Measurement-based modeling of inter-arrivals for the simulation of highway vehicular networks

Marco Gramaglia, Marco Fiore, *Member, IEEE,* Maria Calderon

***Abstract*—Simulation is the standard approach to the performance evaluation of vehicular network solutions. However, the relevance of its results is often questioned by the adoption of unrealistic vehicular mobility representations. In this letter, we propose a measurement-based Hidden Markov Model that generates realistic synthetic inter-arrivals. The proposed methodology accurately captures important traffic dynamics, and provides a reliable input for the simulation of highway vehicular networks.**

***Index Terms*—Vehicular networks, simulation, inter-arrivals.**

I. Introduction

Vehicular networks are recognized as a key enabling technology for next-generation Intelligent Transportation Systems, which are expected to improve road safety and ease traffic management. However, field testing of vehicular network solutions is often impractical, as it requires the large-scale deployment of new technologies in highway or urban environments. Simulation is thus widely adopted as a standard approach to the performance evaluation of protocols and architectures dedicated to vehicular environments.

A critical aspect to account for when simulating vehicular networks are road traffic dynamics. In highway environments, it has been recently shown that realistic models of the mobility of vehicles can unveil unique connectivity properties [1]. It is thus important that faithful representations of road traffic are employed in the simulations of highway vehicular networks. Specifically, that requirement mandates (i) a correct microscopic description of the in-flow of vehicles, and (ii) a faithful model of the behavior of individual drivers.

While there exist many validated models for the second aspect above [2], the first has received small attention, as transportation research focused on macroscopic studies [3]. A limited number of networking studies tackled this problem. Bai *et al.* [4] study short-term perturbations in the vehicular traffic flow, including microscopic modeling of headways. However, they consider aggregate inter-arrivals instead of per-lane ones as we do. Moreover, their model consists in a fitting of the empirical distribution: we will show that this approach cannot fully capture the interactions among vehicles within the traffic flow. The last consideration also holds in the case of more recent works, i.e., [5]–[7], all of which generate inter-arrivals from models based on fittings of empirical distributions.

In this letter, we propose a novel technique to generate microscopic inter-arrival times of vehicles on highways, which captures correlations that are lost in distribution-fitted models. Our measurement-based Hidden Markov Model (HMM) can faithfully mimic fine-grained inter-arrivals collected on real-world highways, and, via its parametrization, can represent traffic in free-flow, congested and mixed states. Thus, it can be employed for a reliable simulation of vehicular networks.

Marco Gramaglia and Marco Fiore are with IEIIT-CNR, Turin, Italy (e-mail: *name.surname*@ieiit.cnr.it). Maria Calderon is with the Telematics Department, University Carlos III of Madrid (e-mail: maria@it.uc3m.es).

This work has been funded by the Spanish MICINN through the ADAS-ROAD project (TRA2013-48314), and by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement n.630211. The authors would like to acknowledge the Spanish Directorate General of Traffic for kindly proving us with the empirical traces used in this work.

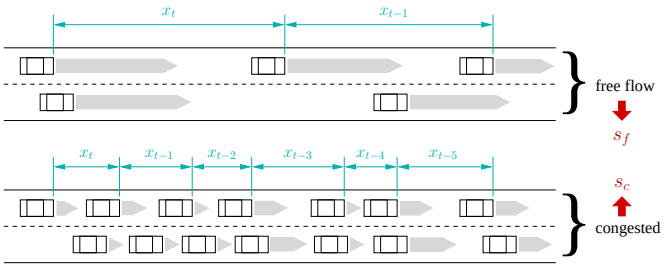


Fig. 1. Reference system: free flow (top) and congested (bottom) traffic.

II. Measurement-driven inter-arrival model

A. System and model formalism

We consider a generic highway segment, where road traffic can present different levels of congestion. Consistently with well-known results from transportation theory, we assume that road traffic conditions range from pure free flow to heavy congestion [3]. In the former case, traffic is so sparse that vehicles travel far apart and do not affect each other's speed, as in the top plot of Fig. 1. In the latter case, traffic is dense enough to force all vehicles to travel at the same low speed, as in the bottom plot of Fig. 1. Clearly, intermediate conditions, such as synchronized traffic [3], can occur.

Previous works have shown that free flow and congested traffic map to different distributions of vehicle inter-arrival times on each lane of a highway segment. Namely, free flow conditions induce exponential inter-arrivals, while congested conditions yield gaussian inter-arrivals. Traffic conditions in between those two extremes result in inter-arrivals that can be fitted to a mixture of the two distributions [5].

Our objective is to build a model capable of generating per-lane synthetic inter-arrival events $\bar{x} = \{x_t\}$ that capture diverse realistic traffic conditions, as portrayed in Fig. 1 in both free flow and congested cases. To that end, we can leverage observations of the real-world system, i.e., measurements of per-lane inter-arrivals collected by induction loops deployed on the highway segment of interest. More formally, system observations are thus sequences $\bar{o} = \{o_t\}$ of inter-arrival times[1] on each highway lane.

[1]For the sake of clarity, we employ two different notations for inter-arrival times. On the one hand, $\bar{x}$ denotes a sequence of *synthetic* inter-arrival times generated by a model, such as the HMM we propose. On the other hand, $\bar{o}$ denotes a sequence of *real-world* inter-arrival times, obtained from experimental measurements and used to formulate and calibrate the model.

A naive model of inter-arrivals is provided by the empirical probability distribution obtained by aggregating the observations in $\bar{o}$. As mentioned above, such a distribution has a mixed exponential-gaussian shape, whose precise mixture depends on the level of traffic congestion. One can extract any number of realizations from the empirical distribution and use them as synthetic samples x_t of the inter-arrivals. This approach assumes independence of the inter-arrivals, i.e., that x_t and x_{t+1} are i.i.d., $\forall t$. However, in real-world traffic, subsequent vehicles influence each other's headway distance and speed, generating temporal correlations among inter-arrivals. Extracting $\bar{x}$ from the empirical probability distribution cannot capture this effect, and thus leads to unreliable results, as we will later demonstrate in Sec. III.

In order to account for the correlations that characterize the vehicular mobility, we model the system as a Hidden Markov Model (HMM) depicted in Fig. 2(a). There, we consider that the road traffic system can be in one of the states of a set $\bar{s} = \{s_i\} = (s_f, s_c)$: state s_f maps to free flow traffic, while state s_c corresponds to congested traffic. Transitions between these two states are regulated by a probability matrix

$$\bar{\bar{a}} = \{a_{ij}\} = \begin{pmatrix} a_{ff} & a_{fc} \\ a_{cf} & a_{cc} \end{pmatrix},$$

which describes the temporal correlations in the traffic state.

The system above is *hidden* from our perspective, as one cannot measure it directly. Instead, each state generates observable inter-arrival times, according to emission probability distributions $\bar{p} = \{p_i(x)\} = (p_f(x), p_c(x))$. There, $p_f(x)$ and $p_c(x)$ describe the probability that free flow and congested traffic conditions, respectively, result into an inter-arrival time of duration x. As already pointed out, we know from the previous studies in [5] that $p_f(x) = \lambda\, exp(-\lambda(x-\kappa))$, and $p_c(x) = (1/\sigma\sqrt{2\pi})\, exp(-(x-\mu)^2/2\sigma^2)$.

Building a HMM that generates realistic inter-arrival time sequences $\bar{x}$ implies calibrating the model in Fig. 2(a). To that end, one needs to assign the correct values to the coefficients in the transition matrix $\bar{\bar{a}}$ and to the parameters of the emission probability distributions $\bar{p}$, i.e., λ, κ, μ, σ. Moreover, as the transient evolution of the model depends on its initial state, it is also important to calibrate the initial state probability distribution $\bar{\pi} = (\pi_f, \pi_c)$, which describes the probability π_f (resp., π_c) that the system starts in free flow (resp., congested) traffic conditions. Next, we discuss how all these values can be inferred from real-world observations.

B. Model calibration

The calibration of the Hidden Markov Model $\mathcal{M} = (\bar{\bar{a}}, \bar{p}, \bar{\pi})$ from a known sequence of real-world observations $\bar{o}$ is performed via a modified version of the iterative Baum-Welch algorithm [8], whose pseudo-code is shown in Alg. 1.

At each iteration, the algorithm first performs *forward-backward* procedures (lines 3–4 of Alg. 1). The *forward* procedure calculates, for each observation $o_t \in \bar{o}$, the probability $\alpha_i(t)$ that o_t is generated while in state $s_i \in \bar{s}$, given the sequence of past observations $(o_1, o_2, \ldots, o_{t-1})$ and the current model $\mathcal{M}$. Conversely, the *backward* procedure calculates the probabilities $\beta_i(t)$ that o_t is generated while in state $s_i \in \bar{s}$, given the sequence of observations that follow [2] o_t, $(o_{t+1}, o_{t+2}, \ldots, o_T)$, and the current model $\mathcal{M}$.

Both procedures generate trellis-like structures, as portrayed in Fig. 2(b). Let us focus on the forward procedure, illustrated by the top structure in the figure[3]. Starting from the initial probabilities π_f and π_c, the procedure multiplies them by the probabilities $p_f(o_1)$ and $p_c(o_1)$ that the first observation o_1 is actually generated while in s_f and s_c, respectively. The results are normalized so as to obtain the probability distribution described by $\alpha_f(1)$ and $\alpha_c(1)$. Then, $\alpha_f(1)$ and $\alpha_c(1)$ are combined with the probabilities to remain in the current state or transit to the other, according to the matrix $\bar{\bar{a}}$. E.g., there are probabilities a_{ff} (assuming initial free flow traffic) and a_{cf} (assuming initial congested traffic) that the next state is still s_f (red arrows in Fig. 2(b)): the result is then multiplied by the probability $p_f(o_2)$ that o_2 is generated in s_f. The process is the same for state s_c (cyan arrows in Fig. 2(b)), and the resulting values are normalized again so that $\alpha_f(2)$ and $\alpha_c(2)$ are a probability distribution. The operation is then repeated at each step. Formally, at the generic step t, $\alpha_i(t)$ (with $t > 1$ and i such that $s_i \in \bar{s}$) is computed as

$$\alpha_i(t) = P(o_1, \ldots, o_T, q_t = s_i | \mathcal{M}) = \frac{\tilde{\alpha}_i(t)}{\tilde{\alpha}_f(t) + \tilde{\alpha}_c(t)}\ , \quad (1)$$

where q_t is the state of the system at step t, and

$$\tilde{\alpha}_f(t) = p_f(o_t)\ [\alpha_f(t-1)\, a_{ff} + \alpha_c(t-1)\, a_{cf}]\ , \quad (2)$$

$$\tilde{\alpha}_c(t) = p_c(o_t)\ [\alpha_c(t-1)\, a_{cc} + \alpha_f(t-1)\, a_{fc}]\ . \quad (3)$$

The resulting probabilities are collected into two $T \times 2$ matrices $\bar{\bar{\alpha}} = \{\alpha_i(t)\}$ and $\bar{\bar{\beta}} = \{\beta_i(t)\}$. These are then used to derive two additional measures, for each time step (lines 5–9 of Alg. 1). The first measure (line 7 of Alg. 1) is the probability $\gamma_i(t)$ of being in state $s_i \in \bar{s}$ at step t, given the observation

[2]Hereinafter, we will denote as T the index of the last observation in $\bar{o}$.

[3]The backward procedure, in the bottom structure of Fig. 2(b), is symmetrical to the forward one, and its discussion is omitted for the sake of brevity.

Algorithm 1 Modified Baum-Welch algorithm.

1: **procedure** $\mathbf{BaumWelch}(\bar{o},\bar{\bar{a}},\bar{p},\bar{\pi})$
2: **do**
3: $\bar{\bar{\alpha}} \leftarrow$ `forward` $(\bar{o}, \bar{\bar{a}}, \bar{p}, \bar{\pi})$
4: $\bar{\bar{\beta}} \leftarrow$ `backward` $(\bar{o}, \bar{\bar{a}}, \bar{p})$
5: **for** $t = 1 \rightarrow T$ **do**
6: **for** $i = s_f \rightarrow s_c$ **do**
7: $\gamma_i(t) \leftarrow$ `calcGamma`$(\bar{\bar{\alpha}}, \bar{\bar{\beta}}, \bar{o}, t, i)$
8: **for** $j = s_f \rightarrow s_c$ **do**
9: $\xi_{ij}(t) \leftarrow$ `calcXi`$(\bar{\bar{\alpha}}, \bar{\bar{\beta}}, \bar{p}, \bar{o}, t, i, j)$
10: **for** $i, j = s_f \rightarrow s_c$ **do**
11: $a_{ij} \leftarrow$ `updateA`$(\bar{\bar{\gamma}}, \bar{\bar{\xi}})$
12: **for** $i = s_f \rightarrow s_c$ **do**
13: $\pi_i \leftarrow \gamma_i(1)$
14: $p_f(x) \leftarrow$ `mleExp`$(\bar{o}, \bar{\bar{\gamma}})$
15: $p_c(x) \leftarrow$ `mleGauss`$(\bar{o}, \bar{\bar{\gamma}})$
16: $\Lambda_{prev} \leftarrow \Lambda$
17: $\Lambda \leftarrow$ `calcLikelihood` $(\bar{o}, \bar{\bar{a}}, \bar{p}, \bar{\pi})$
18: **while** $\Lambda \neq \Lambda_{prev}$
19: **return** $\bar{\bar{a}}, \bar{p}, \bar{\pi}$

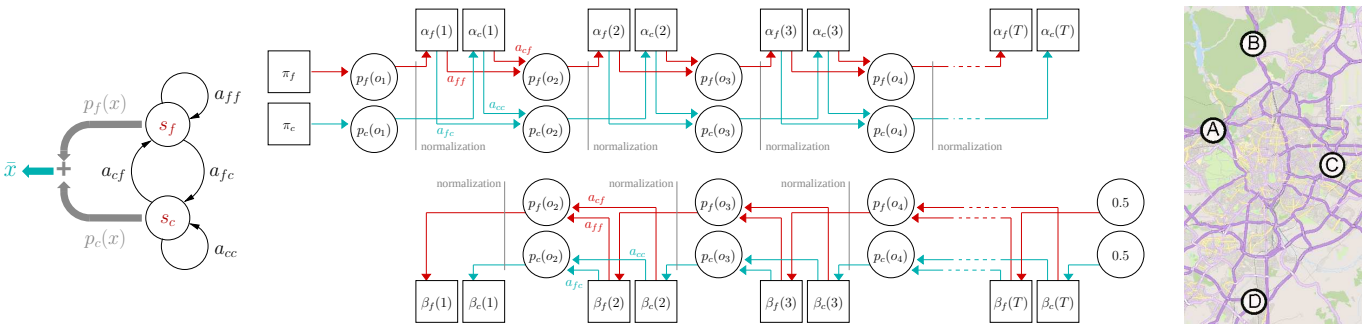


Fig. 2. Inter-arrival modeling and evaluation. (a) Graph representation of the Hidden Markov Model. (b) Trellis-like structures associated to the forward and backward procedures of the modified Baum-Welch algorithm. (c) Geographical localization of the real-world measurement points on four highways.

sequence $\bar{o}$ and the current model $\mathcal{M}$. Formally,

$$\gamma_i(t) = P\left(q_t = s_i | o_1, \ldots, o_T, \mathcal{M}\right) = \frac{\alpha_i(t)\beta_i(t)}{P\left(\bar{o}|\mathcal{M}\right)}\,, \quad (4)$$

where $P\left(\bar{o}|\mathcal{M}\right)$ is the probability of observing the sequence of inter-arrivals $\bar{o}$ according to the model $\mathcal{M}$.

The second measure (line 9 of Alg. 1) is the probability $\xi_{ij}(t)$ of being in state $s_i \in \bar{s}$ at step t and in state $s_j \in \bar{s}$ at the following step $t+1$, given $\bar{o}$ and $\mathcal{M}$. Formally,

$$\begin{aligned}\xi_{ij}(t) &= P\left(q_t = s_i, q_{t+1} = s_j | o_1, \ldots, o_T, \mathcal{M}\right) \\ &= \frac{\alpha_i(t)\, a_{ij}\, p_j(o_{t+1})\, \beta_i(t+1)}{P\left(\bar{o}|\mathcal{M}\right)}\,,\end{aligned} \quad (5)$$

where $p_j(o_{t+1})$ is the emission probability of observation o_{t+1} by state $s_j \in \bar{s}$, i.e., $p_f(x)$ or $p_c(x)$ depending on j.

Using $T \times 2$ matrices $\bar{\bar{\gamma}} = \{\gamma_i(t)\}$ and $\bar{\bar{\xi}} = \{\xi_{ij}(t)\}$, the elements of the transition probability matrix $\bar{a}$ are updated (lines 10–11 of Alg. 1) as

$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_{ij}(t)}{\sum_{t=1}^{T-1} \gamma_i(t)} \quad \forall i, j \;\text{ s.t. }\; s_i, s_j \in \bar{s}. \quad (6)$$

There, the sum over all steps of $\gamma_i(t)$ yields the overall probability of being in state s_i given the observation sequence $\bar{o}$ and the current model $\mathcal{M}$. Similarly, the sum over all steps of $\xi_{ij}(t)$ is the overall transition probability from s_i to s_j, both of which belong to $\bar{s}$, given the observation sequence $\bar{o}$ and the current model $\mathcal{M}$.

Also, since $\gamma_i(1)$ is the probability of being in state s_i at time step 1, the initial state probability distribution $\bar{\pi}$ is updated (lines 12–13 of Alg. 1) as

$$\pi_i = \gamma_i(1) \quad \forall i \;\; s.t. \;\; s_i \in \bar{s}. \quad (7)$$

Having adjusted $\bar{a}$ and $\bar{\pi}$, the last parameters of model $\mathcal{M}$ that need updating are the emission probabilities in $\bar{p}$. However, the standard Baum-Welch algorithm only considers discrete emissions by each state, whereas, in our case, emissions are the output of continuous random variables. In order to account for that aspect, we extend the original algorithm and infer the values of λ, μ and σ[4] by using maximum likelihood estimators (MLEs, in lines 14–15 of Alg. 1). We remark that:

[4]The value of κ, i.e., the last parameter in $\bar{p}$, is updated by sweeping the range $[0, 3]$ s and picking the value that yields the best log-likelihood.

TABLE I
HMM CALIBRATION IN FOUR REPRESENTATIVE SCENARIOS.

Id	Road	Day/Time	$[\lambda, \kappa]$	$[\mu, \sigma]$	$[a_{ff}, a_{fc}, a_{cf}, a_{cc}]$	Traffic
1	B, C	2, peak	$[0.34, 1.3]$	$[0.94, 0.24]$	$[0.68, 0.32, 0.69, 0.31]$	68%, 32%
2	A, L	1, off-peak	$[0.28, 1.8]$	$[1.15, 0.37]$	$[0.62, 0.38, 0.48, 0.52]$	56%, 44%
3	D, R	2, peak	$[0.22, 2.1]$	$[1.28, 0.46]$	$[0.47, 0.53, 0.36, 0.64]$	40%, 60%
4	C, L	3, peak	$[0.27, 1.7]$	$[1.06, 0.36]$	$[0.40, 0.60, 0.23, 0.77]$	28%, 72%

(i) different states are associated to different emission probabilities distributions, thus the actual MLE employed depends on $s_i \in \bar{s}$; (ii) since $\gamma_i(t)$ is the probability of being in state s_i at time t, the MLEs are calculated on the observations $\bar{o}$ weighted by the corresponding $\gamma_i(t)$'s.

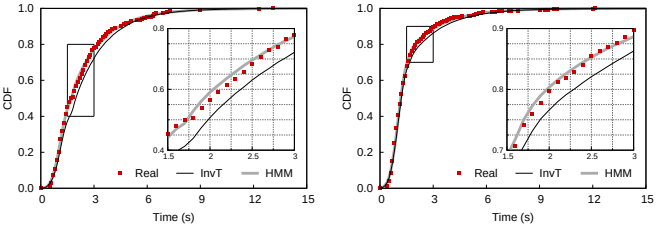
By iterating over these steps, the algorithm converges to the local maximum of the likelihood Λ of the system parameters [8], at which point it is stopped (lines 16–18 of Alg. 1).

III. Performance Evaluation

To calibrate and validate the HMM, we employ fine-grained real-world datasets of vehicle inter-arrival times collected at four measurements locations on peri-urban highways near a major European city, represented by points A to D in Fig. 2(c). The highways feature three lanes, tagged as right (R), center (C) and left (L) in the following. The data is collected via induction loops, especially configured to provide fine-grained information on the time, speed and lane of each passing vehicle. Measurements were performed on four weekdays (Fri 7, and Mon 10 to Wed 12, on May 2010) numbered from 1 to 4. In each such day, two 30-minute datasets were collected, one during the morning traffic peak, and one during off-peak hours. These datasets allow to evaluate the HMM in presence of heterogeneous road traffic conditions.

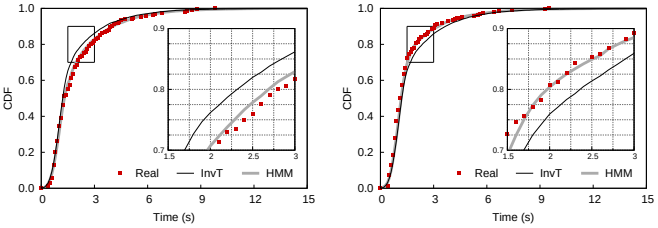
We use the available mobility datasets to train the Hidden Markov Model presented in Sec. II-A according to the algorithm described in Sec. II-B. The headways between vehicles are used as observations $\bar{o}$, for any combination of road, lane, day, and hour at which measurement were taken. For each such scenario, the calibrated HMM is used to generate synthetic inter-arrivals (*"HMM"* in the results), and compared against:

- the original inter-arrivals collected via induction loops in the real world (*"Real"* in the results);
- instances of the mixed exponential-gaussian distribution that best fits each real-world sequence $\bar{o}$, obtained by means of the Inverse Transform algorithm. This is the



(a) Scenario 2, subset $\mathcal{E}$ (b) Scenario 4, subset $\mathcal{G}$

Fig. 3. Sample inter-arrival distributions, first-order correlation.



(a) Scenario 1, subset $\mathcal{EG}$ (b) Scenario 1, subset $\mathcal{GE}$

Fig. 4. Sample inter-arrival distributions, second-order correlation.

naive modeling approach anticipated in Sec. II-A, and it is used as a benchmark (*"InvT"* in the results).

The HMM calibration results are reported in Tab. I, for four representative cases (average speeds of 107.9, 92, 104 and 102.5 Km/h). Although experiments were run for the whole 96 cases. with very similar performance, due to space limitations we here restrain our discussion to these four traces. The latter were selected so as to cover a sensible range of traffic situations: the last column in Tab. I shows the traffic conditions in each scenario, expressed as the percentages of time where free flow and congested traffic is observed within the trace. We can observe that scenarios 1 to 4 are heterogeneous and yield mostly-free-flow to mostly-congested traffic, in that order.

As a baseline test, we evaluate the capability of the HMM and InvT approaches to generate inter-arrivals that faithfully reproduce the original distribution of Real inter-arrivals. As expected, both models successfully pass the Mann-Whitney U test in Tab. II, confirming that the inter-arrivals come from identical distributions. We remark however that the z-score of the HMM is generally better than that of the InvT.

In any case, the aggregate distribution hides temporal correlations among inter-arrivals. In order to explore the impact of such correlations, we first focus on first-order dependencies, i.e., on the relationship between any two subsequent inter-arrivals. To that end, we divide the Real, InvT and HMM inter-arrival populations into two subsets

$$\mathcal{G} = \{o_t | o_{t-1} \geq \tilde{o}\} \text{ and } \mathcal{E} = \{o_t | o_{t-1} < \tilde{o}\}, \ \forall t, \quad (8)$$

where $\tilde{o}$ is a threshold value, set for each scenario so that it yields the best separation between the exponential and gaussian portions of the original mixed distribution[5]. The set $\mathcal{G}$ thus contains inter-arrivals measured after a gaussian event, whereas $\mathcal{E}$ contains those occurring after an exponential event.

By drawing the distributions of inter-arrivals in $\mathcal{G}$ and $\mathcal{E}$ separately, we observe that the InvT curves can significantly differ from the Real ones, as in the samples in Fig. 3. This is a clue that first-order correlations exists among inter-arrivals, since the value of an inter-arrival is affected by that of the previous event. The InvT approach neglects such correlations, and cannot generate values that capture this phenomenon. The HMM yields instead distributions that are nearly identical to the Real ones in Fig. 3. A more rigorous proof is provided by the results of the Mann-Whitney U test in Tab. II. Not

[5]We remark that the origin of the exponential component k is typically not much smaller than $\mu + 3\sigma$ (see, e.g., Tab. I), the latter denoting the value after which gaussian inter-arrivals are unlikely. In other words, the two portions are disjoint enough that a separation $\tilde{o}$ can be clearly identified in most cases.

TABLE II
AVERAGE Z-SCORES FOR THE MANN-WHITNEY U TEST.

Id	Baseline		First order				Second order			
	HMM	InvT	HMM		InvT		HMM			
			$\mathcal{E}$	$\mathcal{G}$	$\mathcal{E}$	$\mathcal{G}$	$\mathcal{EE}$	$\mathcal{EG}$	$\mathcal{GE}$	$\mathcal{GG}$
1	0.46 ✓	0.24 ✓	0.09 ✓	1.67 ✓	0.23 ✓	1.08 ✓	0.47 ✓	0.53 ✓	0.48 ✓	0.47 ✓
2	0.13 ✓	0.33 ✓	0.11 ✓	0.14 ✓	2.38 ✗	1.62 ✓	0.13 ✓	0.13 ✓	0.14 ✓	0.14 ✓
3	0.10 ✓	0.30 ✓	0.17 ✓	0.17 ✓	1.60 ✓	1.32 ✓	0.19 ✓	0.20 ✓	0.21 ✓	0.19 ✓
4	0.64 ✓	1.01 ✓	0.62 ✓	0.39 ✓	2.90 ✗	2.50 ✗	0.12 ✓	0.17 ✓	0.14 ✓	0.13 ✓

only the z-scores are better in the HMM case then in the InvT case, but in some scenarios the InvT method generates distributions such that the null hypothesis is rejected, i.e., that are not representative of the Real ones.

At this point, a rightful question is whether correlations at orders higher that the first exists among inter-arrival times, and if the HMM model is sufficient to capture them. To address this aspect, we consider second-order dependencies among triplets of subsequent inter-arrivals, and separate four subsets, $\mathcal{EE}$, $\mathcal{EG}$, $\mathcal{GE}$, and $\mathcal{GG}$, depending on the nature of the past two inter-arrivals. The results in Fig. 4 and in Tab. II show that second-order correlations are weak, and that the HMM model always passes the Mann-Whitney U test – thus providing an reliable approximation of real-world inter-arrival dynamics.

IV. CONCLUSIONS

We proposed a measurement-based Hidden Markov Model of vehicle inter-arrivals on highways. The model can generate temporally correlated inter-arrivals that are coherent with those observed in the real world. As a result, our proposal is especially suitable for the generation of reliable inter-arrivals in network simulations.

REFERENCES

[1] M. Gramaglia, O. Trullols-Cruces, D. Naboulsi, M. Fiore, M. Calderon, *"Vehicular Networks on Two Madrid Highways,"* IEEE SECON, 2014.
[2] R.E. Wilson, J.A. Ward, *"Car-following models: fifty years of linear stability analysis – a mathematical perspective"*, Transportation Planning and Technology, Taylor & Francis, 2011.
[3] B.S. Kerner, *"The physics of traffic: empirical freeway pattern features, engineering applications, and theory"*, Springer, 2004.
[4] F. Bai and B. Krishnamachari, *"Spatio-temporal variations of vehicle traffic in VANETs: facts and implications"*, ACM VANET, 2009.
[5] M. Gramaglia, P. Serrano, J. Hernandez, M. Calderon, C. J. Bernardos, *"New insights from the analysis of free flow vehicular traffic in highways"*, IEEE WoWMoM, 2011.
[6] L. Cheng and S. Panichpapiboon, *"Effects of intervehicle spacing distributions on connectivity of VANET: a case study from measured highway traffic"*, IEEE Communications Magazine, 2012.
[7] N. Akhtar, S. Coleri Ergen, and O. Ozkasap, *"Vehicle Mobility and Communication Channel Models for Realistic and Efficient Highway VANET Simulation"*, IEEE Transactions on Vehicular Technology, 2014.
[8] L.E. Baum, T. Petrie, *"Statistical Inference for probabilistic functions of finite state Markov chains"*, Ann. Math. Stat, 1966.