introduced the concept of Maximum Possible Relative Error (MPRE) to theoretically investigate the estimation uncertainty and reliability of the OD estimated trips obtained by the entropy model. [Bianco et al. \(2001\)](#) further explored the accuracy of estimated OD matrix bound under different sensor location strategies. [Bierlaire \(2002\)](#) proposed the novel concept of total demand scale as a new measure to examine the quality of estimated OD trip tables from link counts, by maximizing/minimizing the total travel demand satisfying all observations. In addition, [Claudel et al. \(2009\)](#) fused observed probe and fixed sensor data to identify the range of highway travel times by maximizing and minimizing the total number of vehicles present on the highway at the initial condition in their proposed linear programs. Further, based on viability theory, [Claudel and Bayen \(2010, 2011\)](#) and [Canepa and Claudel \(2017\)](#) proposed theoretical approaches to address the Hamilton–Jacobi Equation in the Lighthill–Whitham–Richards (LWR) traffic flow problems with heterogeneous sensor data to estimate the uncertainty of different traffic system states. Note that there are also a number of studies focusing on traffic state estimation, attempting to find the most likely one by least square methods or others. We, however, aims to find the exact estimated state range using min/max methods based on available external multi-source sensor information.

1.2 State estimation and smart card applications in public transportation systems

In public transportation systems, the automatic fare collection system (AFC) or smart card provides much more information on trip information. A comprehensive literature review about smart card data use can be found in the papers ([Pelletier et al., 2011](#); [Ma et al., 2013](#)). [Trépanier et al. \(2007\)](#) estimate the alighting point for each passenger based on the smallest distance to the boarding stop of his/her next route from individually continuous riding records in smart card. [Seaborn et al. \(2009\)](#) proposed maximum elapsed time thresholds to identify transfers for bus-to-underground, underground-to-bus, and bus-to-bus to identify and assess multi-modal trips in London. Meanwhile, [Munizaga and Palma \(2012\)](#) estimated a multimodal transport OD matrix from smartcard and GPS data whiling consider unobserved trips by expansion factors in Santiago, Chile. [Yuan et al. \(2013\)](#) proposed a space alignment approach by aligning the monetary space and geospatial space with the temporal space to infer each passenger’s trajectory and the results improve the detection of uses’ home and work places. [Nassir et al. \(2015\)](#) applied the smart card data to detect activity and identify transfers to estimate the true origins and destinations. Under the situation that passenger’s boarding stop information is not recorded in smart cards, [Ma et al. \(2012\)](#) developed a Markov chain based Bayesian decision tree algorithm to estimate the sequential stops on the bus route and then match those stops with the recorded boarding time to infer passengers’ origin. Further, [Ma et al. \(2015\)](#)

improved their previous algorithms to increase the estimation accuracy and computation efficiency. In addition, in case that buses don't have Automatic Vehicle Location data, Zimmerman et al. (2011) developed a system named Tiramisu that can estimate and predict the real-time bus arrival time by applying the crowd-sourcing data from commuters sharing their GPS-enabled mobile phones.

In urban rail transit systems, a number of studies focus on the route choices and transfer patterns, which can be viewed as different system states required for estimation. Kusakabe et al. (2010) focused on the passengers' train choice behavior by assuming that each passenger aims to minimize the total waiting time at the departure station, loss time at the arrival station, and the transfer frequency. Zhao et al. (2007) chose the logit discrete choice model, but the tight side constraints (e.g. strict vehicle capacity constraint) are still hard to include. Ceapa et al. (2012) mined the regular spatial-temporal trip relations from AFC data to estimate and predict the crowding level for providing more accurate personalized trip planning services. Sun and Xu (2012) estimated the path choice based on the observed overall probability density of journey time and the derived distribution of individual path travel time from the rail transit smart card. Kusakabe and Asakura (2014) proposed a data fusion methodology to consider both the smart card data and person trip survey data by Bayes probabilistic model to estimate behavioral attributes of trips in the smart card data. Based on passenger OD matrix information and vehicle stop time and location data, Zhu et al. (2017a; 2017b) proposed probabilistic models to estimate the individual train loads, left behind probabilities, time-dependent crowding levels at stations under tight vehicle capacity considerations. In addition, Nair et al. (2013) focused on a large-scale bicycle sharing system and analyzed the connection between bicycle usage and public transit systems. Recently, Shang et al. (2019) fused the passenger trip time information from smart card and passenger counts at key locations in urban rail transit systems to estimate the passenger flow by proposing a space-time-state hyper network.

1.3 Potential contributions and structure of this paper

Most recent studies mainly focused on most likely system state estimation rather than system observability quantification, so this paper aims to develop a modeling framework capable of incorporating multi-source sensor data to address the observability quantification problem in public transportation systems. The contributions of our work are specifically listed as follows.

(i) The concept of information space is first adapted from the control theory to represent the multi-source sensor data and physical public transportation system by a system of linear equations and inequality constraints. (ii) Different system states are then connected by mapping the unique information space with the corresponding state projection functions, so the solution for quantifying the uncertainty of one state can be also used for other interested states. (iii) This paper presents a first set of analytical results for studying the observability problem in public transportation systems, and our proposed well-structured space-time network flow models can be decomposed as a number of subproblems by Dantzig-Wolfe decomposition for large-scale networks.

The remainder of this paper is organized in the following manner. The following section conceptually illustrates the general relation between information and system states. Section 3 shows how to construct a space-time network to model the transportation systems, how to generate the information space based on available multi-source information and why Dantzig-Wolfe decomposition is selected to solve our space-time network flow models. The state definitions and observability quantification are provided in Section 4, which also describes the general process of Dantzig-Wolfe algorithms for our specific models with considering measurement errors. Finally, numerical experiments are performed to demonstrate our proposed methodology in Section 5.

2. Conceptual illustration

Through applying some concepts from game theory and control theory into our problem (LaValle, 2012), a state space is defined as all possible internal system state based on the external physical transportation world, and an information space is a place where the internal states live when available external information is involved. As shown in Fig. 2, the information space is formed by the available information, and the states are tightly connected by different projection functions, which mathematically define the states according to the managers' needs. The bound among all possible states represents the state uncertainty to reflect system observability under current available information. When the information space is generated as one single point, it can state that the system is observable; otherwise, a non-empty set or space leads to unobservable or partially observable system. In this section, we aim to build the connection between the

internal states and the external states (observation or information) by information space as a bridge or communication channel, and further quantify the corresponding uncertainty of states defined by users.

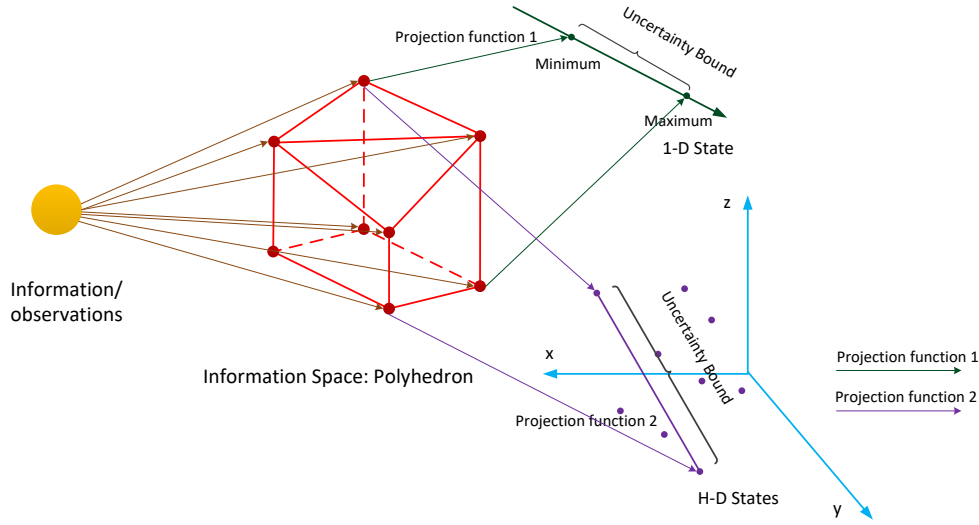


Fig. 2 Relation among information, information space, and flexible states

For illustrative purposes, Fig. 3(a) depicts a simple transportation network with four nodes and five links. The link travel time and capacity are also provided as physical network attributes. Let x_1 , x_2 and x_3 represent the path flow on paths 1, 2 and 3. Based on the tight capacity constraints, the following relation can be obtained: $0 \leq x_1 \leq 2$, $0 \leq x_2 \leq 3$, and $0 \leq x_3 \leq 1$, which defines the system state space shown as a blue cuboid in Fig. 3(b).

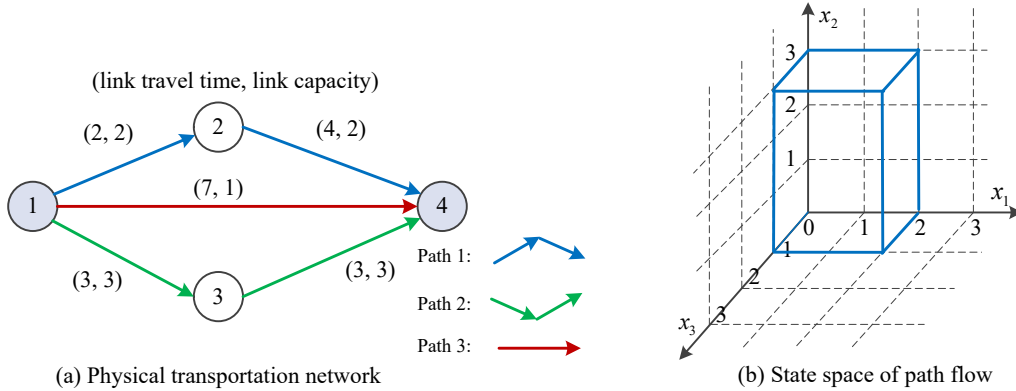


Fig. 3 An illustrative transportation network and its state space

Assume that the OD information is available through survey that there are four vehicles departing from node 1 to node 4. Then, one corresponding constraint will be $x_1 + x_2 + x_3 = 4$. Fig. 4(a) displays the information space as the intersection of the red triangle and the blue cuboid based on the available OD information. Two scenarios are further designed as follows to analyze the relation between system states and available information.

Scenario 1: Assume that there is one flow count detector on link (2, 4) and its link count is 1. The relation gets updated as follows: $1 + x_2 + x_3 = 4$, $0 \leq x_2 \leq 3$, and $0 \leq x_3 \leq 1$, so the corresponding information space is reduced to be the intersection of the red triangle and the green rectangle shown in Fig. 4(b).

Scenario 2: Suppose that the automatic vehicle identification (AVI) detectors are available at nodes 1 and 4. One vehicle's travel time is observed as 7min. Since only path 3's travel time is 7 and its capacity is just 1, it implies that path flow $x_3 = 1$. As a result, the relation changes as follows: $x_1 + x_2 + 1 = 4$, $0 \leq x_1 \leq 2$, and $0 \leq x_2 \leq 3$. The corresponding information space becomes the intersection of the purple rectangle and the red triangle displayed in Fig. 4(c).

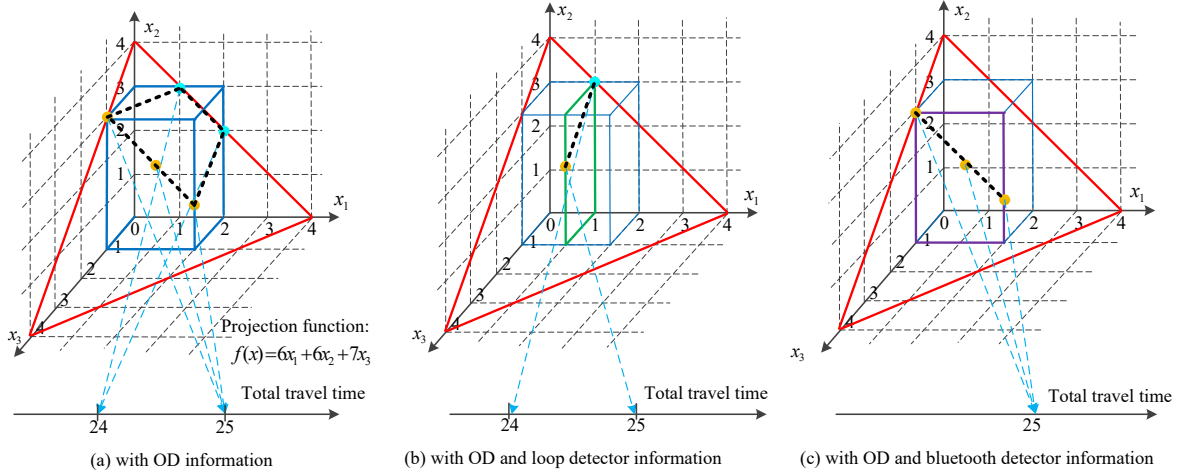


Fig. 4 Information spaces and its projected bound under different available information

As shown in Fig. 4, the information spaces are generated as polyhedrons based on different available information. A projection function is defined to map the information space into one dimension state (total travel time). It should be remarked that, the projection functions could be defined based on different goals interested by the system managers, such as, the total system travel time, the number of vehicles in one area, etc. In Fig. 4(a), one projection function is defined as $f(x) = 6x_1 + 6x_2 + 7x_3$, which means that the total system travel time is the analysis goal. Accordingly, different optimization models are solved by maximizing and minimizing $f(x)$ subject to different information spaces.

As demonstrated in Fig. 4(a), the bound of total travel time formed by projection is [24,25]. When the link count data are added in Fig. 4(b), the information space is reduced, but the projected bound is still the same. It indicates that the new information from link count doesn't contribute to reduce the uncertainty of this state estimation (or increase the observability of this state), even though a smaller information space is generated in Fig. 4(b). It reminds us that the value of "big data" is determined by not only "big volume" but also its usefulness of information to the system.

In Fig. 4(c), the point-to-point Bluetooth data (end-to-end passenger id detector data) makes the projected bound converged to be one unique value, which implies that the point-to-point data is more powerful than the point detector for increasing the observability of system travel time in this case. Therefore, evaluating the values of different information should be based on which states the system manager really cares. One information that seems worthless for one goal may be much useful for other state estimation applications. Moreover, the information space in Fig. 4(c) is larger than that in Fig. 4(b), but the uncertainty of state estimate in Fig. 4(c) is 0 and less than that in Fig. 4(b). Thus, the volume of information space might not be the best criteria to judge the bound of state estimate uncertainty or system observability. Similarly, the network scale may not be the best criteria, either.

Except for information from physical sensors, the previous travel experiences or currently published traffic information from transportation agencies could also take important roles in quantifying observability. For example, if everyone has perfect information over the network attributes based on his/her experiences and intends to find the best route for his/her trip, which is usually entitled as Wardrop's first principle, the information space will be redefined as, $x_1 + x_2 = 4, 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 3$. Compared with these above two scenarios, the information space is further changed by this assumed and (potentially questionable) travel behavior. Therefore, one accurate travel behavior model could provide much different information to determine system states from the information space perspective.

3 Problem Statement

Table 1 lists the general indices, sets, parameters and variables in our proposed models appeared in Sections 3 and 4.

Table 1 Indices, sets, parameters and variables

Indices	Definition
i, j	Index of nodes, $i, j \in N$
(i, j)	Index of physical link between two adjacent nodes, $(i, j) \in L$

a	Index of passenger group, $a \in A$
$o(a)$	Index of origin node of group a
$d(a)$	Index of destination node of group a
t, s	Index of time intervals in the space-time network
τ	Index of time period for the observed passenger flow
p	Index of paths, $p \in P$
r	Index of transit companies
Sets	
N	Set of nodes in the physical transit network
L	Set of links in the physical transit network
A	Set of passenger groups
V	Set of vertices in the space-time network
E	Set of edges/arcs in the space-time network
G	Set of time period for the observed passenger flows
$S_{p,a}$	Set of paths p of group a
$G(i, j, \tau)$	Set of arcs on observed link (i, j) at time period τ
Parameters	
β_1, β_2	The weights on target passengers' trip time and observed link/arc flows, respectively
μ_a	The observed aggregated average trip time of group a from smart card data
$\mu_{i,j,\tau}$	The observed aggregated total passenger count on link (i, j) during time period τ
w_p	The travel time of path p
c_p^r	The earning collected on path p of transit company r
$Cap_{i,j,t,s}$	Capacity of traveling arc (i, j, t, s) in the space-time network
DT^a	The departure time of group a
AT^a	The assumed arrival time of group a
D_a	The number of passengers in group a
$c_{i,j,t,s}$	Travel cost of traveling arc (i, j, t, s) in the space-time network
T	The time horizon in the space-time network
$\delta_{(i,j,t,s)}^{p,a}$	Path-link incidence index of route p of group a on arc (i, j, t, s)
w^p	The path travel time of path p
Variables	
$x_{i,j,t,s}^a$	The number of passengers in group a is assigned on traveling/waiting arc (i, j, t, s) in the space-time network
$\theta_a, \theta_{i,j,\tau}$	Continuous positive deviation variables for group a 's trip time and link (i, j) during time period τ , respectively
x_a^p	The number of passengers of group a choosing their feasible path p
μ_a^*	The preprocessed aggregated trip time of group a from smart card data
$\mu_{i,j,\tau}^*$	The preprocessed aggregated passenger count on link (i, j) during time period τ

3.1 Space-time Network Construction in Public Transit Systems

To properly account for the evolution of system dynamics over time, [Ford and Fulkerson \(1958\)](#) first introduced dynamic network flow models to solve the dynamic maximum flow problem in time extended networks. The space-time network flow models are then widely used in dynamic transportation systems, such as, dynamic system optimal with a point queue model ([Zawack and Thompson, 1987](#)), dynamic user equilibrium with a spatial queue model ([Drissi-Kaitouni and Hameda-Benchekroun, 1992](#)), dynamic system optimal with departure time, route choice and congestion toll ([Yang and Meng, 1998](#)), dynamic user equilibrium with link travel time functions ([Chen and Hsueh, 1998](#)), and activity-based dynamic user equilibrium ([Lam and Yin, 2001](#)). Recently, in order to maximize network accessibility, [Tong et al. \(2015\)](#) proposed a space-time network flow model with binary decision variables, which actually derives a number of agent-based models in space-time networks later.

There are a number of studies providing how to construct specific time-expanded networks for different transportation systems, such as, freeway network ([Lu et al., 2016](#)), road network with signal settings ([Li et al., 2016](#)), urban transit network ([Liu and Zhou, 2016](#)), bike-sharing network ([Lu, 2016](#)), road network with

activity requests (Liu et al., 2018), and vehicle trajectory network (Wei et al., 2017). In this section, we consider a physical urban rail transit network with a set of nodes (stops/stations) N and a set of links L as a starting point. Each link can be denoted as a directed link (i, j) from upstream node i to downstream node j . A deterministic transit schedule is supposed to be obtained from Automatic Vehicle Location (AVL) data from vehicle tracking systems. Then, we construct a space-time network, where V is the set of vertices and E is the set of arcs. Node i is extended to a set of vertices (i, t) at each time interval t in the study horizon, $t = 1, 2, \dots, T$, where T is the length of the optimization horizon. The transit schedule from node i to node j from time t to time s can be represented by a travelling arc (i, j, t, s) where $(s - t)$ is the exact scheduled/running link travel time and should be integer multipliers of one time interval. The capacity of travelling arcs can be viewed as the transit vehicle's carrying capacity. In addition, a waiting arc is built from (i, t) to $(i, t + 1)$ at node i with waiting time of 1 unit of time and its capacity is defined as the station/platform storage capacity.

In urban rail transit systems, individual passengers should have a trip record with origin, departure time, destination and arrival time from the smart card. However, transit agencies may just provide aggregated trip data for groups of passengers to protect the passengers' privacy. Each group a with D_a passengers has a departure time DT^a at origin node $o(a)$ to its destination node $d(a)$. At each destination node, there is one assumed large arrival time T for all groups so that the following proposed model will be one-origin-one-destination problem in the space-time network. It should be noted that the travel cost of waiting arcs on the destination node of each passenger group is 0, which means that once the passengers in a group arrive at the destination, the waiting cost to the super-destination (at larger arrival time T) is 0. Finally, the estimated trip time in the model should be equal to the observed trip time, which will be presented in the following sections.

In addition, one transfer node can be divided as multiple nodes, depending on how many transit lines intersect at this node. One illustrative example is shown in Fig. 5(a) where two lines intersect at node 2 and make it as a transfer station. Then node 2 is split to node 2' and node 2'' and the modified physical network is shown in Fig. 5(b). The travel time of transfer links could be the actual walking time, and its capacity is the maximum passenger throughput at transfer corridors. As a remark, based on the maximum transfer distance accepted by passengers, it is possible to connect different stops by transfer links or extended to multimodal networks. Fig. 5(c) shows the transfer process where all transfer time is assumed to be 1 unit of time. In addition, it is feasible to consider the uncertainty of walking time on transfer links or from station entry to the platform through constructing more service/travelling arcs with different arc travel times. Furthermore, in traffic networks, the road can provide its service at each time interval with a specific arc capacity, and the given signal timing rules whether those service arcs are open or closed, as shown in Fig. 5(d) to show our unified modeling framework.

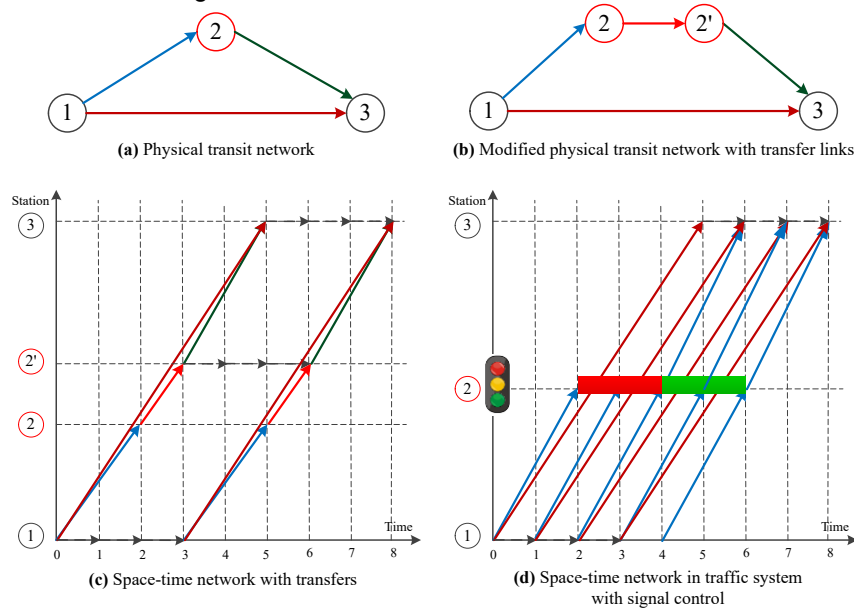


Fig. 5 space-time network construction

3.2 Information space generation based on multi-source sensor data

Information space (LaValle, 2012) works as a communication channel to connect the external physical world and the internal system states. The external physical world is sensed by heterogeneous sensors in terms of different observations or information, which finally forms a corresponding information space. Meanwhile, the internal system states are reflected through the information space based on the specific state definitions. In addition to the physical transit lines and schedules as useful information, smart card in the urban transit systems can provide origin, boarding time, destination, and/or alighting time for each person or each passenger group. With the development of sensing technologies, more available sensor information from big data applications are being used in the public transportation systems, including aggregated passenger flow from video data, cellphone/GPS based point-to-point trajectory data, travel behavior data (e.g. preference) through survey, trip related information from social media.

The specific modeling on generating information space based on those available sensor data is developed as follows. It should be remarked that, the first-in-first-out (FIFO) rule is not incorporated in our proposed space-time network, because it can be violated in transit networks and the multi-source information can better present travelers' decision by reducing the feasible information space. Also, readers who are interested in FIFO in space-time networks can be referred to the details in the paper (Shang et al., 2018).

Taking the rail transit system as our modeling example, we have the following formulation.

(i) According to the physical network, transit schedule, and dynamic OD information from smart card data, the standard flow balance constraint can be given as follows:

$$\sum_{i,t:(i,j,t,s) \in E} x_{i,j,t,s}^a - \sum_{j,t:(j,i,s,t) \in E} x_{j,i,s,t}^a = \begin{cases} -D_a & \forall a, j = o(a), s = DT^a \\ D_a & \forall a, j = d(a), s = T \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

(ii) Strict vehicle and station platform capacity constraint

$$\sum_a x_{i,j,t,s}^a \leq Cap_{i,j,t,s}, \quad \forall (i, j, t, s) \in A \quad (2)$$

(iii) The estimated trip time of each group in the model should be consistent with the observation (average trip time of each group) from smart card.

$$\sum_{(i,j,t,s)} (x_{i,j,t,s}^a \times c_{i,j,t,s}) = D_a \times \mu_a, \quad \forall a \quad (3)$$

(iv) Estimated aggregated passenger flow count on link (i, j) during time period τ is expected to be the observation from video data or counting by person.

$$\sum_a \sum_{t \in \tau} x_{i,j,t,s}^a = \mu_{i,j,\tau}, \quad \forall (i, j, \tau) \quad (4)$$

(v) Non-negative arc flow variables

$$x_{i,j,t,s}^a \geq 0 \quad (5)$$

Compared with the existing space-time network models with tight arc capacity constraints and agent-based bounded rationality constraint, the constraints on observed trip journey time and observed passenger flow counts are new and specific for this observability problem. It should be pointed out that the concept of capacity defined as the maximum number of passengers the facility can take is not easy to be calibrated and has its subjectivity. However, Automated Passenger Counter (APC) used for accurately recording passengers boarding and alighting activities (Furth et al., 2006) can be greatly helpful to capture the capacity value under saturated conditions when passengers have to wait for the next available vehicle due to the in-vehicle congestion.

In addition, in practice, the transportation agencies usually have different levels of network models. When they focus on the regional analysis, all input elements in the whole regional area should be considered. On the other hand, when performing a subarea analysis, it requires to estimate the route behavior of all passengers related to the subarea, so the vehicle capacity will be adjusted to exclude the number of external passengers who utilize vehicles in the subarea. In another way without adjusting capacity, we can build external nodes connected with the subarea, and those external nodes represent the origins/destinations of aggregated external passengers. As a result, there will be four categories of OD pairs for the subarea (Zhou et al., 2006): external-internal OD, internal-external OD, internal-internal OD, and external-external OD where its passengers may pass through this subarea. Then, it needs to first solve an OD trip demand estimation problem for further transportation system observability, which will be considered in our future research.

Note that, if a passenger is regarded as a group of passengers, $x_{i,j,t,s}^a$ will be a binary variable and the above modeling is still available, which is transformed in an agent-based model. Usually, agent-based

trajectory data can provide more point-to-point travel time information rather than just path choice. The above formulation presents an arc-based formulation for constructing the information space. In comparison, a path-based formulation based on the feasible path generation is offered as follows:

(i) Flow balance constraint:

$$\sum_p x_a^p = D_a, \forall a \quad (6)$$

(ii) Capacity constraint:

$$\sum_{(p,a) \in S(p,a)} (\delta_{(i,j,t,s)}^{p,a} \times x_a^p) \leq Cap_{i,j,t,s}, \forall (i,j,t,s) \in A \quad (7)$$

(iii) Trip time constraint:

$$\sum_p x_a^p * w^p = D_a \times \mu_a, \forall a \quad (8)$$

(iv) Aggregated passenger flow count constraint:

$$\sum_a \sum_{t \in \tau} (\delta_{(i,j,t,s)}^{p,a} \times x_a^p) = \mu_{i,j,\tau}, \forall (i,j,\tau) \quad (9)$$

(v) Non-negative path flow:

$$x_a^p \geq 0 \quad (10)$$

Given the time-expanded space-time network constructed in advance, the general feasible path set for a passenger group with specific OD pair and a departure time can be generated by a forward label correcting algorithm from the vertex (origin and departure time) to its destination node based on the observed trip time of that group as a prism. In addition, we assume that the sensor data are currently measured in a perfect way without errors and noise. In section 4.1, we will address the issue of measurement errors, so the revised measurement representation could replace the observations in the models above.

It should be remarked that when considering the bus transit systems, the smart card data usually only have the origin and departure time without passengers' destination and arrival time information. To model this condition that dynamic OD trips are unknown, D_a will be a variable in equation (1) and the summation of D_a with same origin and departure time should be equal to the recorded total trip generation at this origin and departure time from smart card data. If the structure of each OD pair with departure time is given based on the historical OD information, the number of unknown OD variables will be greatly reduced. In addition, from the data mining perspective, interesting readers can be referred to the paper (Ma et al., 2013) to find studies related to estimating destination probabilities.

3.3 Dantzig-Wolfe decomposition for special flow-balance blocks

As shown in the arc-based and path-based formulation in section 3.2, the flow balance constraint is a special block that can be solved by classical shortest path algorithms and further be incorporated by Dantzig-Wolfe decomposition. Actually, this kind of methods had been adopted for static traffic assignment (Larsson and Patriksson, 1992), side constrained traffic equilibrium (Larsson et al., 2004), time constrained shortest path problem (Desrosiers and Lubbecke, 2005), etc. The advantage of this decomposition allows us to solve the special blocks in parallel via independent computation threads to address large-scale networks, especially when the computer hardware has a rapid development in current days. It also has the re-optimization capability if the travel demand, arc performance function or network topology has any changes in future (Larsson and Patriksson, 1992).

Specifically, Dantzig-Wolfe decomposition is originally proposed for solving linear programming problems with special structure. A general primal linear program can be represented as: $\min c^T x$, subject to, $Ax \leq b$, $Dx \leq d$, and $x \geq 0$. According to Minkowski-Weyl's Theorem, given the convex set $X = \{x \in \mathbb{R}^n | Ax \leq b\}$ where $Ax \leq b$ is a **special block**, X can be represented by the extreme points and extreme rays of X : $X = \{x = \sum_i \lambda_i x^i + \sum_j \mu_j y^j | \sum_i \lambda_i = 1, \lambda_i \geq 0, \mu_j \geq 0\}$. When X is a bounded polyhedron, X can be represented by the extreme points, $X = \{x = \sum_i \lambda_i x^i | \sum_i \lambda_i = 1, \lambda_i \geq 0\}$.

Substituting the expression above to the original model leads to the following Master Problem:

$$\min \sum_i c^T \lambda_i x^i \quad (11)$$

$$\text{Subject to, } \sum_i D \lambda_i x^i \leq d, \sum_i \lambda_i = 1 \text{ and } \lambda_i \geq 0 \quad (12)$$

Suppose that a subset of extreme points P is available. The Restricted Master Problem (**RMP**) can be obtained by $\min \sum_{i \in P} c^T \lambda_i x^i$, subject to, $\sum_{i \in P} D \lambda_i x^i \leq d$, $\sum_{i \in P} \lambda_i = 1$ and $\lambda_{i \in P} \geq 0$. Assume that λ^* and (π, ω) is the optimal and dual solutions to the RMP, respectively. The reduced cost is defined as $\gamma(x) = c^T x - \pi^T Ax - \omega$. Then, we solve the subproblem: $\min c^T x - \pi^T Ax - \omega$, subject to $Ax \leq b$ and $x \geq 0$. If the reduced cost is non-negative, the solution is optimal; otherwise, the solution can be viewed as a new extreme point and added to the RMP until the reduced cost is non-negative.

With different objective functions related to different estimated states, our proposed models in section 4.2 based on the generated information space in Section 3.2, will be solved under the framework of Dantzig–Wolfe decomposition in section 4.2. Specifically, based on the flow-balance constraint, the flow on a particular path (or path flow for a passenger group a) can represent one extreme point. A path flow uniquely corresponds to its path, so a particular path implicitly indicates a specific extreme point. This enables us to express the arc flow of group a on arc (i, j, t, s) as $x_{i,j,t,s}^a = \sum_h (\delta_{(i,j,t,s)}^{p(h),a} \times x_a^{p(h)} \times \lambda_{(a,h)})$, where $x_a^{p(h)} = D_a$ for each generated extreme point h , and $\sum_{h \in H(a)} \lambda_{(a,h)} = 1$. Since variable $x_{i,j,t,s}^a$ is continuous rather than discrete, it should be a continuous combination of extreme points and $\lambda_{(h,a)} \geq 0$ according to Minkowski-Weyl's Theorem. On the other hand, the link flow vector of each group $x_{i,j,t,s}^a$ can also be seen as one extreme point, as it is the result of one specific path flow vector. In addition, since the model is a linear program, it is possible to return an unbounded solution or non-unique solution, which has been studied in some classical linear programming literature (Dantzig, 1998).

4. Observability quantification of different states under heterogeneous data sources

The system observability reflected by state uncertainty mainly arises from two sources: one is lack of useful information, which results in the many-to-one mapping between the many possible system states and one partial observation, and the other is the possible measurement error due to the noise and disturbance in sensing systems. This section will address the measurement error issues and further quantify the uncertainty of state estimates.

4.1 Data preprocessing for measurement errors

The analyses on small card data (Trépanier et al., 2007; Barry et al., 2009) show that the data must be thoroughly validated and corrected prior to the practical use. Therefore, it might happen that no feasible solution exists when the observed data are directly used in built models due to its measurement errors. In addition, even each observation is tested in the model and can provide feasible solutions, but it is still possible to have infeasible solutions when heterogeneous sensor data are considered simultaneously, because the inconsistency among different kinds of sensor data may still exist. Hence, this situation leads to the measurement estimation problem, which aims to obtain estimations as close as possible to the corresponding measurements under real-world physical constraints. It is also recognized that there are other different approaches to clean and verify those measurements in advance. The approach adopted in this paper is the generalized least squares. Based on the proposed constraints in section 3.2, a nonlinear estimation model is presented as follows.

$$\text{Min } \beta_1 \sum_a (\sum_{(i,j,t,s)} (x_{i,j,t,s}^a \times c_{i,j,t,s}) - D_a \times \mu_a)^2 + \beta_2 \sum_{(i,j,\tau)} (\sum_a \sum_{t \in \tau} x_{i,j,t,s}^a - \mu_{i,j,\tau})^2 \quad (13)$$

Subject to constraints (1), (2) and (5).

The objective function is to minimize the weighted total deviations between the estimated and observed measurements, where β_1 and β_2 are the weights reflecting different degrees of confidence on observations. Those weights can be viewed as the inverses of the variances of the heterogeneous sources of measurements adopted by Lu et al. (2013).

Another technique used to measure the deviation is to quantify the absolute difference as least absolute deviations (LAD). The corresponding objective function will change to be

$$\text{Min } \beta_1 \sum_a |\sum_{(i,j,t,s)} (x_{i,j,t,s}^a \times c_{i,j,t,s}) - D_a \times \mu_a| + \beta_2 \sum_{(i,j,\tau)} |\sum_a \sum_{t \in \tau} x_{i,j,t,s}^a - \mu_{i,j,\tau}| \quad (14)$$

Least absolute deviation treats all observation equally, but least squares gives more emphasis to large residuals by squaring the residuals, which could be a better choice when dealing with outliers in which estimated values are far from real-world sensor observations. Note that the least absolute deviations can be solved by linear programming through transforming the model. For example, the new model based on objective function (14) will be minimizing $\beta_1 \sum_a \theta_a + \beta_2 \sum_{(i,j,\tau)} \theta_{i,j,\tau}$ while adding new constraints $-\theta_a \leq \sum_{(i,j,t,s)} (x_{i,j,t,s}^a \times c_{i,j,t,s}) - D_a \times \mu_a \leq \theta_a$, $-\theta_{i,j,\tau} \leq \sum_a \sum_{t \in \tau} x_{i,j,t,s}^a - \mu_{i,j,\tau} \leq \theta_{i,j,\tau}$, and $\theta_a \geq 0, \theta_{i,j,\tau} \geq 0$. By assuming that the sensor has 1% relative errors θ_a as the worst case, Canepa and Claudel (2017) added the sensor errors in the hard constraints in their optimization problems. However, in our paper we assume that the ground-truth data will be closest to the sensor data while surely satisfying the physical constraints in transportation systems, which is similar to the data reconciliation problem. Therefore, the measurement values μ_a and $\mu_{i,j,\tau}$ in information space generation and uncertainty quantification in section 3.2 need to be replaced by μ_a^* and $\mu_{i,j,\tau}^*$ from our proposed preprocessing model.

The model is a linearly constrained quadratic programming model. Frank-Wolfe algorithm is usually used to solve the optimization problem where the objective function is convex differentiable real-valued function and the feasible region of side constraints is compact convex. Frank-Wolfe algorithm is explained at Appendix A in detail.

4.2 Projection-function-based observability quantification

As illustrated in Section 3.2, the generated information space can work as a channel to connect available observations with different states. This section will propose different projection functions as the mapping between the feasible information space and specific transit system states. Table 2 introduces our focused states.

Table 2 Focused states and motivations

Focused states	Motivations
(1) Arc flow/density state : passenger density on station platforms, in vehicles, and transfer corridors	(i) identify possible dangerous spots for safety; (ii) make decisions on vehicle updates, line/timetable changes and stop location adjustment
(2) Path flow state : the number of passengers taking one specific line segment	(i) clear the total ticket fare to each company based on the service they provide; (ii) evaluate the current liquidation policy and quantify the unreasonable income bound for each company
(3) Path flow state : the path flow range of each time-dependent OD pair	(i) compare or verify the traditional logit route choice model for better understanding travel behavior
(4) Network-level arc flow/density state : network-level time-dependent passenger flow/density states on several key stations/vehicles	(i) distribute the network-level transit condition and intelligent passenger trip guidance (ii) evaluate network-level control and policy

The focused states and its uncertainty quantification are listed as follows.

Projection function 1 for arc flow state: the number of passengers on one specific arc (i, j, t, s) (station platform, vehicle, transfer corridor) in the space-time network is represented as $\sum_a x_{i,j,t,s}^a$, so the arc flow uncertainty can be quantified by maximizing and minimizing $\sum_a x_{i,j,t,s}^a$, subject to constraints (1) to (5).

Projection function 2 for path flow state 1: the earnings that one transit company r can obtain is represented as $\sum_a \sum_p (x_a^p \times c_p^r)$, where c_p^r is the income of using the segment in company r 's operation area of path p . It can be calculated as a parameter in advance based on the ticket price and segment and path distance. Therefore, the earning bound is estimated by maximizing and minimizing $\sum_a \sum_p (x_a^p \times c_p^r)$ subject to constraints (6) to (10).

Projection function 3 for path flow state 2: the flow rate on path p is $\sum_a x_a^p$, so the uncertainty bound of path flow is measured by maximizing and minimizing $\sum_a x_a^p$, subject to constraints (6) to (10).

Projection function 4 for network-level arc flow state: the passenger flow (density) states on key station platforms at one time index (e.g., at 7:30am) is a high-dimensional vector $\{q(i, t)\}$ where i is one of the key stations. For one specific station i , $q(i, t) = \sum_a x_{i,j,t,s}^a$ is the number of passengers at station i at time t . Since the state is not one dimension anymore, the concept of the Maximal Possible Relative Error (MPRE) first introduced by Yang et al. (1991) is adopted to quantify the state uncertainty of high-dimensional variables. As shown in Fig. 6(a), the state solution (vector $\{x_{i,j,t,s}^a\}$) based on different projection functions for one-dimensional state above is one feasible solution in the information space, so each solution (vector $\{x_{i,j,t,s}^a\}$) can be mapped to high-dimensional states to generate new state points (vector $q(i, t)$) illustrated in Fig. 6(b), which are used as sample points to approximately obtain the MPRE. Specifically, we need to calculate the average relative error between any two points, and find the maximal one as the MPRE.

For example, the average relative error between point 1 and point 2 is calculated as follows (Yang et al., 1991), where $q_1(i, t)$ and $q_2(i, t)$ are m -dimensional vectors recording m stations' passenger flow at time t . The relative deviation between point 1 and point 2 is $\lambda_{(1,2,t)} = \frac{q_1(i,t) - q_2(i,t)}{q_1(i,t)}$ and the average relative

deviation $AV(\lambda_{(1,2,t)}) = \sqrt{\frac{\phi(\lambda_{(1,2,t)})}{m}}$, where $\phi(\lambda_{(1,2,t)}) = \sum_{i=1}^m \lambda_{(1,2,t,i)}^2$ and $\lambda_{(1,2,t)} = \{\lambda_{(1,2,1,t)}, \lambda_{(1,2,2,t)}, \dots, \lambda_{(1,2,m,t)}\}$.

In addition, Yang et al. (1991) defined the concept of Estimation Reliability as a measure about the state uncertainty; that is, $Re = \frac{1}{1+AV(\lambda)}$, which shows that when the $AV(\lambda)$ is 0, the reliability of the estimated state is 1. In contrast, when $AV(\lambda)$ tends to infinity, there is almost no reliability guarantee. Therefore, this index is used to measure the possible estimated flow range rather than the specific flow rate. The result is just based on some sample points, so it is still an approximation approach.

System observability is just the first step for future system controllability related to sensor network design and system optimal control. Therefore, the goal of this research is to inform planners whether the observations from currently available sensors are enough to well observe the system. In other words, if the state uncertainty range in some key locations is pretty large, it means that the current sensor information is not enough and the planners should add more useful sensors to better know what is happening in the system.

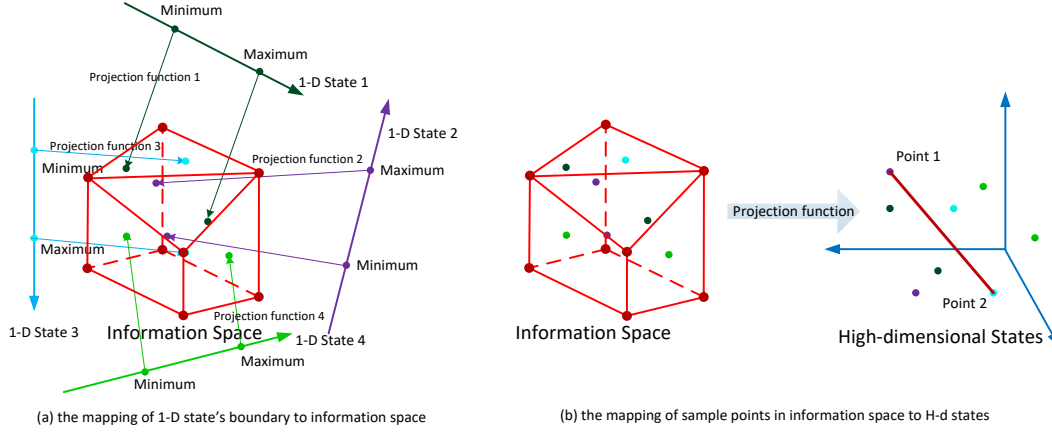


Fig. 6 Relation of information space and different types of state

4.3 The solution procedure of Dantzig-Wolfe decomposition

From the perspective of Dantzig-Wolfe decomposition, based on the master problem in formulas (11) and (12), x^i as one extreme point can be replaced by variable vector $x_{i,j,t,s}^a$ and variable x_a^p for arc-based and path-based models above, respectively, and c^T is the corresponding cost on each variable.

Taking minimizing the flow on arc (i, j, t, s) as an example, the general procedure of the algorithm is described as follows.

Step 1: Initialization. Find one feasible passenger arc flow vector $\{x_{i,j,t,s}^{a,k}\}$ on the shortest path as the k^{th} extreme point for each passenger group a . It indicates that $(k - 1)$ extreme points have been generated before finding one feasible solution, which will be explained after Step 3 as a remark.

Step 2: Solve the restricted master problem to obtain the duals of side constraints.

$$\text{Min } \sum_k \sum_a (x_{i,j,t,s}^{a,k} \times \lambda_{a,k}) \quad (15)$$

Subject to,

$$\sum_k \sum_a (x_{i,j,t,s}^{a,k} \times \lambda_{a,k}) \leq \text{Cap}_{i,j,t,s}, \quad \forall (i, j, t, s) \in A \quad (16)$$

$$\sum_k \sum_{(i,j,t,s)} [(x_{i,j,t,s}^{a,k} \times \lambda_{a,k}) \times c_{i,j,t,s}] = D_a \times \mu_a, \quad \forall a \quad (17)$$

$$\sum_k \sum_a \sum_{t \in \tau} (x_{i,j,t,s}^{a,k} \times \lambda_{a,k}) = \mu_{i,j,\tau}, \quad \forall (i, j, \tau) \quad (18)$$

$$\sum_k \lambda_{a,k} = 1, \quad \forall a \quad (19)$$

$$\lambda_{a,k} \geq 0 \quad (20)$$

$\pi_{i,j,t,s}$, π_a , $\pi_{i,j,\tau}$ and ω_a are the duals of side constraints (16)-(19), respectively.

Step 3: Solve each sub-problem as a time-dependent shortest path problem to calculate its reduced cost for each passenger group, which can be implemented by parallel computing techniques, such as, Multi Process Interface (MPI).

Sub-problem for each passenger group a :

$$\text{Min } (c_{i,j,t,s}^a \times x_{i,j,t,s}^{a,k+1}) - \sum_{(i,j,t,s)} (\pi_{i,j,t,s} \times x_{i,j,t,s}^{a,k+1}) - \pi_a \times \sum_{(i,j,t,s)} (c_{i,j,t,s} \times x_{i,j,t,s}^{a,k+1}) - \sum_{(i,j,\tau)} (\pi_{i,j,\tau} \times \sum_{t \in \tau} x_{i,j,t,s}^{a,k+1}) - \omega_a \quad (21)$$

Subject to

$$\sum_{i,t:(i,j,t,s) \in E} x_{i,j,t,s}^{a,k+1} - \sum_{i,t:(j,i,s,t) \in E} x_{j,i,s,t}^{a,k+1} = \begin{cases} -D_a & j = o(a), s = DT^a \\ D_a & j = d(a), s = T \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

Actually, the reduced cost is $(c_{i,j,t,s}^a \times x_{i,j,t,s}^{a,k+1}) - \sum_{(i,j,t,s)} (\pi_{i,j,t,s} \times x_{i,j,t,s}^{a,k+1}) - \pi_a \times \sum_{(i,j,t,s)} (c_{i,j,t,s} \times x_{i,j,t,s}^{a,k+1}) - \sum_{(i,j,t)} (\pi_{i,j,t} \times \sum_{t \in \tau} x_{i,j,t,s}^{a,k+1}) - \omega_a$. If it is negative, we add the solution of the sub-problem to the restricted master problem at step 2 and begin next iteration. When the reduced costs of all sub-problems are non-negative, the optimal solution is achieved.

In the initialization step, in order to find one feasible solution from the shortest path problem as initial extreme points, we can introduce artificial variables for those coupling constraints and solve the problem by the Dantzig-Wolfe decomposition again (Kalvelagen, 2003). For the maximum problem, we can transform it as a minimum problem by changing the positive arc costs to be negative. Since there is no circle in the space-time network along the time dimension, the label correcting algorithm can always be used to find the shortest path.

In addition, if the k -best space-time paths can be found for each passenger group in advance, those k paths can be treated as k extreme points, and then we only need to solve a restricted master problem to obtain an approximated solution. This approximation approach can greatly reduce the complexity of this problem and is applicable for large-scale networks where the solvers don't have the capability to directly solve the original models. The detailed k -shortest space-time path generation will be explained in section 5.2.

4.4 Discussions on real-time state uncertainty quantification

The uncertainty of real-time system state increases the difficulty of real-time state prediction and optimal control. Compared with the offline state observability in this paper, the challenges in the real-time condition include (i) the real-time rail transit OD travel information is not available and (ii) the state transition along the time is highly required.

(i) Real-time OD demand estimation: Based on day-to-day historical and accurate dynamic OD demands in urban rail transit systems, we can classify k representatives $OD_{o,d,\tau}^k$ for each OD pair at different time periods, so the estimated real-time OD demand is $OD_{o,d,\tau} = \sum_k (w_k \times OD_{o,d,\tau}^k)$, where w_k is a binary variable, which indicates that only one OD candidate k will be chosen. As a result, the dynamic OD travel demand's spatial structure can be well captured, compared with those OD estimation models which mainly optimize one departure time profile for all or one-class total static OD trips. In addition, the real-time trip generation at each station/origin with departure time is available from the smart card data, so $\sum_d OD_{o,d,\tau} = OD_{o,\tau}^{obs}$ provides more information to generate the real-time information space.

(ii) Real-time state transition: the rolling horizon approach has been widely chosen for real-time transportation operations and control (Peeta and Mahmassani, 1995; Zhou and Mahmassani, 2007; Meng and Zhou, 2011). Under this mechanism, when focusing on one time period, it needs a look-back period and a look-ahead period, because the generated passengers from the look-back period could still be in the network during our focused time period, and in the look-ahead period all passengers can arrive at their destination for our network modeling. Along the planning time horizon, once some trips are finished at our focused time period, their true OD information can be obtained in real time, so the corresponding estimated OD trips can be replaced by the real ones, which can also reduce the information space for our state observability quantification.

5. Experiments

5.1 Tests in a hypothetical network

This section will demonstrate the proposed models and algorithms in Sections 4 and implement them in a general purpose optimization package GAMS. All source codes can be downloaded at the website: https://www.researchgate.net/publication/326020738_Observability_Scenarios_1-4. The experiments are performed in the following transit network shown in Fig. 7(a), where seven urban rail lines exist in the transit systems. In order to model the passenger count observation at transfer corridors, specific transfer links are built as shown in Fig. 7(b).

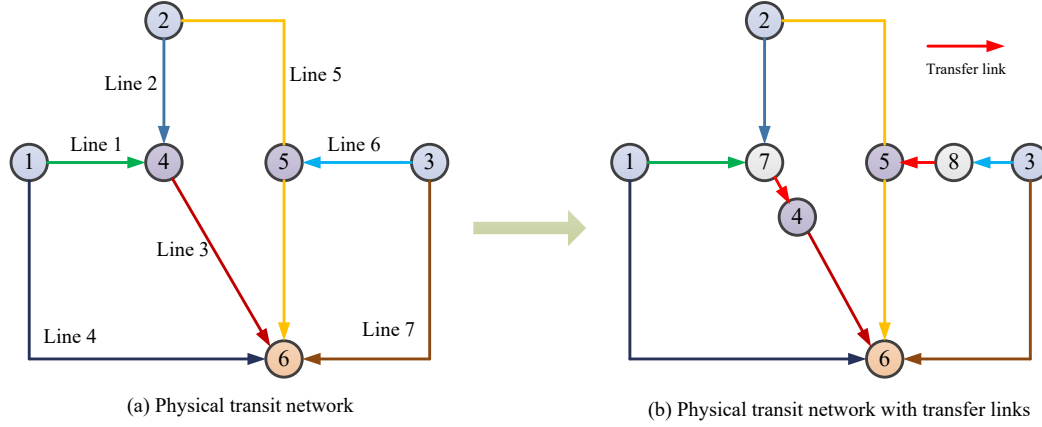


Fig. 7 Hypothetic urban rail transit network

5.1.1 Given multi-source sensor data

- (1) Table 3 lists the existing transit service arcs based on the given timetable of the seven transit lines.
- (2) The origin, destination, departure time and aggregated trip time of each passenger group are listed in Table 4, and each group represents 100 passengers in this test.
- (3) The vehicle capacity of each line is assumed in Table 5, where it can be observed that the capacity of rail transit vehicles could have its adjustment at different time periods by increasing or decreasing the number of train units.
- (4) The passenger count data from video processed data at transfer corridor (7, 4) is available; that is, 450 and 810 passengers are observed at time points 3 and 6.

Table 3 Hypothetic transit service arcs lists

Service Arc	Start Time	End Time	Service Arc	Start Time	End Time
(1,7)	0	3	(1,7)	3	6
(7,4)	3	4	(7,4)	6	7
(4,6)	4	6	(4,6)	7	9
(1,6)	0	8	(1,6)	3	11
(2,7)	0	3	(2,7)	3	6
(2,5)	0	4	(2,5)	3	7
(5,6)	4	7	(5,6)	7	10
(3,8)	0	3	(3,8)	3	6
(8,5)	3	4	(8,5)	6	7
(3,6)	0	8	(3,6)	3	11

Table 4 Trip attributes of each passenger group

Group No	OD Pair	Departure Time	Average Trip Time	Group No	OD Pair	Departure Time	Average Trip Time
1	1 → 6	0	6	15	1 → 6	3	7.5
2	1 → 6	0	7	16	1 → 6	3	7
3	1 → 6	0	8	17	1 → 6	3	8
4	1 → 6	0	6.5	18	2 → 6	3	6
5	2 → 6	0	7	19	2 → 6	3	7
6	2 → 6	0	7.5	20	2 → 6	3	6.5
7	2 → 6	0	6.5	21	2 → 6	3	7.5
8	2 → 6	0	6	22	2 → 6	3	8
9	3 → 6	0	7	23	2 → 6	3	6.8
10	3 → 6	0	7.5	24	3 → 6	3	7
11	3 → 6	0	8	25	3 → 6	3	7.5

12	1 → 6	3	6	26	3 → 6	3	7.4
13	1 → 6	3	7	27	3 → 6	3	7.8
14	1 → 6	3	6.5	28	3 → 6	3	8

Table 5 Vehicle capacity of transit lines

Line No	L1	L2	L3	L4	L5	L6	L7
Capacity of vehicles departing at time 0	300	300	600	200	400	300	200
Capacity of vehicles departing at time 3	400	400	800	300	600	400	300

5.1.2 Focused states for observability quantification

The states focused in this experiment are listed as follows.

- (1) Arc flow state: passenger count (congestion) in transfer corridor (8, 5) at time points 3 and 6, respectively.
- (2) Path flow state 1: the passenger flow departing at node 2 and time 0 to use line 1.
- (3) Path flow state 2: the earning collected in the ticket for company line 1 on its first vehicle.
- (4) Network-level arc flow state: the system-wide passenger count (congestion) on the running vehicles at time point 5.

5.1.3 Scenario design

As a short summary, based on the available supply and demand data, we aim to (i) preprocess the measurements as data reconciliation in step 1, and (ii) quantify the uncertainty of our focused states in step 2. Five scenarios are designed to demonstrate the value of information based on our proposed models.

Scenario 1 (S1: base case): the origin, destination, and departure time of each passenger group is given, and no other information is available.

Scenario 2 (S2: base case + count): based on scenario 1, the passenger count data from video processed data at transfer corridor (7, 4) is available.

Scenario 3 (S3: base case + end-to-end travel time): based on scenario 1, the averaged group trip time from smart card is available.

Scenario 4 (S4: base case + end-to-end travel time + count): based on scenario 1, both the passenger count data and average group trip time data are available.

Scenario 5 (S5: ground truth): since the observed data may have its measurement errors, we assume that a ground truth can be obtained and will be compared with other scenarios. The ground truth is assumed as the system conditions based on maximizing the arc flow at time point 3 in scenario 3.

5.1.4 Result analysis

In step 1, the measurement is preprocessed by the proposed model at section 4.1. In step 2, we compute the uncertainty range of states (1)-(3) by maximizing and minimizing the state goals, and state (4) is addressed based on the solutions from the previous three states as a sample-based approximation. Before analyzing different state results in different scenarios, it is important to clearly illustrate the conditions under which those results are obtained from our proposed models.

(1) In scenario 1, there is no available sensor data, so the measurement doesn't need to be preprocessed.

(2) In scenario 2, the measurement is preprocessed for the passenger count data at transfer corridor (7, 4). The estimated passenger counts at transfer corridor (7, 4) at time points 3 and 6 is 450 and 800, respectively, compared with the observed values of 450 and 810. The total absolute error for the observed passenger count is 10.

(3) In scenario 3, the estimated average group trip time for each group is shown in Table 6. The total absolute error for the average group trip time is 2.58.

(4) In scenario 4, in step 1, there are two different sensor data, so it will require weights on different measurements. As discussed by [Lu et al. \(2013\)](#), the weights should reflect the degrees of confidence on different observed data and can be represented by the inverses of the variances of the distinct sources of measurements. The total absolute errors for observed average group trip time and passenger count are 3.83 and 273, which are greater than the absolute errors in scenario 1 and scenario 2, respectively. It indicates that the inconsistency among multi-source data forces the model to find a balance among those observations.

(5) In scenario 5, the preprocessed group trip time in step 1 is used as the input to maximize the passenger count in transfer corridor (8,5) at time points 3, and the corresponding system condition is assumed as the ground truth in this dynamic transit system.

Table 6 The observed and preprocessed average group trip time for each passenger group

Passenger group No	Observed values	Preprocessed values in scenario 3	Passenger group No	Observed values	Preprocessed values in scenario 3
1	6	6	15	7.5	7.5
2	7	7	16	7	7
3	8	8	17	8	8
4	6.5	6.5	18	6	6
5	7	7	19	7	6.9
6	7.5	7	20	6.5	6.4
7	6.5	6.5	21	7.5	7
8	6	6	22	8	7
9	7	7	23	6.8	6.7
10	7.5	7.5	24	7	7.08
11	8	8	25	7.5	7.57
12	6	6	26	7.4	7.47
13	7	7	27	7.8	7.87
14	6.5	6.5	28	8	8

Fig. 8 shows that the estimated maximum and minimal flow rates on each focused arc under different scenarios. As available information is increased, the range of passenger flow uncertainty on transfer corridor (8, 5) is reduced. Meanwhile, both scenarios 3 and 4 can assert that their estimated state uncertainty is 0 and the state is completely observable. However, the different estimated unique states on arc (8,5,6,7) seem conflicted.

Specifically, in scenario 3, the observed trip time is preprocessed due to its measurement error, and finally the estimated states on transfer corridor (8, 5) is consistent with the states in the ground truth in scenario 5. Note that the estimated states may not be totally consistent with the ground truth, even though the observed data is same as the corresponding data in ground truth, because the observation is only a partial reflection of the whole system condition. It is also possible that the corrected measurement is not consistent with that in this ground truth if other measurement correction approaches rather than the least square method are used in reality in step 1.

In addition, in scenario 4, the inconsistency of observed link count data and observed trip time data makes the corrected measurement different from the corresponding data in the ground truth, so the final estimated unique state in step 2 cannot be the real-world condition anymore. Therefore, in reality, when the transportation system state is estimated by different sensor data, the data quality and assigned weight on each data source in step 1 is important and should be clearly stated. Note that how to balance each observation is beyond the scope of this paper. For more details on knowledge fusion, readers can be referred to the paper (Zheng et al., 2014).

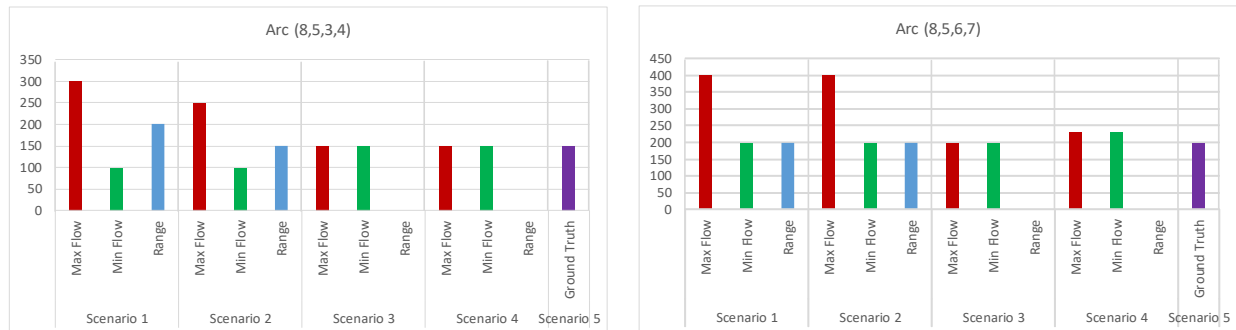


Fig. 8 The estimated flow uncertainty range on each focused arc

Focusing on the passenger flow departing at node 2 and time 0 to use line 1, it is actually the path flow of path (2,0) → (5,4) → (6,7). The path flow uncertainty is shown in Fig. 9. The uncertainty range is similar to the arc flow above. The estimated unique state in scenario 4 is not consistent with the state value in

ground truth. In addition, if line 1 is managed by one company and other lines are managed by other companies, it needs to assign the fare to each company based on their service. However, the number of passengers using one specific line is uncertain in the transit system, so based on our proposed method, we can quantify the uncertainty and estimate the general fare earning for each company rather than just using some simple rules for fare clearing (Gao et al., 2011; Zhou, 2014). For example, one previously simple rule is to just calculate the shortest path and then assume that passengers will choose the shortest path as their selected lines.

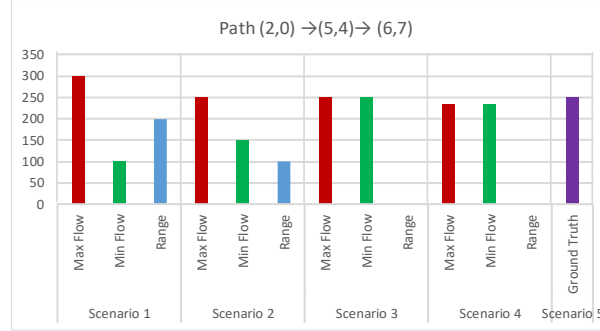


Fig. 9 Estimated flow uncertainty range on the focused path

In each scenario, we maximize and minimize the passenger flow on arc (8,5,3,4), arc (8,5,6,7), and path of line 1 as 6 cases, respectively, so six feasible solutions of $x_{i,j,t,s}^a$ can be obtained as sample points to quantify our defined system-level state observability. For state 4, the system-wide passenger count (congestion) on the running vehicles at time point 5 is be represented by the passenger flow on arcs (1,6,3,11), (1,6,0,8), (1,7,3,6), (2,5,3,7), (2,7,3,6), (3,6,3,11), (3,6,0,8), (3,8,3,6), (4,6,4,6), and (5,6,4,7). Based on the definitions of Maximal Possible Relative Error (MPRE) and Estimation Reliability (Re), the values of MPRE and Re are 16.38 and 5.753%, respectively.

In scenario 2, with passenger count information, the corresponding values of MPRE and Re are 0.267 and 78.93%, respectively. It shows that the estimation reliability gets significantly improved when passenger counts from one key location (transfer corridor) are available, which could avoid a large uncertainty range occurring in scenario 1. In Scenarios 3 and 4, the values of MPRE and Re are 0 and 100%, respectively, but it is still emphasized that the MRPR and Re should be clearly explained with its correspondingly different measurement preprocessing errors (assigned weights) and adopted approach.

In addition, as discussed in section 4.1, one other approach to address the sensor errors is to check the possible relative errors of sensors, and then incorporate it in the hard observation constraints instead of performing the preprocess step. In order to observe the sensitivity of relative sensor errors, a number of experiments are conducted in the following. The ground truth trip time is the corrected data in Table 6. The uncertainty ranges of passenger flows on arc (8,5,3,4) are shown in Fig. 10 under different relative sensor error assumptions. With the increase of relative errors of all trip time data, the uncertainty range correspondingly increase as expected. In addition, even the relative error of all trip time sensors reaches 5%, its estimation uncertainty is still smaller than the result of scenario 2 where only observed count data are available. Therefore, it is important not only to choose high-quality sensors, but also to select the right sensor types under different deployment contexts, for improving specific state observability.

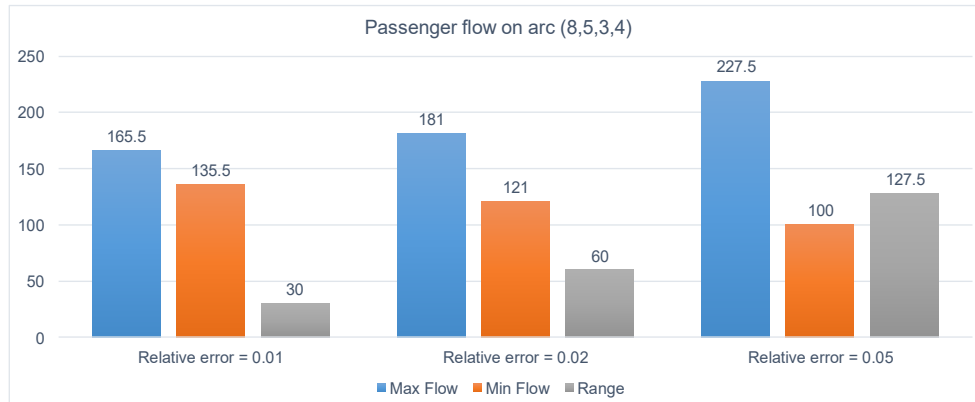


Fig. 10 Sensitivity analysis of relative errors of trip time sensor

5.1.5 Results from Frank-Wolfe algorithm and Dantzig-Wolfe decomposition

In this section, we implement Frank-Wolfe algorithm in step 1 and Dantzig-Wolfe decomposition algorithm in step 2 in GAMS. The case of minimizing the passenger flow on arc (8,5,3,4) in scenario 4 is treated as an example to analyze the performance of those algorithms. The source code can be downloaded at this link (https://www.researchgate.net/publication/324809217_F-W_and_D-W_Observability_Quantification).

In section 5.1.4, the case is solved by the solver MINOS in GAMS directly. In step 1, the solved model is a non-linear programming model, and the minimal total generalized least square error in the objective function is 5.968. When the model is solved by Frank-Wolfe algorithm as a linear programming model, the result finally converges to 7.069 after 20 iterations. The gap is probably caused by the optimal step size, which is found as a constant value at each iteration rather than a constant value vector for each variable. The discussions on improving Frank-Wolfe algorithm can be found in previous literature on static traffic assignment (Fukushima, 1984). That will also be our future research to find better ways to solve quadratic programming models, such as, Alternating Direction Method of Multipliers (ADMM).

In step 2, Dantzig-Wolfe decomposition is applied to generate extreme points for time-dependent OD pairs as subproblems, and the restricted master problem is solved by CPLEX. The solved minimal passenger flow is 142.5 based on the preprocessed measurements by Frank-Wolfe algorithm rather than by the NLP solver. However, if the preprocessed measurements in step 1 are directly obtained from the NLP solver, the final minimal passenger count on arc (8,5,3,4) from the Dantzig-Wolfe decomposition is 150 as well.

In addition, the total computational time of using Dantzig-Wolfe decomposition is 8 s and the average computation time for each subproblem of each passenger group is just 0.1 s. Meanwhile, it needs 0.4 s to directly solve the original problem. Since the subproblem in Dantzig-Wolfe decomposition is independent for each passenger group and has less constraints and variables, its advantage will be reflected when the solvers cannot directly solve the original problem in large-scale networks, which will be demonstrated in next section. In addition, if we focus on the real-time observability and controllability in future, especially in large-scale networks, the values of model complexity reduction and computation efficiency will be further highly respected.

5.2. Tests in a large-scale network

In this section, the public Google Transit Feed Specification (GTFS) data from Alexandria Transit Company in 2015 is used as our tested large-scale transit network (<https://transitfeeds.com/p/alexandria-transit-company>). As shown in Fig.11, it has 12 routes, 1638 trips (866 trips on weekdays, 423 trips on Saturdays, 261 trips on Sundays, and 88 trips on the Christmas day), and 629 stops. In this experiment, the scheduled trips on weekdays are only considered as the provided schedule.



Fig. 11. Alexandria transit network read from GTFS, in Virginia, USA

In order to obtain the time-dependent transit travel demand, we map the traffic analysis zones in the city of Alexandria to the transit network as the activity locations. As a result, 42 OD pairs are matched. In

addition, the time period of 6:00am to 9:00am is divided by 36 time intervals, so the time-dependent OD demand is defined by each 5 minutes. Finally, 1485 time-dependent OD pairs are obtained based on the arc generation rules above. Then, 32,029 vertexes and 713,650 arcs are generated in the corresponding space-time network for one whole weekday. This space-time network includes 636 nodes (stops, passenger origin and destination nodes), and its operation time ranges from 6:00am (time minute interval 360) to 10:20am (time minute interval 620). The arcs include vehicle running arcs, passengers' walking arcs from origin to transit stops and from transit stops to destination, transfer arcs, and waiting arcs. The space-time arc generation rules contain (i) the trip (path) travel time is less than 120 minutes; (ii) the maximum number of transfer times is 3; (iii) the maximum transfer/walking time is 30 minutes; (iv) the maximum transfer/walking distance is 0.5 mile.

For illustrating our modeling and algorithms and testing its computation efficiency in large-scale networks, we have the following assumptions on our data. (i) All transit vehicle capacity is 35 and the walking, waiting and transfer arc capacity is 9999. (ii) The time-dependent demand of each OD pair is assumed to be 1, which means that one passenger will arrive every 5 minutes for each OD pair. (iii) The observed average trip time of each passenger group is generated as a random value between the minimal and the maximal path costs of the 3-shorest paths of its OD pair, and those observed data are accurate and don't need the pre-processing step. However, it should be pointed out that the required real-world data should have the following considerations: (i) The origin and destination of transit demand should be more accurate instead of using the centroid of traffic analysis zones, (ii) the transit OD demand is time-dependent and location-dependent rather than our assumed same value, and (iii) the transit vehicles are also line-dependent and need to be carefully calibrated as mentioned in section 3.2.

Our focused state is the uncertainties of passenger flow count on transfer links from stop 4000644 to stop 4000863 (internal nodes from 370 to 553) and from stop 4000745 to stop 4000509 (internal nodes from 447 to 290) during 6am to 9am. The source codes for the following experiments can be found at https://www.researchgate.net/publication/332545172_Experiments_on_large_scale_networks_k-best_path_based_observability

5.2.1 Computation results of original models solved by standard solver and Dantzig-Wolfe decomposition

As developed in section 4.2, the original model is an arc-based linear programming model, which could be directly solved by standard solvers. Two scenarios are solved by CPLEX in GAMS on one workstation with intel i7@2.80 GHz, 40 threads and 192GB RAM, and its corresponding results are compared in Table 7. Since the variable is arc flow $x(a, i, j, t, s)$ for each passenger group a on arc (i, j, t, s) , the number of original variables is $a \times i \times j \times t \times s$. In scenario 1, this number is $2 \times 636 \times 636 \times 80 \times 80$, around 5.18 billion. Even the matrix of constraints is very sparse, it will take more than 10 min to generate and solve the linear programming model in scenario 1. If we further consider more than 50 passenger groups, it will be beyond the capability of the standard solvers to handle this problem. When the problem in scenario 1 is decomposed by Dantzig-Wolfe decomposition, each passenger group will be solved independently in each subproblem to make the model solvable in large-scale networks. The computational time for solving each subproblem by CPLEX in GAMS is about 4 min 42 s.

Table 7 Model statistics of two scenarios

Model statistics	Scenario 1	Scenario 2
Number of passenger groups	2	5
Number of nodes	636	636
Focused time period (min to min)	360-440	360-440
Number of original variables $x(a, i, j, t, s)$	5.18 billion	12.9 billion
Number of constraints	102,000	255,000
Computation time (s)	10 min 15 s	25 min 21 s

5.2.2 Approximation-based path-based solution

As shown in section 5.2.1, Dantzig-Wolfe decomposition can decompose the model as a number of relatively easy solvable subproblems, which provide the current shortest space-time path as extreme point for each passenger group at each iteration, but the multiple iteration process may take a long computation time and has a high requirement on computer memory used for arc-based models in large-scale networks. Therefore, we will propose an approximation-based approach, which generates k-shortest paths as extreme points for each passenger group (in each OD pair with time-dependent departure time) in advance rather than using Dantzig-Wolfe decomposition to generate extreme point iteration by iteration. That is similar to solving a restricted master problem based on the k extreme points in Dantzig-Wolfe decomposition.

The k-shortest path algorithm for each passenger group (each time-dependent OD pair) is implemented by C++ as follows.

(i) Based on the origin vertex (origin node and departure time) in the space-time network, the label correcting algorithm is used to generate a shortest path tree from the origin vertex to all possible vertexes selected on the basis of the space-time arc generation rules shown at the beginning of section 5.2.

(ii) According to the physical destination location, we can find a number of candidate vertexes (stop id and stop time in schedule) connected to the physical destination node by walking arcs. Then we can add the label costs of those candidate vertexes and its corresponding walking arc costs to the physical destination, so the destination will have a number of vertexes (destination node and arrival time) with different label cost.

(iii) Sort those label costs of the destination node and select k least-cost destination vertexes and back trace to the origin vertex. As a remark, at each vertex, we also record the transfer state (the number of transfer times from its origin vertex), so when back tracing the path to origin vertex, we can obtain different paths which are from one same vertex with same label cost but with different transfer states. Finally, the k -shortest path set can be generated for each time-dependent OD pair.

Compared with the iteratively generated extreme points (shortest space-time paths) to reach the optimal solution in Dantzig-Wolfe decomposition, this approximation method may have no feasible solution or optimal solution if enough paths are not generated. Specifically, if a network has heavy congestion (such as, the Beijing transit systems in peak hours) due to the interaction of high travel demand and limited vehicle carrying capacity, passengers may have to wait for a long time and cannot get on their successively desired vehicles. This kind of endogenous system congestions could affect the selection of “ k ” in our k -best path-based approximation to have feasible solutions. Therefore, we also analyze the impacts of the selection of different “ k ” on the system observability in the following experiments. On the other hand, this path-based approximation could greatly reduce the number of variables in the model and significantly improve the computation efficiency in transit systems where the number of alternative space-time paths for each passenger group is usually limited.

Table 8 lists the model statistics of four scenarios to calculate the uncertainties of passenger flow count on transfer links from stop 4000644 to stop 4000863 (internal nodes from 370 to 553). It is observed that the optimal solution can be reached with the increase of the number of candidate paths generated in k -shortest path algorithm. As the corresponding feasible constraint space is enlarged from scenarios 1 to 4, the maximal passenger flow count increases and the minimal value decreases until scenario 3, which indicates that 5-shortest paths generated for each passenger group may be used to find the optimal solution directly. The total computation time is also greatly reduced compared with solving the arc-based formulation directly in GAMS. In addition, the maximal flow count, the minimal flow count, and the uncertainty range on link from stop 4000745 to stop 4000509 (internal nodes from 447 to 290) in 4 scenarios are shown in Fig. 12

Table 8. Model statistics of four scenarios on link from stop 4000644 to stop 4000863

Model Statistics	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Number of passenger groups (OD pair with departure time)	1485	1485	1485	1485
Focused time period (min to min)	360-620	360-620	360-620	360-620
k-shortest paths of each passenger group	k=3	k=4	k=5	k=6
computation time for generating k paths in C++ (sec)	48	48	48	48
maximal passenger flow count solved in GAMS	26	26.85	27.12	27.12
minimal passenger flow count solved in GAMS	16.32	15	15	15

uncertainty of passenger flow counts on this transfer link	9.68	11.85	12.12	12.12
number of variables x_a^p	4455	5940	7425	8910
number of constraints	12124	12862	13561	14162
computation time for solving the master problem in GAMS (sec)	24	24	24	24
total computation time (sec)	72	72	72	72

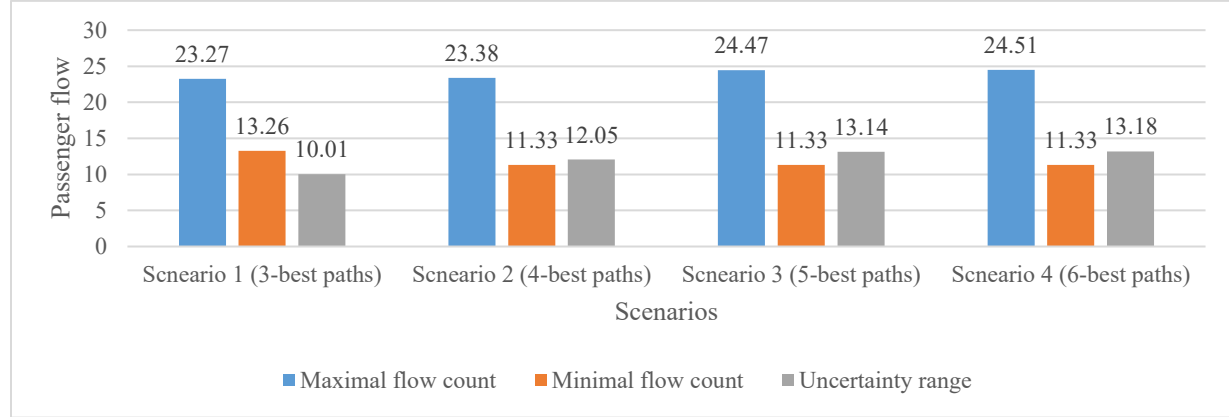


Fig. 12 Uncertainties of passenger flow count at transfer links from stop 4000745 to stop 4000509

7. Conclusion and future research

This research provides insights about the relationship between multi-source information, information space, state estimation, and system observability quantification by taking the urban transit systems as the analysis object. The information space and information errors are highly respected for state observability. Projection-function-based approaches are presented to quantify the observability of different states under the same information space. The proposed models can explain that the value of information highly relies on its aimed specific system states and sensor location rather than its volume. It provides an analysis to show how to better use multi-source information for different state estimate evaluation and how to design the sensor network with measurement errors (Zhou and List, 2006; Xing et al., 2013; Zhu et al, 2018; Wu et al., 2018) for future system state observability improvement.

The observability quantification based on different states is just the first step for better observing and controlling the system. The following questions are currently under our considerations for future research: (1) what is the balance among the system observability, the minimally needed information, and the required accuracy of future controls? (2) What is the balance of the sensor data cost, value of information and its computational efficiency in proposed models and algorithms? (3) How to integrate the heterogeneous sensor network design with the real-time system control? (4) How to use multi-source data to better model travel behavior and further improve system observability (Wu et al., 2018)? (5) How to visualize the real-time uncertainty of different system states in a straightforward way for the public?

Acknowledge

We would like to thank Dr. Yingyan Lou at Arizona State University and Dr. Jianrui Miao at Beijing Jiaotong University for their great comments and helps on our experiments. This paper is supported by National Science Foundation – United States under Grant No. CMMI 1538105 “Collaborative Research: Improving Spatial Observability of Dynamic Traffic Systems through Active Mobile Sensor Networks and Crowdsourced Data”, and partially funded by National Science Foundation – United States under Grant No. CMMI 1663657 “Real-time Management of Large Fleets of Self-Driving Vehicles Using Virtual Cyber Tracks”. The work presented in this paper remains the sole responsibility of the authors.

Appendix A: Frank-Wolfe algorithm for nonlinear programming models

The algorithm procedure is described as follows.

Step 1: initialization: $k = 0$, and find one feasible solution as x_0 ;

Step 2: Based on the first-order Taylor approximation of $f(x)$ around x_k , minimizing the linear approximation: $\min s_k^T \nabla f(x_k)$, and s_k is subject to all constraints.

Step 3: Find γ that minimizes $f(x_k + \gamma(s_k - x_k))$, subject to $0 \leq \gamma \leq 1$.

Step 4: Update: $x_{k+1} = x_k + \gamma(s_k - x_k)$. If $|x_{k+1} - x_k| \leq \Delta$ or $k = K$, stop. Otherwise, $k = k + 1$ and go to step 2.

Specifically, at step 2, $\nabla f(x_k) = \beta_1 \sum_a \sum_{(i,j,t,s)} [2 \times c_{i,j,t,s} \times c_{i,j,t,s} \times x_{i,j,t,s}^a(k)] + \beta_2 \sum_{(i,j,\tau)} [\sum_a \sum_{t \in \tau} (2 \times x_{i,j,t,s}^a(k))] \times s_{i,j,t,s}^a(k)$ is constant, so $s_k^T \nabla f(x_k) = \beta_1 \sum_a \sum_{(i,j,t,s)} [2 \times c_{i,j,t,s} \times c_{i,j,t,s} \times x_{i,j,t,s}^a(k) \times s_{i,j,t,s}^a(k)] + \beta_2 \sum_{(i,j,\tau)} [\sum_a \sum_{t \in \tau} (2 \times x_{i,j,t,s}^a(k) \times s_{i,j,t,s}^a(k))]$. Finally, the model proposed in step 2 is a linear programming model with the flow-balance constraint. In addition, at step 1, for finding one feasible solution, we can define a simple linear objective function, so the model will be a linear programming model with the flow balance constraint, which can be solved by the Dantzig-Wolfe algorithm due to the special block structure as well.

References

- Barry, J., Freimer, R. and Slavin, H., 2009. Use of entry-only automatic fare collection data to estimate linked transit trips in New York City. *Transportation Research Record: Journal of the Transportation Research Board*, (2112), 53-61.
- Bianco, L., Confessore, G., Reverberi, P., 2001. A Network Based Model for Traffic Sensor Location with Implications on O/D Matrix Estimates. *Transportation Science* 35(2), 50-60
- Bierlaire, M., 2002. The total demand scale: a new measure of quality for static and dynamic origin–destination trip tables. *Transportation Research Part B: Methodological*, 36(9), 837-850.
- Canepa, E.S., Claudel, C.G., 2017. Networked traffic state estimation involving mixed fixed-mobile sensor data using Hamilton-Jacobi equations. *Transportation Research Part B: Methodological*, 104, 686-709.
- Castillo, E., Conejo, A.J., Pruneda R., Solares C., 2007. Observability in linear systems of equations and inequalities: Applications. *Computer & Operations Research*, 34, 1708-1720.
- Castillo, E., Conejo, A.J., Menéndez, J.M. and Jiménez, P., 2008. The observability problem in traffic network models. *Computer-Aided Civil and Infrastructure Engineering*, 23(3), 208-222.
- Castillo, E., Grande, Z., Calviño, A., Szeto, W.Y. and Lo, H.K., 2015. A state-of-the-art review of the sensor location, flow observability, estimation, and prediction problems in traffic networks. *Journal of Sensors*, 2015.
- Ceapa, I., Smith, C. and Capra, L., 2012. Avoiding the crowds: understanding tube station congestion patterns from trip data. In *Proceedings of the ACM SIGKDD international workshop on urban computing* (134-141). ACM.
- Chen, H.K. and Hsueh, C.F., 1998. A model and an algorithm for the dynamic user-optimal route choice problem. *Transportation Research Part B: Methodological*, 32(3), 219-234.
- Claudel, C., Hofleitner, A., Mignerey, N., Bayen, A., 2009, January. Guaranteed bounds on highway travel times using probe and fixed data. In *88th TRB Annual Meeting Compendium of Papers*.
- Claudel, C.G., Bayen, A.M., 2010. Lax–Hopf based incorporation of internal boundary conditions into Hamilton–Jacobi equation. Part I: Theory. *IEEE Transactions on Automatic Control*, 55(5), 1142-1157.
- Claudel, C.G., Bayen, A.M., 2011. Convex formulations of data assimilation problems for a class of Hamilton–Jacobi equations. *SIAM Journal on Control and Optimization*, 49(2), 383-402.
- Daganzo, C.F., 2007. Urban gridlock: Macroscopic modeling and mitigation approaches. *Transportation Research Part B: Methodological*, 41(1), 49-62.
- Dantzig, G.B., 1998. *Linear programming and extensions*. Princeton university press.
- Desrosiers, J. and Lübbecke, M.E., 2005. A primer in column generation. In *Column generation* (1-32). Springer US.
- Drissi-Kaitouni, O., Hameda-Benchekroun, A., 1992. A dynamic traffic assignment model and a solution algorithm. *Transportation Science* 26, 119–128.
- Ford Jr, L.R. and Fulkerson, D.R., 1958. A suggested computation for maximal multi-commodity network flows. *Management Science*, 5(1), 97-101.
- Fukushima, M., 1984. A modified Frank-Wolfe algorithm for solving the traffic assignment problem. *Transportation Research Part B: Methodological*, 18(2), 169-177.
- Furth, P.G., Hemily, B.J., Muller, T.H.J., Strathman, J.G., 2006. Uses of archived AVL-APC data to improve transit performance and management: *Transportation Research Board, TCRP Report No. 113*

- Gao, L., Wang, B., and Zhang, C., 2011. Comparison of urban rail transit fare clearing model based on travel survey. *Modern urban transit*, 11, 97-99.
- Gentili, M. and Mirchandani, P.B., 2012. Locating sensors on traffic networks: Models, challenges and research opportunities. *Transportation research part C: emerging technologies*, 24, 227-255.
- Kalvelagen, E., 2003. Dantzig-Wolfe Decomposition with GAMS.
- Kusakabe, T. and Asakura, Y., 2014. Behavioural data mining of transit smart card data: A data fusion approach. *Transportation Research Part C: Emerging Technologies*, 46, 179-191.
- Kusakabe, T., Iryo, T. and Asakura, Y., 2010. Estimation method for railway passengers' train choice behavior with smart card transaction data. *Transportation*, 37(5), 731-749.
- Lam, W.H. and Yin, Y., 2001. An activity-based time-dependent traffic assignment model. *Transportation Research Part B: Methodological*, 35(6), 549-574.
- Larsson, T. and Patriksson, M., 1992. Simplicial decomposition with disaggregated representation for the traffic assignment problem. *Transportation Science*, 26(1), 4-17.
- Larsson, T., Patriksson, M. and Rydgergren, C., 2004. A column generation procedure for the side constrained traffic equilibrium problem. *Transportation Research Part B: Methodological*, 38(1), 17-38.
- LaValle, S.M., 2012. Sensing and filtering: A fresh perspective based on preimages and information spaces. *Foundations and Trends® in Robotics*, 1(4), 253-372.
- Li, P., Mirchandani, P., Zhou, X., 2015. Solving simultaneous route guidance and traffic signal optimization problem using space-phase-time hypernetwork. *Transportation Research Part B* 81(1), 103–130.
- Liu, J., Kang, J.E., Zhou, X. and Pendyala, R., 2018. Network-oriented household activity pattern problem for system optimization. *Transportation Research Part C: Emerging Technologies*, 94, 250-269.
- Liu, J., Zhou, X., 2016. Capacitated transit service network design with boundedly rational agents. *Transportation Research Part B: Methodological*, 93, 225-250.
- Lu, C.C., 2016. Robust multi-period fleet allocation models for bike-sharing systems. *Networks and Spatial Economics*, 16(1), 61-82.
- Lu, C.C., Liu, J., Qu, Y., Peeta, S., Roupail, N.M., Zhou, X., 2016. Eco-system optimal time-dependent flow assignment in a congested network. *Transportation Research Part B* 94, 217–239.
- Lu, C.C., Zhou, X., Zhang, K., 2013. Dynamic origin–destination demand flow estimation under congested traffic conditions. *Transportation Research Part C*, 34, 16-37.
- Ma, X.L., Liu, C., Liu, J.F., Chen, F. Yu, H., 2015. Boarding stop inference based on transit IC card data. *Journal of Transportation Systems Engineering and Information Technology*, 15(4), 78-84.
- Ma, X.L., Wang, Y.H., Chen, F., Liu, J.F., 2012. Transit smart card data mining for passenger origin information extraction. *Journal of Zhejiang University Science C*, 13(10), 750-760.
- Ma, X.L., Wu, Y.J., Wang, Y., Chen, F. and Liu, J., 2013. Mining smart card data for transit riders' travel patterns. *Transportation Research Part C: Emerging Technologies*, 36, 1-12
- Meng, L. and Zhou, X., 2011. Robust single-track train dispatching model under a dynamic and stochastic environment: a scenario-based rolling horizon solution approach. *Transportation Research Part B: Methodological*, 45(7), 1080-1102.
- Munizaga, M.A. and 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.
- Nair, R., Miller-Hooks, E., Hampshire, R.C. and Bušić, A., 2013. Large-scale vehicle sharing systems: analysis of Vélolib'. *International Journal of Sustainable Transportation*, 7(1), 85-106.
- Nassir, N., Hickman, M. and Ma, Z.L., 2015. Activity detection and transfer identification for public transit fare card data. *Transportation*, 42(4), 683-705.
- Pelletier, M.P., Trépanier, M. and Morency, C., 2011. Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4), 557-568.
- Peeta, S. and Mahmassani, H.S., 1995. Multiple user classes real-time traffic assignment for online operations: a rolling horizon solution framework. *Transportation Research Part C: Emerging Technologies*, 3(2), 83-98.
- Seaborn, Catherine, John Attanucci, and Nigel H. M. Wilson.2009. Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data. *Transportation Research Record: Journal of the Transportation Research Board* 2121, 55–62.
- Shang, P., Li, R., Guo, j., Xian, K., Zhou, X., 2019. Integrating Lagrangian and Eulerian observations for passenger flow state estimation in an urban rail transit network: A space-time-state hyper network-based assignment approach. *Transportation Research Part B*, 121, 135-167.

- Shang, P., Li, R., Liu, Z., Yang, L. and Wang, Y., 2018. Equity-oriented skip-stopping schedule optimization in an oversaturated urban rail transit network. *Transportation Research Part C: Emerging Technologies*, 89, 321-343.
- Sun, Y., Xu, R., 2012. Rail transit travel time reliability and estimation of passenger route choice behavior: Analysis using automatic fare collection data. *Transportation Research Record* (2275), 58-67.
- Tong, L., Zhou, X. and Miller, H.J., 2015. Transportation network design for maximizing space-time accessibility. *Transportation Research Part B: Methodological*, 81, 555-576.
- Trépanier, M., Tranchant, N. and Chapleau, R., 2007. Individual trip destination estimation in a transit smart card automated fare collection system. *Journal of Intelligent Transportation Systems*, 11(1), 1-14.
- Wei, Y., Avci, C., Liu, J., Belezamo, B., Aydın, N., Li, P.T. and Zhou, X., 2017. Dynamic programming-based multi-vehicle longitudinal trajectory optimization with simplified car following models. *Transportation Research Part B: Methodological*, 106, 102-129.
- Wu, X., Guo, J., Xian, K., Zhou, X., 2018. Hierarchical travel demand estimation using multiple data sources: A forward and backward propagation algorithmic framework on a layered computational graph. *Transportation Research Part C: Emerging Technologies*, 96, 321-346.
- Xing, T., Zhou, X., Taylor, J., 2013. Designing Heterogeneous Sensor Networks for Estimating and Predicting Path Travel Time Dynamics: An Information-Theoretic Modeling Approach. *Transportation Research Part B*, 57, 66-90.
- Yang, H., Iida, Y. and Sasaki, T., 1991. An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts. *Transportation Research Part B: Methodological*, 25(5), 351-363.
- Yang, H. and Meng, Q., 1998. Departure time, route choice and congestion toll in a queuing network with elastic demand. *Transportation Research Part B: Methodological*, 32(4), 247-260.
- Yuan, N.J., Wang, Y., Zhang, F., Xie, X. and Sun, G., 2013. Reconstructing individual mobility from smart card transactions: A space alignment approach. In *Data Mining (ICDM), 2013 IEEE 13th International Conference*, 877-886.
- Zawack, D.J. and Thompson, G.L., 1987. A dynamic space-time network flow model for city traffic congestion. *Transportation Science*, 21(3), 153-162.
- Zhao, J., Rahbee, A., Wilson, N.H., 2007. Estimating a Rail Passenger Trip Origin-Destination Matrix Using Automatic Data Collection Systems. *Computer-Aided Civil and Infrastructure Engineering*, 22(5), 376-387.
- Zheng, Y., Capra, L., Wolfson, O. and Yang, H., 2014. Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(3), 38.
- Zhou, L., 2014. Verification and optimization of network clearing model. *Urban mass transit*, 11, 59-66.
- Zhou, X., Erdoğan, S., Mahmassani, H.S., 2006. Dynamic origin-destination trip demand estimation for subarea analysis. *Transportation Research Record*, 1964(1), 176-184.
- Zhou, X., List, G.F., 2010. An information-theoretic sensor location model for traffic origin-destination demand estimation applications. *Transportation Science*, 44(2), 254-273.
- Zhou, X. and Mahmassani, H.S., 2007. A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework. *Transportation Research Part B: Methodological*, 41(8), 823-840.
- Zhu, N., Fu, C., Ma, S., 2018. Data-driven distributionally robust optimization approach for reliable travel-time-information-gain-oriented traffic sensor location model. *Transportation Research Part B: Methodological*, 113, 91-120.
- Zhu, Y., Koutsopoulos, H.N. and Wilson, N.H., 2017a. A probabilistic Passenger-to-Train Assignment Model based on automated data. *Transportation Research Part B: Methodological*, 104, 522-542.
- Zhu, Y., Koutsopoulos, H.N. and Wilson, N.H., 2017b. Inferring left behind passengers in congested metro systems from automated data. *Transportation Research Part C: Emerging Technologies*.
- Zimmerman, J., Tomasic, A., Garrod, C., Yoo, D., Hiruncharoenvate, C., Aziz, R., Thiruvengadam, N.R., Huang, Y. and Steinfeld, A., 2011. Field trial of TIRAMISU: crowd-sourcing bus arrival times to spur co-design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (1677-1686). ACM.