

# Effective Prediction of Road Segment Occupancy for the Route-Reservation Architecture

C. Menelaou\* P. Kolios\* S. Timotheou\* C.G. Panayiotou\*

*\* KIOS Research and Innovation Center of Excellence, and the Department of Electrical and Computer Engineering, University of Cyprus, {cmenel02, pkolios, timotheou.stelios, christosp}@ucy.ac.cy*

**Abstract:** The effect of traffic congestion within an urban region can be eliminated by curbing the number of vehicles that use the congested part of the transport network. To do so, earlier work by the authors introduced a novel route-reservation architecture that manages the vehicle departure times and makes appropriate routing decisions so that vehicles arrive at their destination at their earliest possible time and avoiding congested parts of the network Menelaou et al. (2017a,b). Under the proposed route-reservation architecture, the future state of each road segment is estimated based on the received reservations and assuming that all vehicles travel at a constant speed which is set equal with the free-flow speed. Even within a homogeneous region, this assumption is not always valid which can lead to wrong predictions of the time that a vehicle occupies a road segment which in turn affects the accuracy of the overall reservation architecture. Therefore, the key objective of this work is to investigate different estimation schemes that can improve the prediction of the vehicle transit time within each road segment. In this article we explore two prediction methods, an Exponential Moving Average (EMA) and a Multiple Linear Regression (MLR) method. The performance of the two prediction schemes is also presented. Finally, realistic simulations results across an urban region of San Francisco are presented to demonstrate the gains can be achieved from the proposed prediction methods.

*Keywords:* Connected and Automated Vehicles; Intelligent Transportation Systems; Modeling, Control and Optimization of Transportation Systems; Congestion Avoidance.

## 1. INTRODUCTION

Traffic congestion constitutes a major problem in modern cities causing a variety of socio-economic drawbacks including environmental pollution, driver delays and frustration. The major cause of congestion is the lack of efficient traffic management to support the changing demand that affects the network capacity, Chen et al. (2001). Even if real-time network-state information is given to the drivers, traffic congestion would still persist as all rational drivers will seek to navigate through non-congested road-segments producing state oscillations between congested and non-congested road segments, Çolak et al. (2016).

In this work, we adopt the research-reservation architecture proposed in, Menelaou et al. (2017a,b) which, in the context of connected and autonomous vehicles, prevents the occurrence of oscillations by assigning vehicles to non-congested routes, while the reservations are also exploited to produce estimates of the future network state. In the proposed reservation architecture, when a vehicle is about to start its journey, it informs a Road Side

Unit (RSU) about its origin and destination ( $O - D$ ) and its expected starting time. The RSU responds with the vehicle's assigned departure time from  $O$  which can be a little later than the preferred start time, and the best possible route that the vehicle should follow such that the vehicle will avoid traversing road segments that are expected to be at their critical capacity (these are referred to as non-admissible road segments). The objective of the RSU is not only to avoid the non-admissible segments but also to find the path that would allow the vehicle to arrive at its destination at the earliest possible time. Thus, vehicles may be instructed to wait at their origin before commencing their trip, if this action will lead to a shorter path becoming available which can lead to an earlier arrival time. In this way, the proposed reservation architecture avoids excessively longer paths that would also imply higher energy cost. In fact, the RSU can adopt different objective functions such that delayed arrivals can be traded for energy savings. Once the RSU has identified the vehicle's best possible path, it also reserves each road segment at the time-slots that the vehicle is expected to traverse each road segment. Therefore, for each vehicle there is a detailed reservation plan along the exact route that the vehicle should follow from  $O$  to  $D$ .

An important assumption of the proposed architecture is that each segment's transit-time is calculated assuming a constant speed, typically the free-flow speed. However,

\* This work has been supported by the European Union's Horizon 2020 research and innovation programme under grant agreement No 739551 (KIOS CoE), the Government of the Republic of Cyprus through the Directorate General for European Programmes, Coordination and Development and the European Research Council Advanced Grant FAULT-ADAPTIVE.

this assumption is quite restrictive. Even if it is guaranteed that the state of all road segments is below their critical capacity, it does not mean that speed in each segment will always be the free-flow speed. For example, at intersections between a higher priority and a lower priority road segment, vehicles in the lower priority segment will have to slow down resulting in an average speed well below the free-flow speed.. Furthermore, interactions between vehicles introduce randomness in the transit time while significant randomness is inherent in the stochastic nature of any transport network. The inaccuracy in the vehicle speed, leads to an error in the estimation of the vehicle's transit time in the road segment which in turn causes an error in the future predicted state of the network. This error will result in an inaccurate route reservation plan which can have a negative impact on the effectiveness of the proposed architecture.

The key objective of this work is to investigate better methods for estimating the time that vehicles take to traverse each road segment. Towards that end, the RSU also collects probing data from the connected vehicles which allows it to more accurately estimate the average traversal time of each road segment. Recent advances in connected and autonomous vehicle technologies definitely support the development of such a scheme. In this context, when a vehicle exits a road-segment it sends to the RSU the time it took to traverse this particular road segment. The RSU utilizes this information to make near-future predictions of the transit-times of all road segments without requiring to know the average vehicle's speeds in the network or any other information about the network's state. Therefore, as time goes by, it is expected that the road segment traversal time estimates will be closer to the true values and thus the future state predictions will be more accurate and therefore the reservations will be more effective. This paper investigates two transit-time prediction methods which constitute the main contribution of this paper. Furthermore, the improved road segment traversal time estimates, are utilized to demonstrate the improvement in the prediction of the future state of the network which in turn results to more effective path reservations.

The transit-time of each road segment can be interpreted as a time-series where samples are affected by the segment's instantaneous density. As the density of the segment increases the speed decreases and hence, the transit-time increases, Immers and Logghe (2003) exhibiting seasonal changes (depending on density). Considering such behavior, this work investigates two simple prediction methods to address the misprediction problem under the route-reservation architecture. The first method utilizes an Exponential Moving Average (EMA) approach to capture the dynamics as the transit-time within a road segment gradually increases or decreases. Therefore, the segment's transit-time for a near-future upcoming request is defined by the recent measured transit-times as EMA apply more weight to the latter observations. An important feature of the EMA is that, depending on its parameter, it can react faster to recent changes improving the reservation accuracy, Crowder (1987). The second method is based on the Multiple Linear Regression (MLR) technique that uses several explanatory variables to predict the outcome of a response variable as the segment's transit-time depends on

more than one factors. In this work, one factor is segment's current density while the second factor is the density of all neighboring segments that share the same exit junction, Bryman and Cramer (1994). Hence, there is seasonality on transit-time response, meaning that there is a pattern of transit-times that is related with these two depended variables (factors). Furthermore, this approach capture the interactions between vehicles at intersections.

The remainder of this work is organized as follows: Section 2 discusses the relevant literature and Section 3 briefly describes the mathematical formulation of the route-reservation architecture and the Earliest Arrival Time at Destination problem (similar definitions are also presented in, Menelaou et al. (2017a)). Section 4 introduces the investigated transit-time prediction methods while extensive simulations results are illustrated in Section 5 demonstrating the significant improvements that can be achieved when such predictions approaches are utilized. Finally, Section 6 concludes this work and discusses future research directions.

## 2. RELATED WORK

*Gating* and *Perimeter Control* methods constitute the state-of-the-art congestion alleviation approaches. Both control the boundary and external flows in a protected region by restricting the region's input flows, Keyvan-Ekbatani et al. (2012); Haddad and Geroliminis (2012). This can be achieved utilizing various techniques such as street closures, pricing, and traffic light signaling. The traffic dynamics in the protected region are described through the Network Fundamental Diagram (NFD) (extracted from real-time measurements) which provide the observed relation between the network-wide space-mean flow and density. Utilizing the NFD increases the modeling accuracy and ensures a relatively low implementation overhead. Nevertheless, the use of gating or perimeter control approaches may not always be beneficial, Keyvan-Ekbatani et al. (2015) as unwanted queues may be observed at the boundaries of the region. Queues may obstruct the upstream network destinations and the benefits of both approaches will be significantly reduced as the queue lengths cannot be efficiently estimated. On the other hand, the major advantage of the proposed route-reservation architecture is that by reserving the path to be followed by all vehicles, vehicles are forced to avoid non-admissible road segments thus, spill-backs and formation of queues is avoided.

In recent decades travel-time predictions constitutes a useful information for ITS applications which aim to minimize travelers journey time either in an on-line or an off-line manner, Kaysi et al. (1993); Kaufman and Smith (1993). Travel times can be predicted utilizing static or stochastic prediction models based on real-time (or historic) data that can be obtained through various traffic surveillance devices e.g., loop detectors, mobile detectors, radars and cameras, Jiang and Li (2013); Xiao and Lo (2014). The majority of the on-line prediction methods are based on a Kalman Filter-based algorithms which enables the prediction of the state variables to be continually updated as new observations becomes available. Kalman filter estimates require the historic data, Chien and Kuchipudi

(2017) of a previous day (with similar traffic situation) in order to obtain the transition parameters, which describe the relationship between state variables. Unfortunately, Kalman Filter cannot be utilized in our approach as there is no simple and accurate model that can justify the route-reservation architecture. Furthermore, the identifications of real-time and accurate short-term transit-time predictions constitutes a difficult task considering the scale and complexity of the problem. In this way, this work develops two simple and accurate short-term prediction methods provide accurate near-future segments transit-times considering on-line vehicles probe data.

### 3. PROBLEM FORMULATION

We consider a homogeneous region of the transport network expressed as a graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  with vertices  $\mathcal{V}$ ,  $N_V = |\mathcal{V}|$ , denoting the road junctions and edges  $\mathcal{E}$ ,  $N_E = |\mathcal{E}|$ , representing the road segments. Each road segment  $(i, j) \in \mathcal{E}$ ,  $\{i, j\} \in \mathcal{V}$  is described by parameters  $\lambda_{ij}$ , denoting the number of lanes and  $l_{ij}$  representing the segment's length.

The time horizon is quantized into time-slots of duration  $T$  where parameter  $\hat{\tau}_{ij}(t)$  represents the predicted transit-time of segment  $(i, j)$  for time-slot  $t$ , while the parameter  $\tau_{ij}^k(t)$  denotes the transit-time of road-segment  $(i, j)$  as measured by the  $k$ -th vehicle that has passed segment  $(i, j)$  and exits from junction  $j$  at time-slot  $t$ . All regional traffic dynamics are defined according to the NFD parameters  $\rho^C$ ,  $\rho^J$  and  $v^f$  representing the critical density corresponding to the maximum flow, the jam density, and the free-flow speed, respectively. Also, for each segment  $(i, j)$  the traffic dynamics are described according to parameters  $\rho_{ij}(t)$ ,  $\rho_{ij}^C$  and  $\rho_{ij}^J$  representing the instantaneous density at time-slot  $t$ , the critical density and the jam density, respectively. Note that, in this work the parameter  $\rho_{ij}^C$  is approximated with the quantity  $(\rho^C / \rho^J) \rho_{ij}^J$ .

Let, variable  $r_{ij}(t)$  denote the accumulated number of vehicle reservations in segment  $(i, j)$  while  $R(t)$  expresses the network's accumulated number of reservations (i.e.,  $R(t) = \sum_{(i,j) \in \mathcal{E}} r_{ij}(t)$ ), for time-slot  $t$ . Additionally, let  $n_{ij}(t)$  and  $N(t)$  represent the accumulated density and the total network density at time-slot  $t$  (i.e.,  $N(t) = \sum_{(i,j) \in \mathcal{E}} n_{ij}(t)$ ).

The road segment  $(i, j)$  is assumed to be admissible at time  $t$  if a vehicle that enters in junction  $i$  at time-slot  $t$  can traverse the road segment  $(i, j)$  in say  $\hat{\tau}_{ij}(t)$  time units, while the number of accumulated reservations ( $r_{ij}(t)$ ) is no larger than the segment's critical density ( $\rho_{ij}^C$ ) for the entire duration of  $\hat{\tau}_{ij}(t)$ . The *admissibility state* of road segment  $(i, j)$  at time-slot  $t$  is denoted by the variable  $x_{ij}(t)$  where,  $x_{ij}(t) = 1$  if road segment  $(i, j)$  is admissible at time slot  $t$  and  $x_{ij}(t) = 0$ , otherwise. Mathematically,  $x_{ij}(t)$  can be expressed as follows:

$$x_{ij}(t) = \begin{cases} 1, & \text{if } r_{ij}(t+k) \leq \rho_{ij}^C, \forall k = 0, \dots, \hat{\tau}_{ij}(t) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Based on the above notation, vehicles are allowed to traverse road segments during time-slots where  $x_{ij}(t) = 1$  since the RSU is able to make reservations only along

those time-slots. Hence, a vehicle may be instructed to wait at its origin ( $O$ ) until a non-admissible road segment becomes available, or it can traverse through alternative admissible road segments. In addition, a combination of the two aforementioned options may be employed i.e., wait for a short period of time at  $O$  and then follow an alternative route. By allowing waiting at the origin of each vehicle, the cost of traversing a road segment  $c_{ij}(t)$  can mathematically be defined as follows:

$$c_{ij}(t) = \begin{cases} \hat{\tau}_{ij}, & \text{if } x_{ij}(t) = 1, i \neq O \\ \hat{\tau}_{ij} + w, & \text{if } x_{ij}(t) = 0, i = O \\ \infty, & \text{if } x_{ij}(t) = 0, i \neq O \end{cases} \quad (2)$$

where  $w \in \mathbb{N}$  denotes the number of possible time-slots that a vehicle has to wait at its origin. Similar to, Menelaou et al. (2017a) below we defined the Earliest Arrival Time at Destination Problem (EATD) that is solved by the RSU for every vehicle request.

*Earliest Arrival Time at Destination (EATD) problem:* Given an origin-destination ( $O - D$ ) pair, the time-stamp  $t_0$  at which the routing request is made, and the reservation states  $x_{ij}(t)$ ,  $(i, j) \in \mathcal{E}$ ,  $\forall t \geq t_0$ , then the EATD problem requests the earliest-arrival-time-at-destination (from  $O$  to  $D$ ). Let  $d_i$  denote the arrival time at junction  $i$  and  $p_h$  the  $h$ -th path from origin ( $O$ ) to destination ( $D$ ) denoted as  $p_h = (v_0^h, v_1^h), (v_1^h, v_2^h), \dots, (v_{L_h-1}^h, v_{L_h}^h)$ , where  $v_j^h \in \mathcal{V}$  is the  $j$ -th visited vertex in the  $h$ -th path,  $v_0^h = O$ ,  $v_{L_h}^h = D$ , and  $L_h$  is the number of hops of the  $h$ -th path. In the same way, let  $d_j^h$  represent the earliest arrival time at junction  $j$  if the  $h$ -th path is considered assuming that the vehicle was delayed by  $w$  at the origin, which mathematically can be expressed as:

$$\begin{aligned} d_{v_0^h}^h &= t_0 \\ d_{v_1^h}^h &= d_{v_0^h}^h + c_{v_0^h, v_1^h}(d_{v_0^h}^h) \\ &\vdots \\ d_{v_{L_h}^h}^h &= d_{v_{L_h-1}^h}^h + c_{v_{L_h-1}^h, v_{L_h}^h}(d_{v_{L_h-1}^h}^h) \end{aligned} \quad (3)$$

Therefore, the EATD problem can be mathematically expressed as:

$$(II) \ d_D^* = \min_{w, p_h} d_D^h \quad (4)$$

s.t. Constraints (1) – (3) are satisfied.

### 4. TRANSIT-TIME PREDICTION METHODS

This section describes the proposed predictors, EMA and MLR, that are utilized to predict the current transit-time of each road section for improving the accuracy of the reservation scheme. Both predictors are based on statistical techniques where the  $k$ -th observation (i.e.,  $\tau_{ij}^k(t)$ ) is interpreted as a time-series that utilizes the most recent transit-time observations in order to forecast the near future transit-times.

#### 4.1 Exponential Moving Average (EMV) Predictor

Moving averages are lagging indicators which means that if a function is trending down (or up) persistently, then

the moving averages are lagging behind when following such trend as they average the recent observation with older (higher or lower) observations. Therefore, the moving average has difficulty capturing peaks and valleys of a series, failing to forecast non-stationary series. On the other hand, the exponential moving average reduces this lag by putting more weight on the more recent observations. Hence, using the exponential moving average results in transit-time predictions that are more responsive to recent changes.

To make the notation simpler, in this subsection, a single road-segment is considered where the transit time of the  $k$ -th observed vehicle is denoted by parameter  $y_k$  while the variable  $\hat{y}_k$  denotes the predicted value of the next transit-time. When, a new observation is received, it is used to update the forecast transit-time of the next vehicle (*i.e.*,  $y_{k+1} = \hat{y}_k$ ). Notably, this notation can be similarly considered by all road segments. The formula of EMA is mathematically described as follows:

$$\hat{y}_{k+1} = \alpha y_k + (1 - \alpha) \hat{y}_k \quad (5)$$

where,  $0 \leq \alpha \leq 1$  is the EMA parameter that determines the relative importance of recent observations. Values closer to one put more emphasis on recent observations. Expanding equation (5) over  $k$ , yields:

$$\hat{y}_{k+1} = \alpha \sum_{i=0}^{k-1} (1 - \alpha)^i y_{k-i} + (1 - \alpha)^k y_0 \quad (6)$$

where,  $k$  denotes the  $k$ -th vehicle that exits the specific road-segment. From equation (6) we observe that we also require the segment's initial transit-time value. This is assumed to be equal to the time that the vehicle will require to traverse the segment with free-flow speed (*i.e.*,  $y_0 = l_{ij}/v^f$ ).

#### 4.2 Multiple Linear Regression (MLR) Predictor

Again, to make the notation simpler, in this subsection we consider an isolated junction  $j$  and also we consider all road segments  $(p, j)$  that terminate (share) junction  $j$  where  $p = 1, \dots, |\mathcal{P}|$  and  $\mathcal{P} \subseteq \mathcal{V}$ ,  $(p, j) \subseteq \mathcal{E}$  (*i.e.*,  $|\mathcal{P}|$  represents the number of road-segments that share junction  $j$ ). For this section the segment  $(p, j)$  with  $p = 1$  is considered as the segment that its transit-time is going to be predicted at the current time-slot  $t$ . Also, for this section consider that vehicles inform the RSU not only about their exit time from the road segment and their transit-time duration but also they inform the RSU about their entry time on each pre-assigned road-segment<sup>1</sup>. Considering this information, the RSU has an accurate calculation of the current density of each road-segment utilizing only a simple counter,  $n_{p,j}$

$$n_{pj}(t) = \begin{cases} n_{pj}(t-1) + 1 & \text{if vehicle enters } (p, j), \\ n_{pj}(t-1) - 1 & \text{if vehicle exits } (p, j), \\ n_{pj}(t-1) & \text{otherwise} \end{cases} \quad (7)$$

with  $n_{pj}(0) = 0$ .

In transport networks transit times are highly correlated with the network's density, Immers and Logghe (2003).

<sup>1</sup> Note that this information is already known by the RSU. Since it knows the assigned path, the entry time in a road section coincides with the exit time from the previous road section

The segment's transit-time is affected from its current density and also the densities of neighboring segments (*e.g.*, the ones sharing a node) which have higher priority. For example, consider a vehicle that is waiting on a junction because other vehicles on a higher priority segment are passing. In this way, the response variable (segment's transit-time at time-slot  $t$ ) depends linearly on two predictor variables, the current segment's density (*i.e.*,  $n_{1j}(t)$ ), and the maximum density of the neighboring road-segments that share junction  $j$  (*i.e.*,  $\max(n_{pj}(t)) p = 2, \dots, |\mathcal{P}|$ ). Similarly to the EMA method, let the  $k$ -th observed transit-time in segment  $(1, j)$  be denoted by parameter  $y_k$  while the parameter  $\hat{y}_k$  denote the transit-time predicted value. Hence, the next-slot transit-time is forecast as (*i.e.*,  $y_{k+1} = \hat{y}_k$ ). Accordingly, let the vector  $\mathbf{N}_1$  represent all density calculations of segment  $(1, j)$  for all time slots that an observation arise. Hence, the element at the  $k$ -th row (*e.g.*,  $(n_k^1)$ ) contains the segment's density  $n_{1j}(t)$  where,  $t$  equals with the time-slot that the  $k$ -th observation is attained. Similarly, the vector  $\mathbf{N}_2$  represents the maximum value of the estimated densities from all neighboring road segments that share junction  $j$  (*i.e.*,  $\max(n_{pj}(t)) p = 2, \dots, |\mathcal{P}|$ ) for every time slot  $t$  that an observation arise. Note that, this approach is similarly considered for all other road segments. Thus the MLR method that is used is mathematical described as follows:

$$\mathbf{y} = \mathbf{N}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (8)$$

where,

$$\mathbf{N} = [\mathbf{I} \ \mathbf{N}_1 \ \mathbf{N}_2 \ \mathbf{N}_1^T \mathbf{N}_2] = \begin{bmatrix} 1 & n_1^1 & n_1^2 & n_1^1 n_1^2 \\ 1 & n_2^1 & n_2^2 & n_2^1 n_2^2 \\ \dots & \dots & \dots & \dots \\ 1 & n_k^1 & n_k^2 & n_k^1 n_k^2 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_k \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

and  $\boldsymbol{\epsilon}$  are the residuals terms of the model. In this approach we seek to find the least squares parameters that estimate  $\boldsymbol{\beta}$  which minimize the least square error:

$$\sum_{i=1}^k \epsilon_i^2 = (\mathbf{y} - \mathbf{N}\boldsymbol{\beta})^T (\mathbf{y} - \mathbf{N}\boldsymbol{\beta}) \quad (9)$$

The regression repeats on every new observation, trying to identifying the "best" possible  $\hat{\boldsymbol{\beta}}$  in the sense that the sum of squared residuals is minimized. Hence, we can predicted the segment's transit times as follows,

$$\hat{y}_{k+1} = [1 \ n_k^1 \ n_k^2 \ n_k^1 n_k^2] \hat{\boldsymbol{\beta}} \quad (10)$$

where,  $\hat{y}_{k+1}$  are the predicted values in our regression model. Similarly to the EMA method, the predicted initial values are estimated as (*i.e.*,  $y_0 = l_{ij}/v^f$ ).

#### 4.3 EATD Solution

Both prediction methods are evaluated using the Route Reservation Algorithm (RRA) as introduced in, Menelaou et al. (2017a). RRA solves the EATD problem as described in Section 3 in two loops, the inner and outer loop. The inner loop is responsible to identify the earliest-destination-arrival-time path, based on the assumption that a vehicle is allowed to wait at any intermediate node

until an edge's state changes from non-admissible to admissible. Subsequently, in the outer loop if any waiting has been identified in the intermediate nodes of the computed path, then the minimum waiting across all intermediate nodes is transferred at the origin and the vehicle's starting time is updated accordingly. With the new starting time, the inner loop reiterates until a path is computed that does not require any waiting at intermediate nodes but allowing a possible delay only at the originated junction (*i.e.*,  $w \geq 0$ ). Note that the admissibility state of each edge is calculated according to constraint (1). Though in, Menelaou et al. (2017a) the transit-time of each segment is calculated using the constant travel time cost, in this work, the transit-time cost is dynamically obtained by either the EMA or the MLR predictors. As demonstrated in Section 5 both predictions can improve the vehicle's travel times as the accuracy of the reservations is improved.

## 5. SIMULATION RESULTS

The network under consideration is an  $1.8 \text{ km}^2$  unsignalized homogeneous region of down-town San Francisco (consisted of 99 road junctions and 208 single-lane road segments similar as in, Menelaou et al. (2017a)). Traffic mobility is simulated using the SUMO micro-simulator, Behrisch et al. (2011), where the Kraus's car-following model, Krauss et al. (1997), is assumed with the following parameters: vehicle length  $5 \text{ m}$ , maximum speed  $15 \text{ m/s}$ , acceleration  $2.5 \text{ m/s}^2$ , deceleration  $4.5 \text{ m/s}^2$ , driver imperfection  $5\%$ , driver reaction time  $0.5 \text{ s}$ , minimum gap distance  $2.5 \text{ m}$ , and simulation time-step  $0.1 \text{ s}$ . The NFD parameters to be used were set as indemnified in, Menelaou et al. (2017a) with  $\rho_{ij}^C = 40 \text{ veh/km/lane}$  (*i.e.*, around  $40\%$  of the regions total density). For the presented results 5 Monte Carlo realizations were conducted within which the  $(O - D)$  pairs and inter-arrival times were randomly generated, with flow rates varying between  $3000 - 8000 \text{ veh/h}$  over a 2 hours simulation time.

We emphasize that the proposed predictors utilize the recursive route reservation (RRA) algorithm as considered in, Menelaou et al. (2017a). Therefore, simulation results include the comparison of normal RRA algorithm (constant  $v^f$  was used for reservation estimations) against our proposed predictors EMA and LMR (where adaptive transit-time predictions were used). The aforementioned approaches are also compared against the uncontrolled scenario (US) (*i.e.*, where vehicles select their path strictly based on shortest path). Note that, in the proposed solution, new route reservations are computed solely based on information from previous reservations made and not the actual network state. Therefore, the true travel times may vary from the estimated travel times due to various sources of uncertainty as a number of different factors can affect journeys duration. The results in the sequel demonstrate the superior performance that can be achieved with respect to transit-time predictions using either EMA or MLR methods.

Figures 1 (a) and (b) depict the mean value of vehicles travel times and the number of vehicles that have managed to finish their journey within the simulation time, respectively. More specific, the scattered plots in Figure 1 (a) show the mean travel time of each realization, while the

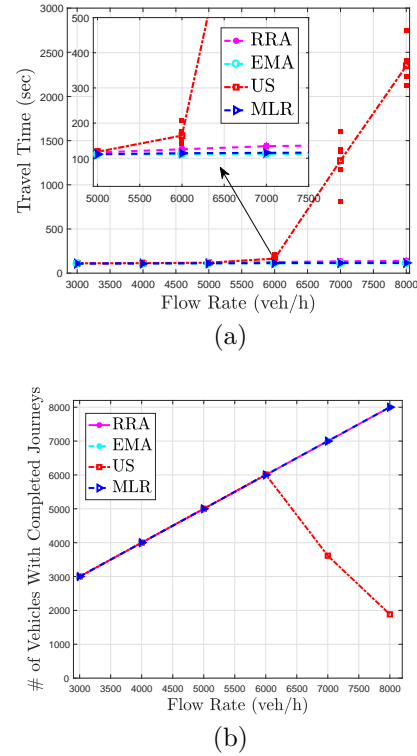


Fig. 1. (a) Average travel time and (b) Number of vehicles with completed journeys.

dashed lines represent the average travel time for all realizations. Similarly, the scattered plots in Figure 1 (b) show the average number of vehicles that have finished their journey while the scattered plots represent the realizations obtained by each simulation. For low flow rates (until  $5000 \text{ veh/h}$ ), no congestion occurs, and all approaches have similar behavior as US. At higher flow rates, EMA and MLR perform better than RRA as the travel times are not affected due to the flow rate increment while the travel time for both approaches slightly increases (around 2s) compared with the lower flow rates. Also, vehicle of all approaches manage to finish their journeys within the simulation time while in case of US vehicles seeks to delay as congestion occurs.

Figure 2 (a) illustrates the travel time distribution for the flow rate of  $8000 \text{ veh/h}$ . As can be identified, both EMA and LMR greatly improved travel time distributions compared to RRA offering a more robust solution as travel time distribution is non-increased irrespective of the higher demand. The mean travel time for normal RRA is  $135.9 \text{ s}$ , for EMA is  $109.7 \text{ s}$ , for LMR  $109.9 \text{ s}$  and for US is  $2123.1 \text{ s}$ . The standard deviation for RRA is  $64.8 \text{ s}$ , for EMA and LMR is  $40.1 \text{ s}$  and for US is  $2774.1 \text{ s}$  illustrating that as congestion of the road segments increases, EMA and LMR methods are more stable than RRA since travel times estimate do not significantly deviate from the actual travel times. Traveling time deviations of normal RRA are observed due to travel time mispredictions. Considering this, Figure 2 (b) depicts with red color the network's road segments that exhibited the highest residual error (*i.e.*,  $\epsilon = n_{ij}(t) - r_{ij}(t)$ ) higher than 13 vehicles. This figure

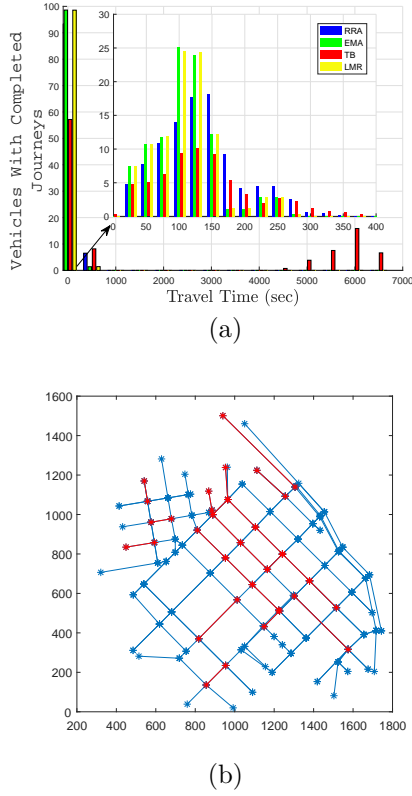


Fig. 2. (a) Travel time distribution of 8000 *veh/h* and (b) Road-segments mispredictions for 8000 *veh/h* (RRA).

indicates that about one third of all segments exhibit high residual errors.

Figures 3 (a), (b), (c) illustrate a comparison of the reservation estimates compared to the actual network density for RRA, EMA and MLR methods, respectively. The green line represents the accumulated number of reservations (*i.e.*,  $R(t)$ ), the yellow denotes the total network density (*i.e.*,  $N(t)$ ) and the blue line represent the accumulated network residual error (*i.e.*,  $E = N(t) - R(t)$ ). Similarly, the per road segment residual errors (*i.e.*,  $\epsilon$ ) are depicted in Figures 3 (d), (e) and (f) for RRA, EMA and MLR methods, respectively. Both set of figures demonstrate the superior improvements in transit-times predictions that EMA and MLR methods can offer. The EMA method seems to be slightly better than the MLR method as the average accumulated residual error for EMA is 3.6 *veh*s and for MLR is 4.3 *veh*s. Note, the significantly higher average accumulate residual error of normal RRA which is up to 31.8 *veh*s.

Figures 4 (a), (b) and (c) show the required waiting duration at the origin before vehicles depart for their journeys for RRA, EMA and MLR, respectively. All methods show that as demand increases, waiting time at origin increases too, with the initial delays for EMA and MLR methods to be about 3 times higher than the normal RRA algorithm. Note that this behavior is expected since better transit-time predictions explicitly imply a decrease in capacity as vehicles requires more time-slots to pass through the road segments rather than the underestimated mispredictions due to constant speed (*i.e.*,  $v^f$ ) assumptions. Even so,

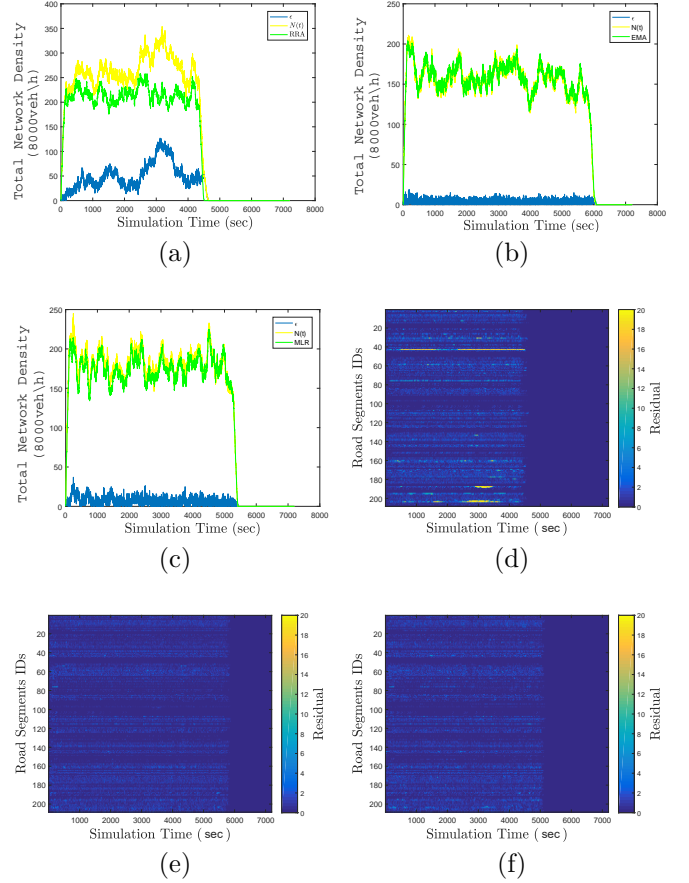


Fig. 3. Accumulated Network residual (a) RRA, (b) EMA, (c) MLR. The per road segment residual of (d) RRA, (e) EMA and (f) LMR (80000 *veh/h*).

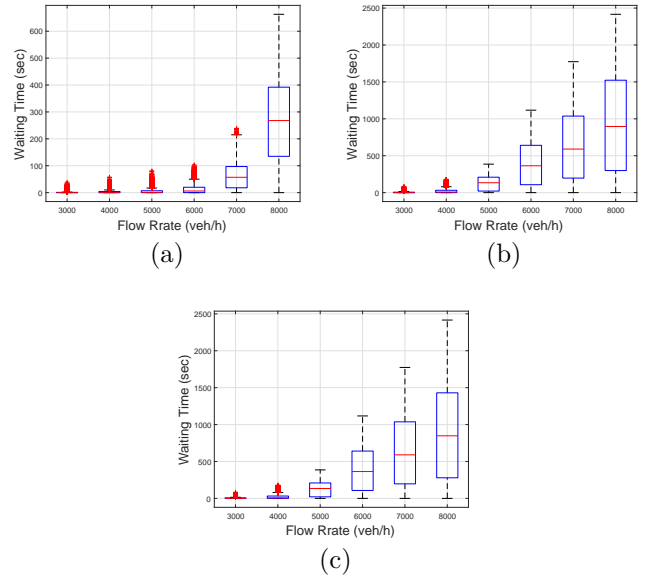


Fig. 4. Origin waiting-time that vehicles need to wait before departing for (a) RRA, (b) EMA and (c) LMR.

the average waiting is within acceptable levels (14 *min* for the highest demand approach) inasmuch as, waiting times

observed at the origin are not comparable to waiting due to congestion. Comparing EMA and MLR approaches the MLR has slightly lower waiting time but with, slightly worst transit-time predictions as illustrate in Figures 3 (b) and (c). A performance tradeoff thus exists where longer waiting times impact more accurate reservations. Depending on the application scenario, one of the two predictors can be employed to improve one of the two competing objectives.

## 6. CONCLUSIONS

This work proposes two simple transit-times predictors concentrated to improve reservations accuracy under the route-reservation architecture. The route-reservation architecture seems to be significantly beneficial for the overall network utilization. Unfortunately, the non-accurate transit-time predictions may lead to unstable solutions as long queues may observed. Hence, the application of reservations on bigger subareas rather than individual segments is prohibited as transit-time mispredictions may lead to network congestion. In this way, the proposed predictors make feasible the investigation of the route-reservation architecture on bigger subareas rather than individual road-segments. This option, can result in a complexity reductions enable more scalable routing solutions considering larger networks. Additionally, better transit-time estimates make feasible the investigation of how network operation can be improved if only a small percentage of vehicles follows their RSU plan.

## REFERENCES

- Aboudolas, K. and Geroliminis, N. (2013). Perimeter and boundary flow control in multi-reservoir heterogeneous networks. *Transportation Research Part B: Methodological*, 55, 265–281.
- Behrisch, M., Bieker, L., Erdmann, J., and Krajzewicz, D. (2011). Sumo-simulation of urban mobility-an overview. In *SIMUL 2011, The Third International Conference on Advances in System Simulation*, 55–60.
- Bryman, A. and Cramer, D. (1994). *Quantitative Data Analysis for Social Scientists*. Routledge, New York, NY, 10001.
- Chen, C., Jia, Z., and Varaiya, P. (2001). Causes and cures of highway congestion. *IEEE Control Systems Magazine*, 21(6), 26–32.
- Chien, S.I.J. and Kuchipudi, C.M. (2017). Dynamic travel time prediction with real-time and historic data. ASCE.
- Çolak, S., Lima, A., and González, M.C. (2016). Understanding congested travel in urban areas. *Nature communications*, 7(10793).
- Crowder, S.V. (1987). A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics*, 29(4), 401–407.
- Haddad, J. and Geroliminis, N. (2012). On the stability of traffic perimeter control in two-region urban cities. *Transportation Research Part B: Methodological*, 46(9), 1159–1176.
- Immers, L. and Logghe, S. (2003). *Traffic flow theory*. Department of Civil engineering Section Traffic and Infrastructure, Kastekpark Arenberg 40, B-3001 heverlee, Belgium. Course H 111.
- Jiang, Y. and Li, X. (2013). Travel time prediction based on historical trajectory data. *Annals of GIS*, 19(1), 27–35. doi:10.1080/19475683.2012.758173.
- Kaufman, D.E. and Smith, R.L. (1993). Fastest paths in time-dependent networks for intelligent vehicle-highway systems application\*. *Journal of Intelligent Transportation Systems*, 1(1), 1–11.
- Kaysi, I., Ben-Akiva, M., and Koutsopoulos, H. (1993). *An integrated approach to vehicle routing and congestion prediction for real-time driver guidance*. Ph.D. thesis, Doctoral dissertation, Massachusetts Institute of Technology Cambridge, Mass.
- Keyvan-Ekbatani, M., Kouvelas, A., Papamichail, I., and Papageorgiou, M. (2012). Exploiting the fundamental diagram of urban networks for feedback-based gating. *Transportation Research Part B: Methodological*, 46(10), 1393–1403.
- Keyvan-Ekbatani, M., Yildirimoglu, M., Geroliminis, N., and Papageorgiou, M. (2015). Multiple concentric gating traffic control in large-scale urban networks. *IEEE Transactions on Intelligent Transportation Systems*, 16(4), 2141–2154.
- Krauss, S., Wagner, P., and Gawron, C. (1997). Metastable states in a microscopic model of traffic flow. *Physical Review E*, 55(5), 5597.
- Menelaou, C., Kolios, P., Timotheou, S., Panayiotou, C., and Polycarpou, M. (2017a). Controlling road congestion via a low-complexity route reservation approach. *Transportation Research Part C: Emerging Technologies*, 81, 118–136.
- Menelaou, C., Timotheou, S., Kolios, P., and Panayiotou, C.G. (2017b). Improved road usage through congestion-free route reservations. *Journal of the Transportation Research Board*.
- Xiao, L. and Lo, H.K. (2014). Adaptive vehicle navigation with en route stochastic traffic information. *IEEE Transactions on Intelligent Transportation Systems*, 15(5), 1900–1912.