

Regular Paper

MC-TES: An Efficient Mobile Phone Based Context-Aware Traffic State Estimation Framework

QUANG TRAN MINH^{1,a)} MUHAMMAD ARIFF BAHARUDIN¹ EIJI KAMIOKA¹

Received: April 6, 2012, Accepted: October 10, 2012

Abstract: This paper proposes a notable mobile phone based context-aware traffic state estimation (MC-TES) framework whereby the essential issues of low and uncertain penetration rate are thoroughly resolved. A novel intelligent context-aware velocity-density inference circuit (ICIC) and a practical artificial neural network (ANN) based prediction approach are proposed. The ICIC model not only improves the traffic state estimation effectiveness but also minimizes the critical penetration rate required in the mobile phone based traffic state estimation (M-TES). The ANN-based prediction approach is considered as a complement of the ICIC in cases of an unacceptably low or unknown penetration rate. In addition, the difficulty in selecting the “right” traffic state estimation model, namely among the ICIC and the ANN, under the condition of an uncertain penetration rate is resolved. The experimental evaluations confirm the effectiveness, the feasibility as well as the robustness of the proposed approaches. As a result, this research contributes to accelerating the realization of mobile phone-based intelligent transportation systems (M-ITSs) or of the M-TES systems in specific.

Keywords: mobile phone based ITS, traffic state estimation, mobile context-aware, penetration rate, inference circuit

1. Introduction

Transportation and road traffic are important parts of any economy all over the world. Therefore, research in ITS has attracted a numerous number of researchers in various fields including transportations, civil engineering, statistical study, computational science, communication engineering, and so forth. However, besides advanced achievements in ITS, traffic congestion still remains as a serious issue in modern cities. Traffic jam is not only the cause of economic loss but also the source of pollution (air, noise pollution, etc.), violence and other social issues. The Urban Mobility Report [1] reported in 2007 that traffic congestion causes 4.2 billion hours of extra travel time requiring 2.9 billion extra gallons of fuel, which cost the US tax-payers an additional \$78 billion [2]. The Ministry of Land Infrastructure and Transport of Japan reported in 2006 that the economic loss caused by traffic jams is around \$100 billion annually [3]. In addition, the situations where ambulances are hopelessly stuck on the way to hospitals; shops along the road sides have to be closed; students, teachers, workers, officers cannot go to school/work in time because of traffic jams, are not unusual in big cities. Such uncomfortable and even dangerous traffic environments lower the citizen’s quality of life (QoL).

The advances of mobile phone technologies catalyze researches on M-ITS by which mobile phones are utilized as traffic probes [4], [5], [6], [8], [9]. Since mobile phones are available everywhere and the mobile phone network has already been deployed, this approach is advantageous in coverage, real-time ef-

fect, investment and maintenance cost. These advantages help to accelerate the realization of M-TES systems, thus realize a safe and green (no traffic jam) traffic environment.

However, in mobile phone based traffic state estimation systems such as the MC-TES, the traffic state is estimated based on individual data reported by mobile phones. Obviously, the more data are collected, the higher accuracy the estimation is, and vice versa [2], [10]. In the road-side fixed sensor traffic state estimation systems, the presence, velocity, image, etc., of any vehicle is almost successfully obtained when it passes over a road-side fixed sensor such as a loop detector [11], [12], a RFID reader [13], [14], [15], a video camera [16], [17], [18], and so forth. In contrast, there is no way to compel every mobile phone user to report the data to the estimation server. That is to say, the penetration rate, namely the number of vehicles that report data to the estimation server out of the total number of vehicles traveling in the considered road segment, is commonly low, especially when the system has just been launched, which affects the traffic state estimation accuracy significantly. In addition, the “actual” penetration rate is uncertain at the estimation time implying difficulties in employing optimization models for traffic state estimation. Therefore, low and uncertain penetration rate issues are the most difficulties impeding the realization of mobile phone based traffic state estimation systems. Unfortunately, to the best of our knowledge, there is no relevant research investigating these issues thoroughly. Instead, existing researches assumed that the penetration rate is always relevant [5], [6], [8]. This article aims at investigating and solving these issues with primary contributions as follows:

- Investigating the effect of a *low* and *uncertain* penetration rate

¹ Shibaura Institute of Technology, Koto, Tokyo 135–8548, Japan

^{a)} m709504@shibaura-it.ac.jp

on traffic state estimation.

- Proposing a notable intelligent context-aware velocity-density inference circuit (ICIC) whereby useful contexts extracted from the data reported by mobile devices are utilized to accurately estimate traffic state even in cases of low penetration rates.

- Proposing a practical artificial neural network (ANN) based prediction approach to ensure the effectiveness of the MC-TES model when the penetration rate becomes unacceptably low, namely just a few percent or even 0%.

- Proposing a simple yet effective traffic state estimation model selection method under the condition of an *uncertain* penetration rate.

This paper is organized as follows: Section 2 reviews the related work. The system model and the problem formulation are presented in Section 3. Sections 4 and 5 propose the ICIC and the ANN-based prediction approaches, respectively. An effective traffic state estimation model selection method under the condition of an uncertain penetration rate is proposed in Section 6. Section 7 thoroughly evaluates the effectiveness, the feasibility as well as the robustness of the proposed solutions. Section 8 concludes this work and draws out future research directions.

2. Related Work

Existing TES systems such as VICS [19], NAVITIME [20] in Japan, the ITS project at Kansas, USA [21] primarily rely on road-side fixed sensors for traffic data collection. However, the road-side fixed sensor technologies such as loop detector [11], [12], RFID [13], [14], video camera [16], and so forth, disclosed their essential weakness in terms of coverage limitation as well as investment cost. The ad-hoc network technology [3], [22] theoretically helps to improve the coverage but it has not matured enough for real-world applications.

The Mobile Millennium Project (MMP) [4] is closely related to this work which employs GPS-enabled mobile phones as traffic probes. However, the MMP estimates the traffic state by analyzing vehicle flows on road networks using a dynamical theory. The dynamical theory may work effectively in an environment of narrow and short flows but it may reveal serious errors when being applied to a complex environment such as wide road networks. Moreover, the issue of a low penetration rate was not discussed thoroughly.

Another fundamental requirement for a TES system is to accurately estimate the traffic state level at each considered road segment. J. Yoon et al. [5], proposed to divide the road network into separate road segments and the traffic state was estimated based on complete traces of participating vehicles. This approach, however, could not be applied in the cases of serious congestions where no complete traffic trace can be obtained. Our previous works in Refs. [6], [8] where the traffic state could be granularly quantified using the data reported from mobile phones constitute the foundations for this research. However, these works still assumed that every mobile phone sends data to the estimation server, thus the issue of a low penetration rate was still left as an open research question.

In addition, in the M-TES, controlling the data transmission

load while keeping the completeness of data collection (no missing of useful data) is crucial in assuring the performance of the whole estimation model. The work in Ref. [7] proposed to pre-set Virtual Trip Lines (VTLs) on the road network by which the traffic data is collected and reported only when a vehicle passes a VTL. Obviously, this approach reduces the number of data collections, thus alleviates the data transmission load. However, the useful data may be missed since there is no relation between VTLs and the places where a traffic congestion actually occurs. Moreover, the VTL setting criteria such as where on the road a VTL should be set, how far should be the two consecutive VTLs, and so forth, are matters of argument. The work in Ref. [8] proposed the so-called “3R” approach by which only the Right data is collected at the Right time by the Right mobile devices. For example, the mobile phones carried by walkers are refrained from reporting their data since walkers do not affect the traffic state. While this approach alleviates both the data transmission load and the data redundancy, it might be affected significantly by the *low* and *uncertain* penetration rate issues.

The work in Ref. [2] focused on ensuring the traffic estimation effectiveness even if the penetration rate becomes unacceptably low. A statistical learning model based on historical data was proposed to estimate/forecast the traffic state in terms of *travel time* and *congestion state*. The work stated that the logistic regression model works effectively even if the penetration rate is quite low (the estimation accuracy was higher than 70% even if the penetration rate is definitely low as 5%). However, several issues were left which need to be clarified. First, the work utilized the VTL concept proposed in Ref. [7] while, as mentioned before, the effectiveness of the VTL itself is still a matter of argument. Second, the work did not consider the density in estimating the traffic state. Third, the congestion state was defined as a “binary” indicator accepting only two states, namely “*congested*” and “*not-congested*”. This setting may bias the estimation accuracy since any “blind” guessing approach can also reach 50% of accuracy. Fourth, the Paramics simulator [23] was employed to generate synthetic data for evaluations. This simulator gave information about every vehicle. To imitate a low penetration rate dataset, namely 5% for example, a large portion of data (95%) was removed. In fact, this process could not generate the appropriate low penetration rate dataset as it was defined in the work. Therefore, the estimation error with regards to the penetration rate should be clarified.

Our previous works in Refs. [10], [24] developed a foundation on solving the issues of *low* and *uncertain* penetration rate. These works investigated the effect of a low penetration rate on the estimation effectiveness. In addition, two “velocity-density inference circuits” (VDICs), namely the adaptive and the adaptive feedback VDICs were proposed to improve the M-TES’s effectiveness even in cases of a low penetration rate. In the VDICs, both the average velocity and the density which are calculated directly from the *sensed* data (i.e., the GPS data obtained by mobile phones) and inferred directly using the Greenshields model [25] were utilized. However, the impact coefficients of the corresponding parameters in the whole estimation model were not optimized. This article proposed a sophisticated ICIC model, whereby useful contexts

extracted from the data reported by mobile devices are utilized to optimize the impact coefficients, thus optimize the whole estimation model. The details of this framework will be presented in the remainder of this paper.

3. System Modeling and Problem Formulation

3.1 System Modeling: the M-TES

Considering a road network with a total of N road segments, the set of all road segments is denoted as $V = \{i | i = 1 \dots N\}$. For any road segment $i \in V$, traffic data is available at any time t . However, the obtained GPS data is the event-based data which cannot be directly transformed into a traffic state. Therefore, the traffic state should be aggregated in predefined time intervals, namely in t -second windows. Concretely, the traffic state is estimated at times $k = 0, t, 2t, \dots$, etc. The task here is to effectively estimate the traffic state of the considered road segment i at time interval k based on the data reported by mobile devices.

Obviously, the velocity and the density directly reflect the traffic state of a road segment. In this work, a notable traffic state quantification (TSQ) model is proposed by which the velocity and the density are independently and directly estimated using the *sensed* data reported by mobile phones before being integrated to quantify the traffic state.

Definition 1: The *average velocity* of a traffic flow in the road segment i during time k , denoted as $V_{Avg}^{k,i}$, is the average velocity of all vehicles traveling in the considered road segment during time k .

The average velocity defined above can be formally expressed in Eq. (1). Here, $V_{m,j}^{k,i}$ is the velocity of any individual vehicle j ($j = 1 \dots q$) detected at time t_m ($m = 1, 2, 3 \dots, r$) during time interval k ($[k-1]t \leq t_m < kt$), q is the total number of vehicles, and r is the total number of detection times during time interval k .

$$V_{Avg}^{k,i} = \frac{\sum_{j=1}^q \sum_{m=1}^r V_{m,j}^{k,i}}{qr}, (k-1)t \leq t_m < kt \quad (1)$$

Since the limited speed varies from road segments to road segments, the absolute average velocity defined above may not appropriately represent the traffic state in terms of travel time. In order to solve this issue, a new term, namely the *mean speed capacity*, is proposed.

Definition 2: The *mean speed capacity* of the road segment i during time k , denoted as $M_V^{k,i}$, is defined in Eq. (2), where V_{max}^i is the limited speed of the road segment i .

$$M_V^{k,i} = \frac{V_{Avg}^{k,i}}{V_{max}^i} \quad (2)$$

Observation 1: The higher the M_V is, the better the traffic state is, and vice versa. In this work, the threshold of M_V is set to 0.6 for a good traffic condition in term of travel times [8].

Definition 3: The *density* at the road segment i during time k , denoted as $D^{k,i}$, is the fraction of the number of the vehicles traveling through the considered road segment during time k out of the capacity of the considered road segment. The density is defined in Eq. (3)

$$D^{k,i} = \frac{q^{k,i}}{C^{k,i}} \quad (3)$$

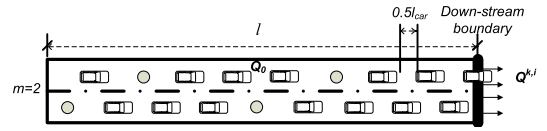


Fig. 1 Capacity, $C^{k,i} = Q_0^i + Q^{k,i}$, of a road segment i with 2 lanes.

Here, $q^{k,i}$ is the total number of vehicles traveling through the road segment i during time k which is estimated based on the sensed data reported by mobile phones, and $C^{k,i}$ is the *flow capacity* of the road segment i .

Definition 4: The *flow capacity*, denoted as $C^{k,i}$, is the maximum number of vehicles which can pass through the road segment i during time k under the best traffic condition. The *flow capacity* can be calculated in Eq. (4).

$$C^{k,i} = Q_0^i + Q^{k,i} \quad (4)$$

In this equation, Q_0^i is the maximum number of vehicles that can be arranged (without moving) in the road segment i , and $Q^{k,i}$ is the volume of the traffic flow passing the downstream boundary of the road segment i during time k in the best traffic condition. These parameters are calculated in Eqs. (5) and (6), respectively.

$$Q_0^i = m \frac{l}{1.5l_c} \quad (5)$$

$$Q^{k,i} = m \frac{k}{\bar{t}} = mk \frac{V_{max}^i}{1.5l_c} \quad (6)$$

Here, m is the number of lanes, l is the length of the road segment i , and l_c is the average length of a car [8], [26]. The value of 1.5 is the coefficient describing the space which must be yielded between two cars in the worst congested area. In Eq. (6), \bar{t} is the average elapse time between two consecutive vehicles, namely vehicles i^{th} and $(i+1)^{th}$, passing the downstream boundary. **Figure 1** illustrates these parameters.

For convenience in estimating the traffic state in the manner of *density*, a new term called the *free space ratio* is proposed.

Definition 5: The *free space ratio* of the road segment i during time k , denoted as $\sigma_s^{k,i}$, is calculated in Eq. (7), where $D^{k,i}$ and $C^{k,i}$ were defined in Eqs. (3) and (4), respectively.

$$\sigma_s^{k,i} = \frac{C^{k,i} - D^{k,i}}{C^{k,i}} = 1 - \frac{D^{k,i}}{C^{k,i}} \quad (7)$$

Observation 2: The higher the *free space ratio* is, the better the traffic state is, and vice versa. In this work, the threshold of σ_s is set to 0.4 for a good traffic condition [8].

After being obtained, $M_V^{k,i}$ and $\sigma_s^{k,i}$ should be integrated in an appropriate way to quantify the traffic state. Here, a term called the *goodness value* which can be used to quantify the traffic state is defined.

Definition 6: The *goodness value* of the road segment i during the estimation time k , denoted as $G^{k,i}(M_V^{k,i}, \sigma_s^{k,i})$, is calculated in Eq. (8), where M_{V0} and σ_{s0} are the thresholds of *mean speed capacity* and *free space ratio*.

$$G^{k,i}(M_V^{k,i}, \sigma_s^{k,i}) = (M_V^{k,i} - M_{V0}) + (\sigma_s^{k,i} - \sigma_{s0}) \quad (8)$$

Observation 3: The *goodness value* is a continuous value, ranging from -1 to 1 , representing from the worst to the best

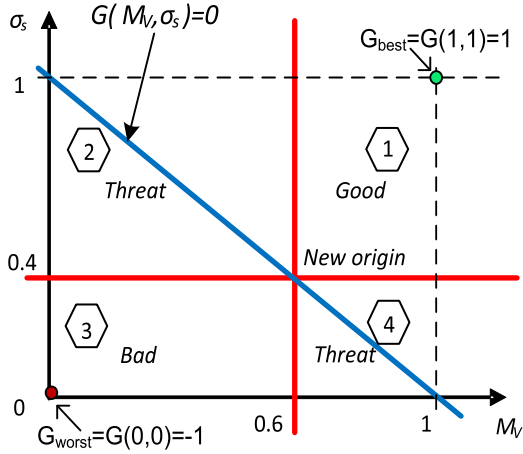


Fig. 2 The goodness value used to quantify traffic state.

traffic states. It is quite adequate for granularly comparing traffic state levels.

Here, $M_v^{k,i}$, $\sigma_s^{k,i}$ and their thresholds, namely M_{v0} and σ_{s0} form the so-called “traffic state quadrant space” as shown in Fig. 2. The original coordination is formed by 2 axes represented by $M_v^{k,i}$ and $\sigma_s^{k,i}$. The values in each axis range from 0 to 1 representing from the worst to the best traffic conditions in corresponding manners, namely in travel time and in density, respectively. The traffic state quadrant space is created by two new axes that are orthogonal to the original axes at the threshold values M_{v0} and σ_{s0} (i.e., 0.6 and 0.4). The 1st quadrant represents good traffic states while the bad traffic states fell in the 3rd one. The 2nd and 4th quadrants represent the threat traffic conditions. More concretely, the 2nd quadrant expresses situations where few vehicles travel slowly due to some special physical conditions such as road construction, bad weather, etc. The 4th quadrant represents the opposite situations, that is to say, at a good speed but high density.

Obviously, the goodness value (as defined above) of each traffic state can be concretely calculated. This value can be used to granularly compare the traffic state levels of different road segments. For example, the goodness values of the worst traffic state (presented by the original origin, namely at the point (0,0)) and of the best traffic state (presented by the rightmost upward point on the quadrant space, namely at the point (1,1)) are calculated in Eqs. (9) and (10), respectively.

$$G_{worst} = G(0, 0) = (0 - 0.6) + (0 - 0.4) = -1 \quad (9)$$

$$G_{best} = G(1, 1) = (1 - 0.6) + (1 - 0.4) = 1 \quad (10)$$

3.2 Problem Formulation with a Low and Uncertain Penetration Rate Consideration

In the M-TES, the traffic state is estimated based on the traffic data reported by mobile phones. However, in practice it is not necessary that every mobile phone reports data to the estimation server. In addition, there is no way to know exactly the actual portion of vehicles (out of the total number of participating vehicles) that report the data. Therefore, the M-TES faces the essential issues of a *low* and *uncertain* penetration rate.

Definition 7: The *penetration rate* at the road segment i during time k , denoted as $\rho^{k,i}$ in Eq. (11), is the fraction of vehicles that

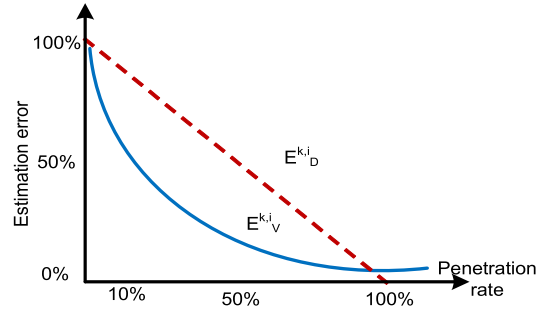


Fig. 3 Effect of the penetration rate on velocity and density estimations.

report data to the estimation server (p) out of the total number of vehicles traveling through the considered road segment (q).

$$\rho^{k,i} = \frac{p}{q} \quad (11)$$

For example, as shown in Fig. 1, 4 vehicles (denoted as circles) among 18 ones do not report data to the server revealing the penetration rate of 14/18 ($\approx 77.7\%$).

Given a penetration rate $\rho^{k,i}$, the average velocity estimation model described in Eq. (1) must be replaced by Eq. (12). Here, $V_{Avg}^{k,i}$, q and r were defined in Eq. (1).

$$V_{Avg}^{k,i,\rho^{k,i}} = \frac{\sum_{j=1}^{\rho^{k,i}q} \sum_{m=1}^r V_{t_{m,j}}^{k,i}}{\rho^{k,i}qr}, (k-1)t \leq t_m < kt \quad (12)$$

Consequently, the average velocity estimation error, $E_V^{k,i}$, can be expressed in Eq. (13), where $V_{Avg}^{k,i}$ is the “actual” average velocity estimated when every vehicle reports data (Eq. (1)), and $V_{Avg}^{k,i,\rho^{k,i}}$ is the average velocity estimated under the given penetration rate $\rho^{k,i}$ (Eq. (12)).

$$E_V^{k,i} = \left| 1 - \frac{V_{Avg}^{k,i,\rho^{k,i}}}{V_{Avg}^{k,i}} \right| = \left| 1 - \frac{1}{\rho^{k,i}} \cdot \frac{\sum_{j=1}^{\rho^{k,i}q} \sum_{m=1}^r V_{t_{m,j}}^{k,i}}{\sum_{j=1}^q \sum_{m=1}^r V_{t_{m,j}}^{k,i}} \right| \quad (13)$$

Similar to the average velocity, the density estimation is also affected by the penetration rate. According to the density definition in Eq. (3), the density estimation error, denoted as $E_D^{k,i}$, is directly affected by the penetration rate $\rho^{k,i}$ as expressed in Eq. (14).

$$E_D^{k,i} = (1 - \rho^{k,i})100\% \quad (14)$$

Figure 3 shows the effect of the penetration rate on velocity and density estimations. Both the $E_V^{k,i}$ and $E_D^{k,i}$ are affected significantly by the low penetration rate. While the $E_V^{k,i}$ is depicted as a curve along the change of ρ , $E_D^{k,i}$ is linearly affected by the change of ρ . This figure also reveals that if an optimal traffic state estimation model is proposed, the estimation effectiveness can be validated based on the given penetration rate. For example, if the penetration rate is around 50%, $E_V^{k,i}$ is around 10% and $E_D^{k,i}$ is around 50%. However, in practice the “actual” penetration rate cannot be measured at the estimation time. It means the uncertain penetration rate issue impedes the design of an optimal traffic state estimation model. This dilemma will be solved in the remainder of this paper.

4. Intelligent Context-Aware Velocity-Density Inference Circuit

In order to alleviate errors rooted from a low penetration rate, a

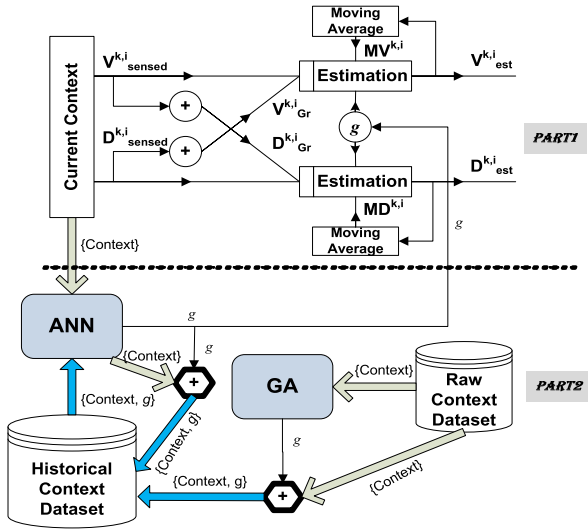


Fig. 4 The intelligent context-aware velocity-density inference circuit (ICIC).

novel intelligent context-aware velocity-density inference circuit (ICIC) is proposed as follows.

4.1 General Circuit

The proposed ICIC model is illustrated in Fig. 4 which consists of 2 parts. Part 1 is the velocity-density inference circuit (VDIC) [10] and part 2 is the mechanism that identifies the optimal parameters for the VDIC in part 1.

In part 1, the velocity and the density calculated directly from the sensed data reported by mobile phones, namely $V_{sensed}^{k,i}$ and $D_{sensed}^{k,i}$, serve as primary inputs. The circuit provides the final estimated velocity and density, namely $V_{est}^{k,i}$ and $D_{est}^{k,i}$. The intermediate inferred velocity and density, namely $V_{Gr}^{k,i}$, $D_{Gr}^{k,i}$, obtained by applying the Greenshields model [25] are also taken into account. In addition, moving average values of estimated velocity and density at time k , namely $MV^{k,i}$ and $MD^{k,i}$, are fed back to the estimation model.

The philosophies behind this VDIC are as follows: 1) The velocity and the density calculated independently from sensed data help to avoid any error propagation. 2) The Greenshields model [25] used to infer density from estimated velocity and vice versa, can help to avoid the over-error of the density estimation when the penetration rate becomes unacceptably low. 3) The current traffic state contains inherent relations with previous traffic states at the same road segment. 4) All of the estimation approaches (direct estimation using sensed data, inference using the Greenshields model, inference using the previous estimated data) may uphold their advantages while diminishing their inherent disadvantages if being appropriately integrated.

The overall VDIC is formally presented in Eqs. (15) and (16), where α , β , and γ are the impact coefficients of the corresponding parameters. It should be noted that α , β , and γ are encapsulated in a simplified parameter g ($g = \{\alpha, \beta, \gamma\}$) in Fig. 4. These coefficients are real-value numbers in the range of [0, 1] and the sum of them is 1 as described in Eq. (17).

$$V_{est}^{k,i} = \alpha V_{sensed}^{k,i} + \beta MV^{k,i} + \gamma V_{Gr}^{k,i} \quad (15)$$

$$D_{est}^{k,i} = \gamma D_{sensed}^{k,i} + \beta MD^{k,i} + \alpha D_{Gr}^{k,i} \quad (16)$$

$$\alpha + \beta + \gamma = 1 \quad (17)$$

The moving average velocity and density at time k , $MV^{k,i}$, $MD^{k,i}$ are calculated in Eqs. (18) and (19), where ξ is the sliding window which can be set by domain experts or by using simulation data.

$$MV^{k,i} = \frac{\sum_{j=k-\xi}^{k-1} V_{est}^{j,i}}{\xi} \quad (18)$$

$$MD^{k,i} = \frac{\sum_{j=k-\xi}^{k-1} D_{est}^{j,i}}{\xi} \quad (19)$$

The intermediate inferred velocity and density using the Greenshields model, $V_{Gr}^{k,i}$ and $D_{Gr}^{k,i}$, are calculated in Eqs. (20) and (21), where D_{max}^i and V_{max}^i are the maximum density and the limited velocity of the road segment i .

$$V_{Gr}^{k,i} = V_{max}^i \left(1 - \frac{D_{sensed}^{k,i}}{D_{max}^i} \right) \quad (20)$$

$$D_{Gr}^{k,i} = D_{max}^i \left(1 - \frac{V_{sensed}^{k,i}}{V_{max}^i} \right) \quad (21)$$

As discussed, one of the most important keys constituting the optimization of the estimation models described in Eqs. (15) and (16) is the optimization of coefficients α , β , γ . Obviously, these parameters have some relations with the penetration rate. For example, when the penetration rate is relevant, the accuracy of $V_{sensed}^{k,i}$ is high revealing that $V_{sensed}^{k,i}$ is almost equal to $V_{est}^{k,i}$. It means that α is prominent as around 1, and β , γ are minor as around 0. However, the difficulty here is that, the actual penetration rate cannot be known at the estimation time. The work in Ref. [10] proposed a simple estimation model by which a single coefficient, namely α , was utilized. A statistical model based on historical data to approximately identify this coefficient was proposed. The weakness of this approach, however, is that α was empirically approximated as only two values, namely 0.6 or 0.8, whereby some sophisticated details were overlooked. It could be more effective if α was granularly determined at any value in the range of [0, 1]. In solving this difficulty, more contextual data about the considered road segment as well as the current traffic flow could be useful. In this article, an intelligent context-aware inference approach is proposed to optimize impact coefficients α , β , γ , thus optimizes the whole ICIC model. Details of this approach are presented in the remainder of this section.

4.2 Intelligent Context-Aware Inference Approach

Obviously, the current state of a traffic flow might have some relations with its surrounding contexts. Some of these contexts can be obtained directly from the road segment while other contexts can be extracted from the data reported by mobile phones. Useful contexts which are relevant for inferring the optimal set of coefficients α , β , γ will be discussed in this subsection.

The notion of context has been observed in numerous areas including linguistics, knowledge discovery and presentation, artificial intelligence, information retrieval, reasoning, robotics, theory of communication [27], [28], [29], and so forth. At a high level of abstraction, a context is defined as "that which surrounds, and gives meaning to, something else" [27]. In this definition,

Table 1 Coefficients α, β, γ inferred from current contextual data.

V_{sensed}/V_{max}	V_{sensed}/V_{Gr}	D_{sensed}/D_{max}	MV^t	α	β	γ
0.2349	0.2779	0.1545	66.4811	0.6660	0.0395	0.2945
0.1682	0.2104	0.2004	41.4499	0.6621	0.2881	0.0497
0.1115	0.1454	0.2331	26.7230	0.5543	0.3969	0.0488
0.1172	0.1536	0.2369	14.5153	0.5752	0.3973	0.0275
0.1444	0.1977	0.2697	13.0120	0.8554	0.1446	0.0000
0.5807	0.6252	0.0712	74.7853	0.6468	0.1996	0.1536
0.2935	0.3378	0.1311	60.6069	0.6293	0.2867	0.0840
0.0938	0.1258	0.2547	15.3041	0.7937	0.1498	0.0565
0.4878	0.6231	0.1086	74.3697	0.4981	0.2393	0.2625
0.2031	0.4155	0.2556	56.6439	0.3786	0.1972	0.4242

“something” can be an artifact, a building, a person, a computer system or even an assertion in logic. Dey [29] gives a more practical definition of context as “Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.” The “situation” or the “status” of an entity may be clearly represented by the entity itself. However, in practice, an entity’s status may be a variable that needs to be “inferred.” In this work, the “status” to be inferred is the optimal combination of α, β, γ which leads to the optimal traffic state estimation.

As shown on part 2, Fig. 4, an ANN model is employed to predict the coefficients, $g = \{\alpha, \beta, \gamma\}$, based on the current context of the considered traffic flow and the knowledge learned from the historical context data. In this model, there are two historical datasets. The “raw context dataset”, denoted as $\{context\}$, contains the “raw” contextual data such as the velocity and density calculated directly from data reported by mobile phones, the actual traffic state, namely the actual velocity and density, and so forth. This data is fed to a genetic algorithm based (GA) model to identify the optimal set of g forming a new form of context, namely the combination of $\{context, g\}$. This data is stored in the second dataset, namely the “historical context dataset,” which is used to train the ANN model. Finally, at the estimation time, based on the current context, the ANN can infer the set of coefficient g which is forwarded to the VDIC in part 1. In addition, any new pair of $\{context, g\}$ generated by the ANN is restored back in the “historical context dataset” to enrich this knowledge-based database so that the ANN model can work better in latter iterations.

Obviously, it is not necessary that every context is useful. In addition, the possibility of sensing (obtaining) useful contexts must be considered under the given infrastructural conditions of the proposed application. In the MC-TEs’s circumstance, only GPS data is given, thus useful contexts must be carefully selected from this limited data. The selected contexts and the reasons behind these selections are presented as follows:

- V_{sensed}/V_{max} represents information related to both the real-time data, V_{sensed} , and the physical features of the considered road segment, namely the limited speed, V_{max} .
- V_{sensed}/V_{Gr} reflects the “quality” of the sensed data under the effect of the penetration rate.
- D_{sensed}/D_{max} represents the current recognition of the den-

sity based on the real-time data. This factor is identical to the V_{sensed}/V_{max} but in terms of density.

- MV^t (the moving average of velocity at the estimation time t) serves as a factor that transits previous traffic states to the current traffic state.

Table 1 presents a portion of data extracted from evaluated data showing that coefficients α, β, γ were successfully inferred from the contextual data mentioned above.

In addition to the ANN component, the GA component (please refer to part 2, Fig. 4) also plays an important role in providing the ANN component a qualified set of $\{contexts, g\}$. Therefore, an appropriate GA [30] mechanism which can effectively work with the set of raw contextual data mentioned above is proposed.

GA is a technique that mimics the natural selection (Darwin’s theory of survival) which is effectively applied to the optimization and global search [30], [31], [32]. The process of a GA can be illustrated as a flowchart in **Fig. 5**. An initial population, P (a set of chromosome) in a suitable size is randomly generated. Each chromosome consists of one or several genes. A gene is a string of bits or a string of real-value numbers. Each chromosome represents a solution of the problem solved by the GA. In each population, candidates (chromosomes) are selected based on their fitness values so that those individuals that are more competitive have a larger chance to survive and to keep the genetic information to their offspring. High competitive chromosomes (the survived ones) are duplicated with a corresponding probability based on their fitness to replace the low competitive candidates and to keep the size of population unchanged. After that, any two selected candidates are mated randomly producing two alternative children under the crossover procedure. The mutation helps to prevent GA from falling into local extreme, so that the global solution may be found earlier. The GA would not stop until the fitness of the best chromosome in the current population satisfies a present threshold or when the iterations pass the preset time. The design of the proposed GA is presented as follows.

As mentioned before, the $\{context\}$ in the “raw context dataset” contains actual traffic state data which includes the actual velocity, V_{act} . The proposed GA optimizes the set of $g = \{\alpha, \beta, \gamma\}$ by which the variation between V_{est} (estimated by applying the proposed estimation model described in Eq. (13) with the coefficient set g) and V_{act} is minimal. The schema of a chromosome in the proposed GA is coded as $g = \{\alpha, \beta, \gamma\}$ (the chromosome of 3 genes). Candidates are evaluated using the fitness function proposed in Eq. (22).

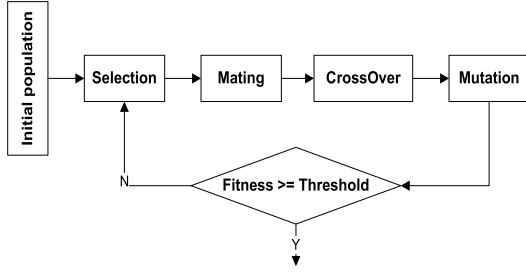


Fig. 5 The flowchart of a GA's process.

$$f(g_i) = \frac{e(g_i)}{\bar{e}(g_j, \forall g_j \in \text{population})} \quad (22)$$

Here, $e(g_i)$ is the evaluation of candidate g_i ($g_i = \{\alpha_i, \beta_i, \gamma_i\}$) and $\bar{e}(g_j)$ is the average evaluation of all individuals g_j in the current population. It is the estimation error in Eq. (15) caused by selecting g_i as the set of coefficients. The evaluation $e(g_i)$ is defined as in Eq. (23), where V_{est-g_i} is the estimated velocity by applying g_i ($g_i = \{\alpha_i, \beta_i, \gamma_i\}$) as coefficients, and V_{act} is the “actual” velocity.

$$e(g_i) = \frac{|V_{est-g_i} - V_{act}|}{V_{act}} \quad (23)$$

Regarding the crossover and mutation procedures, since the chromosome schema is built on real-value numbers, the arithmetic crossover [31], [32] and the non-uniform mutation approach [33], [34] are applied. The arithmetic crossover procedure is described in Eq. (24), where g_1, g_2 are the two parent chromosomes and g'_1, g'_2 are their children; and λ is a random coefficient in the range of [0, 1].

$$\begin{cases} g'_1 = \lambda g_1 + (1 - \lambda)g_2 \\ g'_2 = (1 - \lambda)g_1 + \lambda g_2 \end{cases} \quad (24)$$

The non-uniform mutation approach is described as follow. Assuming that a gene, namely α_i , of the chromosome $g'_i = \{\alpha_i, \beta_i, \gamma_i\}$ is mutated, the mutation is described as in Eq. (25).

$$\alpha''_i = \begin{cases} \alpha_i + \Delta(t, UB - \alpha_i), & \text{if a random } \zeta \text{ is } 0 \\ \alpha_i - \Delta(t, \alpha_i - LB), & \text{if a random } \zeta \text{ is } 1 \end{cases} \quad (25)$$

Here, UB and LB are the upper and lower bounds, namely 1 and 0, respectively, of the variable α_i . The function $\Delta(t, x)$, where t describes the current generation, returns a value in the range of $(0, x)$ which is defined as in Eq. (26) [33].

$$\Delta(t, x) = x(1 - r^{(1-\frac{t}{T})^b}) \quad (26)$$

In Eq. (26), r is a uniform random number ranging in [0, 1], T is the maximal generation number, and b is a system parameter determining the degree of dependency on the iteration number.

Reaching this stage, the proposed ICIC model can be summarized as follows: The GA component provides an optimal set of $\{contexts, g\}$ to train the ANN model. At the estimation time, the current contextual data is fed to the ANN model so that it can identify the optimal coefficients, $g = \{\alpha, \beta, \gamma\}$. These coefficients are forwarded to the VDIC in the part 1 to effectively estimate the traffic state of a certain road segment. The effectiveness of the ICIC model will be evaluated in Section 7.

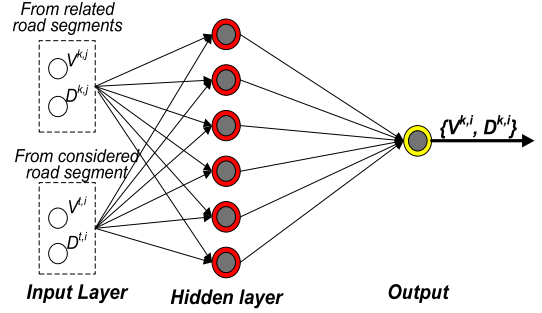


Fig. 6 The ANN-based prediction model dealing with an unacceptably low penetration rate.

5. Traffic State Prediction Under an Unacceptably Low Penetration Rate

In general, the ICIC model proposed in Section 4 improves the effectiveness of the MC-TEs framework significantly, even in cases of a low penetration rate. However, when the penetration rate becomes unacceptably low, namely just a few percent or even zero, the ICIC cannot work properly. It is due to the fact that when the penetration rate is unacceptably low the deviation between the “real” data and the obtained data is large. Consequently, the ANN component in the ICIC model cannot infer the appropriate set of coefficients α, β, γ based on the current low qualified contexts. To address this issue, a prediction model is proposed. In this approach, when the penetration rate of the considered road segment is unacceptably low, the contexts mentioned in Section 4 are not applied. Instead, other contexts such as previous traffic states of the considered road segment, the current traffic state of related road segments, and so forth, are utilized to predict the traffic state of the considered road segment.

Obviously, the traffic state of a road segment is affected by the traffic state of nearby road segments as well as by its previous states. If these *spatial-temporal* relations (rules) are known in advance, the traffic state of the considered road segment can be predicted. These rules can be learned by any machine learning technique using historical traffic state data. In this article, a practical ANN with a multilayer perceptron (MLP) [34], [35] is proposed. This approach is illustrated in Fig. 6. Here, j ($j \in V$) is any related road segment (i.e., the segments which are directly connected to or are close to/nearby the considered road segment) of the considered road segment i ($j \neq i$). The velocity and density of the related road segment j at time k , denoted as $V^{k,j}, D^{k,j}$, are the input data of the ANN model. The other inputs are the previous velocity and density of the considered road segment i , namely $V^{t,i}, D^{t,i}$, respectively, where $t < k$.

One of the essential issues here is how to effectively identify the related road segments in a wide road network. In fact, not only the directly connected road segments but also the ones that indirectly connect to the considered road segment may affect the considered road segment’s traffic state. However, it is impractical to consider all road segments in the whole road network. In order to trade off the prediction effectiveness and the computational cost, only the first three orders related road segments are employed. The related road segments at different orders are defined as follows:

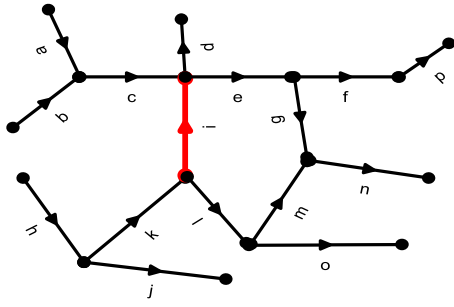


Fig. 7 Related road segments of a considered road segment i .

Let j be a road segment that directly connects to the considered road segment i , and E is the set of directly connected link ij . The road network can be defined as a directed graph $G = (V, E)$, where V is the set of all road segments on the road network. The order (level) of related road segments to the considered road segment i is defined in Eq. (27). Here, the 1st order related road segments, denoted as $N^1(i)$, are all the road segments that directly connect to the considered road segment i (i.e., $ij \in E$ or $ji \in E$). The n th order related road segments, denoted as $N^n(i)$, are recursively defined as the union of all the j 's 1st order related road segments, where j is any road segment which belongs to $N^{n-1}(i)$. For example, in Fig. 7, $\{c, d, e, k, l\}$ are the 1st order and $\{a, b, h, j, m, g, f, o\}$ are the 2nd order related road segments of the given road segment i .

$$\begin{cases} N^1(i) = \{j : j \neq i, j \in V \wedge (ij \in E \vee ji \in E)\} \\ N^n(i) = \bigcup N^1(j) : \forall j \in N^{n-1}(i) \end{cases} \quad (27)$$

Finally, the proposed ANN-based prediction model can be formulated in Eq. (28).

$$\{V_{pre}^{k,i}, D_{pre}^{k,i}\} = \text{predict}(V^{k,j}, D^{k,j}, V^{t,i}, D^{t,i}), \\ t < k, j \neq i, j \neq i, j \in \bigcup_{s=1}^{w-1} N^s(i) \quad (28)$$

In this model, the velocity and density of the considered road segment i at time k , denoted as $V_{pre}^{k,i}, D_{pre}^{k,i}$, are predicted by the so-called $\text{predict}()$ function, where the current traffic state of related road segments ($V^{k,j}, D^{k,j}$) and the previous traffic state of the considered road segment ($V^{t,i}, D^{t,i}$) are served as input parameters.

6. Appropriate Traffic State Estimation Model Selection

Two novel traffic state estimation models considering the issues of a low penetration rate, namely the ICIC and the ANN approaches, have been proposed in Sections 4 and 5. An emerging question here is that which model should be selected in the real application at a particular estimation.

As discussed in Sections 4, and 5 and illustrated in Fig. 8, the estimation error of the ICIC decreases significantly when the penetration rate increases, while the estimation error of the ANN model is not affected by the penetration rate of the considered road segment. Consequently, there is a point where these two estimation errors are identical. This point is called the *critical point* and the penetration rate at that point is called the *critical penetration rate*. According to this analysis, it seems simple to select the appropriate traffic state estimation model. That is to say, if

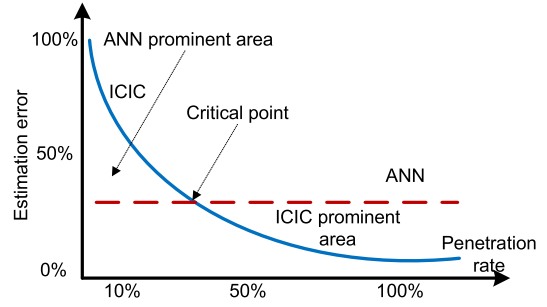


Fig. 8 Relationship between the ICIC and the ANN effectiveness under the condition of the penetration rate.

the current penetration rate is larger than the *critical* one the ICIC model should be selected, and vice versa.

However, the crucial issue here is that there is no way to estimate the “actual” penetration rate accurately. Moreover, under the condition of an *uncertain* penetration rate the ICIC model cannot be applied since its accuracy cannot be evaluated. Consequently, only the ANN model is useful while its prediction accuracy is restricted to a constant value and cannot be improved even if the actual penetration rate is high. This section proposes a simple yet effective traffic state estimation model selection approach whereby the “right” traffic state estimation model can be chosen without any information about the actual penetration rate.

Even though there is no way to select the appropriate traffic state estimation model directly due to the uncertain penetration rate, there are clear relations between estimation errors of the two estimation approaches. As shown in Fig. 8, the gap between the two curves, namely the ICIC and the ANN, changes along with the change of penetration rate. For instance, this gap is narrow and seems to be stable in cases of a relevant penetration rate (i.e., larger than the critical one). On the other hand, this gap is large in cases of a low penetration rate. Therefore, instead of using the current penetration rate, which is uncertain at the estimation time, the relation between the evaluated values provided by the two estimation models should be used to select the “right” estimation model. However, the estimation errors of the two estimation approaches are still variables since the two estimation models just provide the estimated values, while the actual value is unknown. For example, let denote V_{act} , V_{ann} , and V_{icic} the “actual” velocity, and the estimated velocities provided by the ANN and the ICIC models, respectively. Since V_{act} is unknown, the estimation errors of the two estimation models cannot be known either. This section proposes an approximate selection (AS) method to effectively select the appropriate estimation model. The detail of this method is described as follows:

Although the actual velocity (V_{act}) is a variable, the relations between V_{act} and the estimated velocities, V_{ann} and V_{icic} , can be figured out which are shown in Fig. 9. Here, V_{act} is virtually represented as a straight line representing a constant value (e.g., 50 km/h) regardless of the penetration rate. As shown in Fig. 8, the estimation error in the ANN approach is stable regardless of the penetration rate. Therefore, in Fig. 9, V_{ann} is represented by 2 straight dotted lines, namely V_{annUB} and V_{annLB} ($V_{annLB} < V_{act} < V_{annUB}$) paralleling with V_{act} . Different from the ANN approach, the ICIC’s estimation error changes along

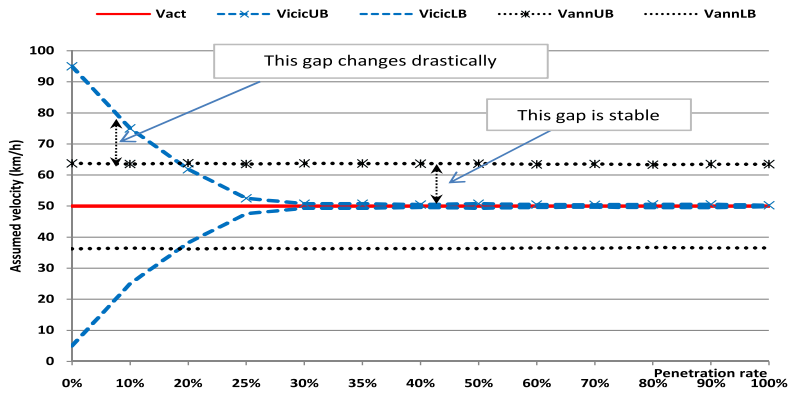


Fig. 9 The relation between the “actual” and estimated velocities considering different penetration rates.

with the change of penetration rate. Therefore, V_{icic} lies on one of the 2 curves representing V_{icicUB} and V_{icicLB} depicted in Fig. 9 ($V_{icicLB} < V_{act} < V_{icicUB}$). As shown, the gap between V_{ann} (namely V_{annUB} or V_{annLB}) and V_{act} is kept constant regardless of the penetration rate. In contrast, the gap between V_{ann} and V_{icic} drastically increases when the penetration rate becomes lower than the critical one. On the other hand, this gap is narrow and almost stable when the penetration rate is relevant (larger than the critical one). It means that the gap between V_{ann} and V_{icic} reflects the current penetration rate. Therefore, it can be utilized to approximately identify the appropriate estimation model. This approximation method can be described in Eq. (29), where \bar{V} is defined in Eq. (30).

$$V_{act} = \begin{cases} V_{ann}, & \text{if } \frac{|V_{icic} - V_{ann}|}{\bar{V}} \geq \theta\% \\ V_{icic}, & \text{other cases} \end{cases} \quad (29)$$

$$\bar{V} = \frac{|V_{ann} + V_{icic}|}{2} \quad (30)$$

Equation (29) represents that if the gap between V_{ann} and V_{icic} is larger than a threshold value, namely $\theta\%$, the ANN approach is selected since the penetration rate is low, while the ICIC model is selected in the other cases (relevant penetration rates). The threshold θ can be determined by experimental data as shown in the evaluation section.

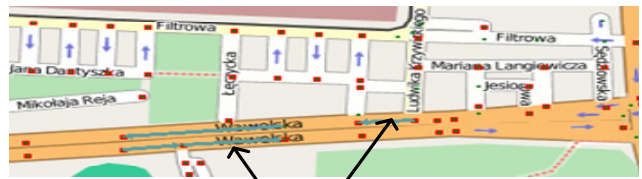
7. Effectiveness Evaluation and Analysis

This section analyzes and evaluates the effectiveness of the proposed solutions, namely the ICIC model, the ANN-based prediction approach, and the traffic state estimation model selection method.

7.1 Experimental Environment and the Data structures

To evaluate the effectiveness of the proposed approaches, a large amount of data is required. The TSF simulator [36] was utilized to generate synthetic data for evaluations. Different road segments were selected randomly as shown in Fig. 10. For each selected road segment, two types of data were created concurrently as shown in Table 2.

It should be noted that different penetration rates, namely 10%, 20%, 25%, 30%, and so forth, were configured by the TSF for each selected road segment. For a preset penetration rate, only such a percentage, namely 20% for example, of random vehi-



The road segments
Fig. 10 Road segmentation in the TSF.

cles reported data to the server. This data is called the detailed GPS data (type a) as described in Table 2. The frequency of the data report timing, which can be configured using the TSF flexibly, was set to every 3 s in this work to mimic the common GPS signal frequency. The summarized data (type b) consists of the actual traffic state information, namely the average velocity and the density of the considered road segment presented by the *road segment Id*. These data were used to evaluate the accuracy of the proposed traffic state estimation approaches, namely the ICIC and the ANN models when these methods process the detailed GPS data. The time interval for recording summarized traffic state data was set to every minute.

7.2 Effectiveness and Feasibility of the ICIC Model

This section evaluates the effectiveness and the feasibility of the proposed ICIC model. To prepare evaluation data, 10 random road segments were selected and associated with different penetration rates. For each pair of {road segment i , penetration rate ρ }, 10 one-hour simulations were conducted. These data were applied to the ICIC model to estimate the average velocity and density of the 10 selected road segments. After that, the average differences between the estimated values (i.e., the estimated velocities and densities) and the “actual” values (provided by the TSF simulator - type b data in Table 2) were drawn out and shown in Fig. 11 and Fig. 12.

Figures 11 and 12 show the effectiveness of the ICIC model compared to that of its counterparts, namely, the conventional estimation model and the VDIC approach proposed in Ref. [10], respectively. The ICIC’s velocity and density estimation accuracies are denoted as ICIC_V and ICIC_D, respectively. The corresponding accuracies in the conventional model (i.e., the *sensed* data is applied directly to the conventional estimation models described in Eqs. (1) and (3), in Section 3) are denoted as Normal_V and Normal_D. Similarly, VDIC_V and VDIC_D represent veloc-

Table 2 Data structure of the data generated by TSF.

Data type	Data structure
a) Detailed GPS data (reported by an individual vehicle)	{time stamp (in second), road segment Id, position (longitude, latitude), current velocity, vehicle Id}
b) Summarized traffic state data	{time interval Id (in minute), road segment Id, average velocity, density}

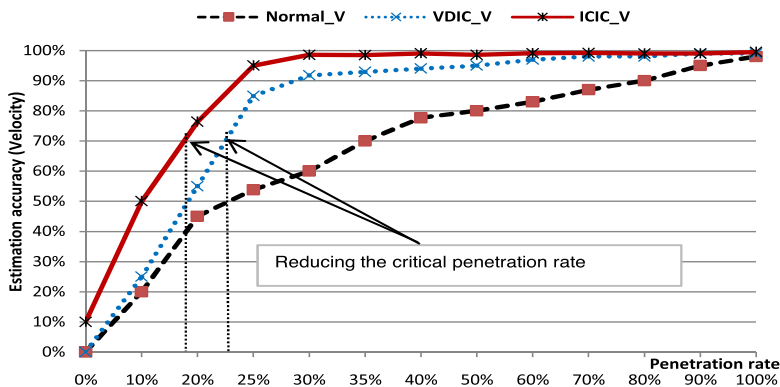


Fig. 11 Effectiveness of the ICIC model - Average velocity estimation.

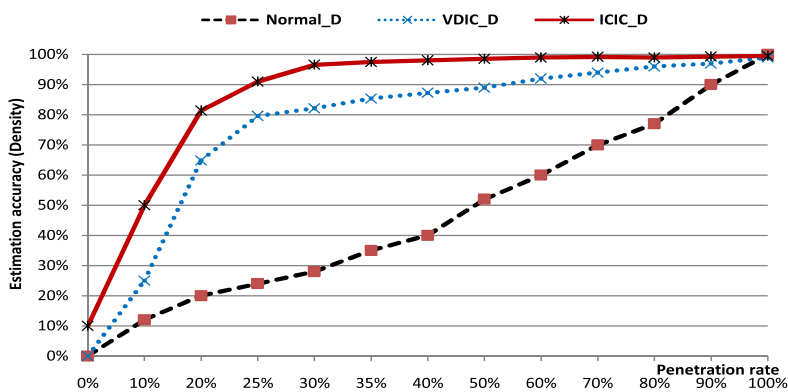


Fig. 12 Effectiveness of the ICIC model - Density estimation.

ity and density estimation accuracies in the VDIC model. Both the two figures show that among these approaches, the ICIC is prominent. In all of the cases when the penetration rate is low but still relevant, namely higher than 25%, the ICIC is completely optimized so that its accuracies are almost higher than 95%. In addition, the ICIC model also minimizes the *acceptable* penetration rate required for a particularly *expected* accuracy. For example, if the *expected* accuracy is set to 70%, the ICIC reduces the *acceptable* penetration rate from 35% (in the normal approach), or from 22% (in the VDIC) to lower than 18% (please refer to Fig. 11). This improvement is practically worth since it improves the scalability of the estimation system in terms of low penetration rates.

In addition to the overall effectiveness evaluated above, the feasibility of the ICIC model was also evaluated. As depicted in Fig. 4, Section 4, the GA component plays an important role in processing a large amount of historical data to provide the ANN component an optimal set of {contexts, g}. For example, in this work, a dataset which consists of 500,000 raw contextual data (please refer to Fig. 4) records was applied to the GA component to generate the optimal set of {contexts, g} to train the ANN model. Therefore, the feasibility (represented by the convergence

rate) as well as the computation time of the GA component should be evaluated.

The GA component was implemented with the error threshold of 0.1% (i.e., the maximum accepted distance between V_{est} and V_{act} is 0.1% of V_{act}), and the maximum number of iterations was set to 500,000. The GA is convergent if the distance between V_{est} and V_{act} reaches the preset error threshold before the maximum iteration number is hit. The convergence and computation time of the proposed GA are affected by the quality and the volume of the evaluated dataset. These qualifiers are affected by the penetration rate and depicted in Fig. 13.

Figure 13 shows that when the penetration rate is low, lower than 20% for instance, the convergence rate is definitely low (lower than 50%) and the computation time is small, namely less than a second. This effect is due to the missing data caused by a low penetration rate. When a lot of useful data are missed, the GA cannot find out the optimal solution (cannot converge). Moreover, the dataset in this case is small thus less computation time is required. On the other hand, when the penetration rate is relevant, larger than 30% for instance, the convergence rate reaches 100% while a longer computation time is required. Nevertheless, as shown, the maximum computation time is limited

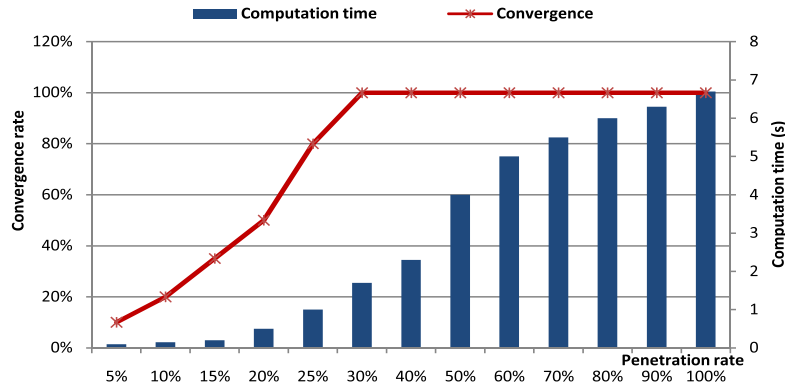


Fig. 13 The GA convergence and computation time.

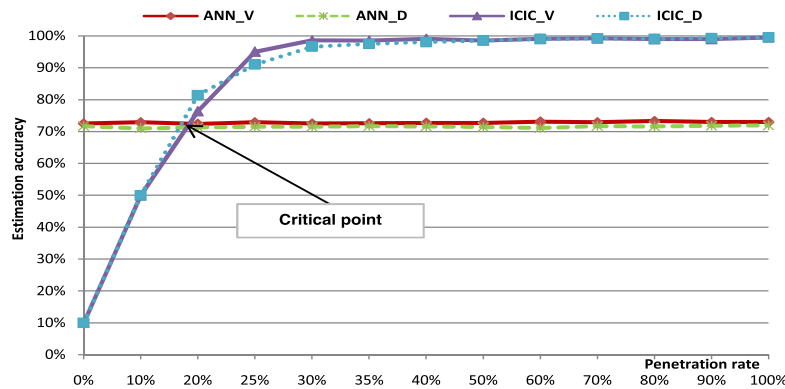


Fig. 14 Effectiveness of the ANN-based prediction model compared to that of the ICIC model.

to 7 s (when the penetration rate is 100%), which is fast enough. This evaluation result also implies that to prepare an optimal set of {contexts, g} for training the ANN in the ICIC model, the “raw context dataset” (please refer to Fig. 4) should be well validated by which the penetration rate must be larger than 30%. Obviously, this condition can be satisfied in real-world applications since the “raw context dataset” is prepared prior to the launch of the ICIC model.

7.3 Effectiveness of the ANN-based Prediction Model

This section evaluates the effectiveness of the ANN-based prediction model. Five road segments were randomly selected as the desired ones whose average velocity and density are required to be predicted. In addition, three different, namely the 1st, 2nd, and 3rd, order related road segments of each desired road segment were also identified so that their corresponding traffic state data (average velocity and density) were recorded. For each pair of {road segment i , penetration rate ρ }, ten 1-hour simulations were performed by which the average velocity and density of not only the road segment i but also of any related road segment j ($j \in \bigcup_{s=1}^3 N^s(i)$) were recorded. Concretely, at each time interval k , the velocity and density of the considered road segment i and its related road segment j , denoted as $V^{k,i}$, $D^{k,i}$, $V^{k,j}$, $D^{k,j}$, respectively, were obtained. Finally, each record in the evaluation dataset consists of: 1) $V^{k,i}$, $D^{k,i}$ serving as the target elements (the outputs of the ANN model); and 2) $V^{k,j}$, $D^{k,j}$, $V^{t,i}$, $D^{t,i}$ $t < k$, $j \in \bigcup_{s=1}^3 N^s(i)$ serving as the inputs for the ANN model as described in Eq. (28). The summarized traffic state information mentioned above was recorded in each minute revealing that 60

data patterns were generated in each simulation (1 hour). As a result, for each pair of {road segment i , penetration rate ρ }, 600 data patterns (in 10 1-hour simulations) were created. The generated dataset (3,000 records) was then divided into two parts as the portion of 75% and 25% for the training and the testing datasets, respectively. An ANN with 6 hidden nodes was trained and evaluated by these training and testing datasets. The average predicted accuracies for both velocity and density estimations of the 5 desired road segments mentioned above were calculated to evaluate the effectiveness of the ANN-based prediction model. The results of these evaluations are shown in Fig. 14.

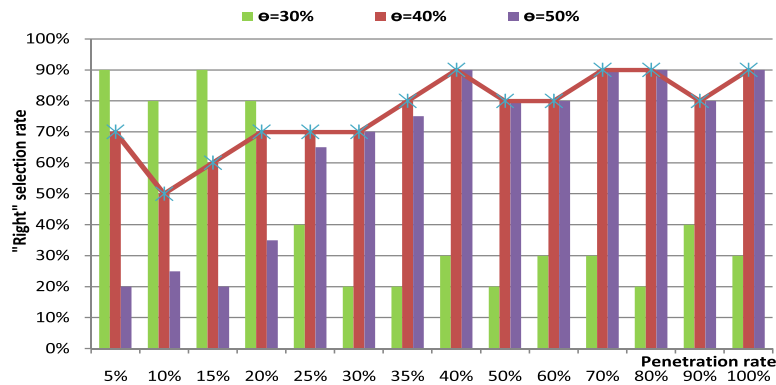
Figure 14 shows that in the ANN-based prediction method, accuracies for both velocity and density estimations, denoted as ANN_V and ANN_D, respectively, are around 73% regardless of the penetration rate. This figure also reveals that even though the ANN model is not affected by the penetration rate, its accuracy is limited (i.e., around 73%). Therefore, this model is suited in the cases of an unacceptably low penetration rate, namely lower than the critical one which is around 18%. When the penetration rate is relevant which is larger than the critical one, the ICIC approach (denoted as ICIC_V and ICIC_D for the velocity and density estimation accuracies, respectively) is dominant.

7.4 Effectiveness of the Traffic State Estimation Model Selection Method

To evaluate the traffic state estimation model selection method (the AS method) proposed in Section 6, another large dataset was generated by the TSF. For each preset penetration rate, 10 road segments were randomly selected. For example, 10 different road

Table 3 Summary of the dataset used for evaluating the AS method.

Measured parameter	Value
Number of different penetration rates (ρ)	14 (from 5% to 100%) - For $\rho \leq 25\%$, the interval was set to 5%. - For $\rho > 25\%$, the interval was set to 10%.
Number of road segments	10
Number of $\{\rho, \theta\}$ pairs	140
Decision threshold ($\theta\%$)	30%, 40%, 50%
Number of records for each pair of $\{\rho, \theta\}$	3,000 (10 road segments x 5 simulations x 60 record per each 1-hour simulation)


Fig. 15 Effectiveness of the proposed method for selecting the “Right” traffic state estimation model.

segments were selected randomly for a penetration rate of 30%, while other 10 different road segments were randomly selected for a penetration rate of 40%, and so forth. For each pair of road segment i , penetration rate ρ , five 1-hour simulations were conducted by which for both type a and type b data mentioned in Section 7.1, Table 2 was recorded. The proposed ANN and ICIC models were applied to estimate V_{ann} and V_{icic} , respectively. These two values were then utilized to evaluate the AS method. It should be noted that the selection is performed based on a given decision threshold θ (please refer to Eq. (29)). In this evaluation two questions should be clarified as follows: 1) what value is most suited for the decision threshold θ ?; and 2) at the optimal threshold how effective is the AS method?

To identify the suited decision threshold, the right selection rate of each pair of {penetration rate ρ , selection threshold θ } was measured. This rate is defined as the portion of the correct selections (i.e., the right traffic state estimation model was selected) out of the total number of selections. There were 3,000 instances (5 simulation * 60 data patterns extracted from 1-hour simulation * 10 road segments) of data instance for each pair of $\{\rho, \theta\}$. More concretely, the right selection rate in each pair of $\{\rho, \theta\}$ is the number of correct selections per 3,000 selections. This process was repeated for every pair of $\{\rho, \theta\}$, namely $\{\rho = 20\%, \theta = 30\%$ for example. In this evaluation, the penetration rate ranged from 5% to 100% and the selection threshold was set to 30%, 40% and 50%. The evaluated dataset mentioned above is summarized in **Table 3**. An interesting point that is worth to be noted here is that the interval of penetration rates for preparing evaluated data is set at 25% ($\rho = 25\%$), but not at the critical penetration rate ($\rho = 18\%$). Obviously, around the critical point the accuracy is still fluctuant (please refer to Fig. 11). Meanwhile, from $\rho = 25\%$ the accuracy becomes more stable. Therefore, $\rho = 25\%$ was considered as an anchor point. More concretely, when $\rho < 25\%$ the

Table 4 The average accuracy of the AS method.

Penetration rate	Selection thresholds and accuracy (%)		
	30	40	50
All the cases	41.36%	71.79%	63.50%
$\leq 25\%$	67.80%	60.60%	33.00%
$> 25\%$	26.67%	78.00%	80.44%

accuracy is still fluctuant and as a result more granular data is needed.

The effectiveness of the AS model is shown in **Fig. 15** and summarized in **Table 4**.

In Fig. 15, the right selection rates with regard to different pairs of $\{\rho, \theta\}$ are shown. For example, at $\rho = 35\%$, the right selection rates are 20%, 80%, 75% for thresholds of 30%, 40%, and 50%, respectively. In addition, this figure also reveals that the low threshold supports the selection method in the cases of a low penetration rate, and vice versa. More concretely, with the threshold of 30% ($\theta = 30\%$) for example, the right selection rate is almost higher than 80% if the penetration rate is low ($\leq 20\%$), while this rate decreases drastically ($\leq 40\%$) when the penetration rate is relevant ($\geq 25\%$). If a completely high threshold, namely $\theta = 50\%$, is chosen, the right selection rate is high when the penetration rate is relevant while this rate is drastically decreased when the penetration rate is low. This figure also shows that the AS method works well (with a high right selection rate) in both the high and low penetration rates when the threshold is set to 40%. In this case, the right selection rate is quite high (around 80% or 90%) in cases of a relevant penetration rate, while this rate in the cases of a low penetration rate is a little bit lower but still high enough (around 60%).

Table 4 summarizes the effectiveness of the AS method with regards to different thresholds. This table shows the right selection rate in a high penetration rate ($> 25\%$), low penetration rate

($\leq 25\%$), and in all the cases (consolidation of both the former cases). As shown, with a low threshold ($\theta = 30\%$), the right selection rate is high (67.8%) in cases of a low penetration rate, but it declines drastically to 26.67% in cases of a high penetration rate. The overall right selection rate in this case is 41.36% which is considered unacceptably low. In contrast, with a high penetration rate ($\theta = 50\%$), the relation between the penetration rate and the right selection rate is in the converse way. That is to say, the right selection rate is high (80.44%) in cases of a high penetration rate, while it is low in cases of a low penetration rate (33%) leading to an overall right selection rate of about 63.5%. The AS method is optimal with the threshold of 40% where it works well for both the low and high penetration rates. More concretely, the right selection rates are 60.6% and 78% with regard to low and high penetration rates, respectively, resulting in a high overall right selection rate of about 71.79%. Therefore, if the AS method is applied in real-world applications, the selection threshold should be 40% which provides high enough selection accuracy of around 71.79%.

It would be worth to recall the purpose of the proposed AS method which aims at determining the appropriate traffic state estimation method, namely between the ANN and the ICIC method in accordance with a particular situation. As mentioned above, on average, the right selection rate is 71.79% when the threshold is 40%. This rate is high enough which satisfies the purpose of this research. The traffic state estimation accuracy of the “right” and the “wrong” selected estimation methods can be found in Fig. 14. For example, when the penetration rate is 35%, if the ICIC model was selected then this selection was “right” revealing that the traffic state estimation accuracy is larger than 90% (please refer to Fig. 14). In contrast, if the ANN method was selected then this selection was “wrong” revealing that the traffic state estimation accuracy is limited to 73%. The interesting thing here is that, even the selection was wrong the estimation accuracy was ensured to be around 73%.

8. Conclusions and Future Work

This article proposed a notable mobile phone based context-aware traffic state estimation (MC-TES) framework by which inherent issues of *low* and *uncertain* penetration rate were thoroughly resolved. More concretely, a novel intelligent context-aware velocity-density inference circuit (ICIC) was proposed whereby useful contexts extracted from the data reported by mobile devices were utilized to improve traffic state estimation accuracy even in cases of a low penetration rate. This model also minimizes the *critical* penetration rate, thus enhances the reliability as well as the scalability of the traffic state estimation system. This work also proposed a notable ANN-based prediction approach to deal with the issues of an unacceptably *low* and *unknown* penetration rate. This approach is considered as a complement of the ICIC model.

The difficulties on what traffic state estimation model, namely the ICIC or the ANN approach, should be employed to estimate the traffic state in a practical situation was also resolved in this research. A simple yet effective traffic state estimation model selection approach, namely the Approximate Selection method

(AS) was proposed whereby the “right” traffic state estimation model can be selected accurately without any information about the actual penetration rate.

The experimental results reveal the effectiveness of both the ICIC and the ANN models, as well as of the AS method. More concretely, the ICIC model is effective when the penetration rate is low but still relevant. Especially, when the penetration rate is higher than 25%, the ICIC is completely optimized so that the estimation errors are almost lower than 5%. In addition, the ICIC model also minimizes the *acceptable* penetration rate required for a particularly *expected* accuracy. This improvement is practically meaningful since it improves the scalability of the estimation system in terms of low penetration rates. The ANN approach can ensure the estimation accuracy at around 73% regardless of the penetration rate. The AS method provides a high enough traffic state estimation selection accuracy at around 71.79% without any information about the actual penetration rate. The proposed approaches, namely the ICIC model, the ANN approach, and the AS method constitute a complete context-aware traffic state estimation (MC-TES) framework.

This paper primarily proposes solutions for low and uncertain penetration rate issues in mobile phone based traffic state estimation systems. It is due to the fact that penetration rate related issues are inherent issues in mobile phone based applications. In addition, the proposed approaches can be effectively applied in in-vehicle communications systems, vehicle ad-hoc network systems or even in static traffic state estimation systems.

One of the remaining issues in the traffic state estimation selection method, however, is that it still requires estimated values, namely V_{ann} and V_{icic} provided by the candidates (i.e., ICIC and ANN models). It may lower the computational performance. Finding suitable estimation methods where no prior estimation value is required is an interesting research direction. In addition, the effectiveness as well as the robustness of the proposed solutions should be confirmed by more real-field experiments before being applied to real-world applications.

References

- [1] TTI, Texas Transportation Institute: Urban Mobility Information: 2007 Annual Urban Mobility Report, available from (<http://mobility.tamu.edu/ums/>).
- [2] Herring, R., Hofeitner, A., Amin, S., Nasr, T.A., Khalek, A.A., Abbeel, P. and Bayen, A.: *Using Mobile Phones to Forecast Arterial Traffic through Statistical Learning*, The 89th Annual Meeting of the Transportation Research Board (Aug. 2009).
- [3] Shibata, N., Terauchi, T., Kitani, T., Yasumoto, K., Ito, M. and Higashishino, T.: *A Method for Sharing Traffic jam Information Using Inter-Vehicle Communication, ubiquitous, 3rd Annual International Conference on Mobile and Ubiquitous Systems - Workshops*, pp.1–7 (2006).
- [4] University of California, Berkeley: Mobile Millennium Project, available from (<http://traffic.berkeley.edu/theproject.html>).
- [5] Yoon, J., Noble, B. and Liu, M.: *Surface Street Traffic Estimation, MobiSys 07*, San Juan, Puerto Rico, USA. (June 2007).
- [6] Quang, T.M. and Kamioka, E.: *Granular Quantifying Traffic State Using Mobile Probes, IEEE 72nd Vehicle Technology Conference: VTC2010-Fall*, Ottawa, Canada, CD-edition (Sep. 2010).
- [7] Hoh, B., Gruteser, M., Herring, R., Ban, J., Work, D., Herrera, J.C., Bayen, A., Annavaram, M. and Jacobson, Q.: *Virtual Trip Lines for Distributed Privacy-Preserving Traffic Monitoring, MobiSys'08*, Breckenridge, Colorado, USA (June 2008).
- [8] Quang, T.M. and Kamioka, E.: *Traffic State Estimation with Mobile Phones Based on The “3R” Philosophy, IEICE Trans. Communica-*

- tions, Vol.E94-B, No.12, pp.3447–3458 (2011).
- [9] Schneider, W. and Mrakotsky, E.: Mobile Phones as a Basis for Traffic State Information, *Proc. Intelligent Transportation Systems* (2005).
- [10] Quang, T.M. and Kamioka, E.: Adaptive Approaches in Mobile Phone Based Traffic State Estimation with Low Penetration Rate, *Journal of Information Processing (IPJS)*, Vol.20, No.11, pp.297–307 (Jan. 2012).
- [11] Coifman, B.: Improved velocity estimation using single loop detectors, *Trans. Res. A.*, Vol.35, No.10, pp.863–880 (Dec. 2001).
- [12] Tang, S. and Wang, F.Y.: A PCI-based evaluation method for level of services for traffic operational systems, *IEEE Trans. Intell. Transp. Syst.*, Vol.7, No.4, pp.494–499 (Dec. 2006).
- [13] Wright, J. and Dahlgren, J.: Using vehicles equipped with toll tags as probes for providing travel times, California PATH Working Paper UCB-ITSPWP-2001-13, Institute of Transportation Studies, University of California, Berkeley, CA. (2001).
- [14] Ban, X., Li, Y., Skabardonis, A. and Margulici, J.D.: Performance evaluation of travel time methods for real time traffic applications, *11th World Conf. on Transport Research*, Berkeley, CA. (June 2007).
- [15] Togashi, H. and Yamada, S.: Preliminary study on vehicle-to-roadside system using RFIDs for detecting road shoulders, *Intelligent Vehicles Symposium*, Xian, China, pp.1148–1154 (June 2009).
- [16] Cho, Y. and Rice, J.: Estimating velocity fields on a freeway from lowresolution videos, *IEEE Trans. Intell. Transp. Syst.*, Vol.7, No.4, pp.463–469 (Dec. 2006).
- [17] Semertzidis, T., Dimitropoulos, K., Koutsia, A. and Grammalidis, N.: Video sensor network for real-time traffic monitoring and surveillance, *IET Journals on Intelligent Transport Systems*, Vol.4, No.2, pp.103–112 (2010).
- [18] Buch, N., Velastin, S.A. and Orwell, J.: A Review of Computer Vision Techniques for the Analysis of Urban Traffic, *IEEE Trans. Intelligent Transportation Systems*, Vol.12, No.3, pp.920–939 (2011).
- [19] Ministry of Land Infrastructure and Transport: The System Outline of VICS, available from (<http://www.vics.or.jp>).
- [20] Route search and location based consumer services in Japan, available from (<http://www.navitime.co.jp>).
- [21] Kansas ITS, available from (<http://www.ksdot.org/burtransplan/burovr/intrans.asp>).
- [22] Sofra N., Gkelias, A. and Leung, K.K.: Route Construction for Long Lifetime in VANETS, *IEEE Trans. Vehicular Technology*, Vol.60, No.7, pp.3450–3461 (2006).
- [23] Paramics Microsimulation Software for the simulation and visualisation of road traffic, available from (<http://www.paramics.com>).
- [24] Quang, T.M. and Kamioka, E.: Error-Tolerance in Quantifying Traffic States Using Mobile Phones, *10th International Symposium on Autonomous Decentralized Systems (ISADS2011)*, Kobe, Japan, pp.559–563 (June 2011).
- [25] Greenshields, B.D.: A Study in Highway Capacity, *Highway Research Board Proceedings*, pp.448–477 (1935).
- [26] The average size of a car, available from (http://en.wikipedia.org/wiki/Full-size_car).
- [27] Loke, S.: *Context-Aware Pervasive Systems: Architectures for a New Breed of Applications*, Auerbach Publications (2007).
- [28] Akman, V.: *Context in Artificial Intelligence: A Fleeting Overview*, McGraw-Hill, Milano (2002).
- [29] Dey, A.K.: Understanding and Using Context, *Personal and Ubiquitous Computing Journal*, Vol.5, No.1, pp.5–7 (2001).
- [30] Michalewicz, Z.: *Genetic algorithms + data structures = evolution programs*, 3rd ed., Springer-Verlag, Berlin (1996).
- [31] Yalcinoz, T. and Altun, H.: Economic dispatch solution using a genetic algorithm based on arithmetic crossover, *IEEE Porto Power-Tech'01*, Porto, Portugal, pp.10–13 (Sep. 2001).
- [32] Yalcinoz, T. and Altun, H.: A new genetic algorithm approach with arithmetic crossover to economic-environmental power dispatch, *Engineering Intelligent System*, Vol.13, pp.45–52 (2005).
- [33] Zhao, X., Gao, X.S. and Hu, Z.: Evolutionary Programming Based on Non-Uniform Mutation, Sciences, A.o.M.a.S. (Ed.), *Mathematics-Mechanization Research Preprints*, Chinese Academy of Sciences, Beijing, China, pp.352–374 (2004).
- [34] Martin, T.H., Howard, B.D. and Mark, H.B.: *Neural Network Design*, PWS Publishing Company, Boston (1996).
- [35] Rumelhart, D.E., Widrow, B. and Lehr, M.A.: The basic ideas in neural networks, *Commun. ACM*, Vol.37, No.3, pp.87–92 (1994).
- [36] Gora, P.: Traffic Simulation Framework -A Cellular Automaton-Based Tool for Simulating and Investigating Real City Traffic, *Recent Advances in Intelligent Information Systems*, ISBN 978-83-60434-59-8, pp.641–653 (2009).



Quang Tran Minh is a researcher at Principles of Informatics Research Division, National Institute of Informatics, Japan. He received his Dr. Eng. in Functional Control Systems (2012) from Shibaura Institute of Technology, Japan, his B. Eng. (2001), M. Eng. (2006) in Computer Science and Engineering from

Hochiminh City University of Technology, Vietnam, and Ritsumeikan University, Japan, respectively. He has been a Lecturer at Hochiminh City University of Technology since 2003. His current research interest includes mobile computing, ubiquitous computing, ad hoc network, and advanced information systems such as data mining, intelligent transportation systems.



Muhammad Ariff Baharuddin obtained his B. Eng. (Hons) in Electrical Engineering in 2008 and M. Eng. Electrical (Telecommunication) in 2010 from Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia. He is currently pursuing his Ph.D. in Shibaura Institute of Technology, Japan. He is a staff of the

Faculty of Electrical Engineering, UTM. His research interests include network communication protocols, Smart Grid, and network optimization technique.



Eiji Kamioka is a Professor at Shibaura Institute of Technology. He received his B.S. (1989), M.S. (1991) and D.S. (1997) degrees in Physics from Aoyama Gakuin University. He also has other job experiences, working for SHARP Communications Laboratory (1991–1993), Institute of Space and Astronautical Science

(ISAS) as a JSPS Research Fellow (1997–1998) and National Institute of Informatics (NII) as an Assistant Professor (1998–2006). His current research interests encompass mobile multimedia communications, ubiquitous computing and context-aware applications.