Preceding Vehicle Identification for Cooperative Adaptive Cruise Control Platoon Forming

Zheng Chen and Byungkyu Brian Park, *Senior Member, IEEE*

***Abstract*— Cooperative adaptive cruise control (CACC) has shown great potential in enhancing traffic efficiency and sustainability. While past research efforts focused on the development of CACC systems and their demonstrations, very few of them considered in detail how to form a CACC platoon in real traffic, where the proper identification of preceding vehicle is required. To ensure safe and reliable CACC operations, the following vehicle needs to establish the correct connection with its preceding vehicle. Although this can be done by matching information shared by surrounding vehicles with the ego-vehicle's radar measurements, the existence of sensor/global positioning system (GPS) errors makes it a challenging task. Considering possible sensor/GPS errors in real traffic, this paper proposes a procedure of identifying preceding vehicle under fully connected vehicle environment and evaluates three preceding vehicle identification systems (PVIS), namely, location-based PVIS, distance-based PVIS, and integrated PVIS combining both location and distance information. The mathematical models of PVISs are developed. The performance evaluation of the PVISs is conducted based on real vehicle trajectory data from the Next Generation Simulation (NGSIM) program, which reflects how vehicles' relative positions change in a high-density segment of highway. The feasibility, performance, and potential of the three PVISs are compared. The results show that location-based PVIS requires a relative positioning accuracy below 1.1 m to ensure an acceptable identification time with zero failure rate. The integrated PVIS has the best performance, providing 99% confidence in identifying preceding vehicle within 1.3 s under typical sensor error settings.**

***Index Terms*— Cooperative adaptive cruise control (CACC), platoon forming, preceding vehicle, identification, sensor errors.**

I. Introduction

DEPLOYMENT of connected and automated vehicle (CAV) is considered the key factor to enhance traffic safety and efficiency in intelligent transportation systems (ITS). Automated vehicle (AV) has been developed for long time and some partly-automated driving applications such as adaptive cruise control (ACC) have been commercialized and equipped in a portion of new cars [1]. Using onboard sensors (e.g. radar) ACC can automatically control the vehicle's longitudinal speed to maintain a safe gap from the preceding vehicle. AV can be upgraded to CAV when vehicle-to-vehicle (V2V) and vehicle-infrastructure (V2I) communications are enabled. To promote vehicles' connectivity, the U.S., European Union, and Asian countries have respectively specified dedicated spectrum for automotive use. In 2010, Institute of Electrical and Electronics Engineers (IEEE) approved the IEEE 802.11p standard to enable wireless access in vehicular environments (WAVE).

Cooperative Adaptive Cruise Control (CACC), as the extension of ACC by introducing V2V communications, has been proven a promising CAV application. CACC allows vehicles to travel in more compact and stable platoons than ACC or human driver can do. Previous studies revealed the potential benefit of CACC, including significant improvement in the roadway throughput [2], [3], and reduction in fuel consumptions [4], [5].

A variety of CACC systems with different communication protocols, control algorithms and applied scenarios were proposed [6]. Most CACC systems require the following vehicle to communicate with its nearest preceding vehicle or/and the leading vehicle of the platoon. In the past decade, there have been considerable research efforts on the development and validation of CACC systems. Field test with uniform vehicles and control structure can be seen in [4], [5], and [7]–[10]. The Grand Cooperative Driving Challenge (GCDC) held by the Netherlands Organization for Applied Scientific Research (TNO) in 2011 was the first attempt to implement heterogeneous CACC vehicles together in realistic scenarios [11]. A positive finding in the GCDC supported by communication and positioning devices is that CACC implementation built on top of an existing ACC system is both technically and economically feasible [12]. Although the GCDC organizer required equipment of high-performance positioning device such as real-time kinematic global positioning system (RTK-GPS), it has been proven in previous CACC field tests that high-cost differential GPS is not indispensable for maintaining stable CACC platoons [7], [8], [10], and GPS positioning is even not needed when CACC vehicles only communicate with its immediate preceding vehicle [9]. This means CACC can be performed on partly-automated vehicles without substantially increasing hardware cost. However, several lingering issues of CACC have been found in GCDC including communication unreliability and flawed or missing data from other vehicles [13]. These problems must be adequately addressed before large-scale deployment of CACC, and thus they have drawn additional research attentions [14].

Manuscript received March 12, 2018; revised June 20, 2018, September 3, 2018, and October 25, 2018; accepted December 14, 2018. Date of publication January 25, 2019; date of current version December 31, 2019. This work was supported in part by the GRL Program through NRF of Korea under Grant 2013K1A1A2A02078326 and in part by KAIA Grant funded by the Korean Government (MOLIT) under Project 17TLRP-B117133-02. The Associate Editor for this paper was J. E. Naranjo. *(Corresponding author: Zheng Chen.)*

The authors are with the Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA 22904 USA (e-mail: zc4ac@virginia.edu; bp6v@virginia.edu).

Digital Object Identifier 10.1109/TITS.2019.2891353

Another fundamental issue that has neither been fully identified nor addressed is how to guarantee the correct connection between each pair of following vehicle and preceding vehicle, especially at the beginning of forming a CACC platoon. To be more specific, when a vehicle (denoted as subject vehicle) detects another vehicle ahead and starts platooning with this preceding vehicle, it has to first find the communication identity of the preceding vehicle. This identification process is absolutely necessary for vehicles that want to join the CACC platoon by connecting to the nearest preceding vehicle. While it might seem trivial, it is not an easy task at all. In the real traffic of connected vehicles, the subject vehicle is likely to find multiple connected vehicles in the vicinity that can be communicated with, and it may be difficult to identify the preceding vehicle due to sensor/GPS errors. In practice, the surrounding connected vehicles can share their GPS position with the subject vehicle. The straightforward way is to find the closest vehicle ahead in the same lane according to the reported GPS positions of surrounding vehicles. This can be seen in GPS/GNSS-based CACC [4] which doesn't have a radar. A major risk in this method is that when the actual preceding vehicle reports a false lane number, the subject vehicle would take the vehicle further ahead as the preceding vehicle and conduct an incorrect gap-closing maneuver. This is an even more dangerous error than taking the vehicle in the adjacent lane as the preceding vehicle. Thus, high-performance GPS/GNSS must be there to guarantee in-the-lane positioning. For most of CACC systems built on top of ACC with radar, the less risky approach is to compare the self-reported positions of surrounding vehicles with the radar-measured position of the preceding vehicle. The surrounding vehicle whose self-reported position matches the subject vehicle's sensor measurement should be considered the preceding vehicle. Once the preceding vehicle is successfully identified, such process does not need to be redone until the radar sensor detects a replacement of the preceding vehicle. For the higher probability of implementation, this study is focused on the radar-based CACC.

As noted, the challenge of identifying the preceding vehicle is due to errors of the surrounding vehicles' self-reported positions and subject vehicle's sensor measurements. When these errors are significant enough, false identification of preceding vehicle could happen. In this case, the subject vehicle would perform CACC based on another irrelevant vehicle's message, leading to possibly severe consequences. Previous studies have not explicitly adressed this issue. For example, California Partners for Advanced Transit and Highways (PATH) program has developed and tested two generations of CACC systems [5], [7], [10], but both of them more focused on evaluating an existing CACC platoon instead of forming a platoon. Actually, in their field tests the order of vehicles in the platoon was predefined. This means a subject vehicle could know which vehicle was the preceding vehicle that it should communicate with. Furthermore, the platoon was running in a single lane, which made it a pushover to distinguish two different vehicles based on the self-reported positions because of their large longitudinal distance. Similar field test settings can also be seen in [4], [8], and [9]. In GCDC 2011, field tests of two groups of CACC platoons traveling on two adjacent lanes was conducted. Although high-accuracy positioning had already been provided by RTK-GPS, participants took a more conservative strategy to form the platoon: a blacklist was preset manually in the vehicle's software, based on the predefined vehicle sequence, to block messages from vehicles in the adjacent lane [15]. However, this blacklist is obviously unavailable in real life.

Fully automated vehicles may not be confronted with this problem because of its centi-meter level positioning capability [16], but as previously stated, running CACC requires neither fully automated vehicles nor high-cost positioning devices. There is no need to wait decades for massive production of fully automated vehicles to implement CACC as long as the issue of preceding vehicle identification can be solved. Therefore, it is necessary to investigate how accurate the self-reported vehicle positions should be to build correct connection between the following vehicle and the preceding vehicle when CACC platoon is being formed, and whether the current low-cost positioning device (e.g. plain GPS) can meet the requirement in real traffic environment. It is also important to explore other possible techniques that help identify preceding vehicle more efficiently than pairing self-reported GPS positions and sensor measurements.

The rest of the paper is organized as follows: Section 2 formalizes the identification procedure and proposes three preceding vehicle identification systems (PVIS) including location-based PVIS, distance-based PVIS, and integrated PVIS. In addition, the vehicle trajectory data from US Highway 101 and onboard sensors' characteristics that drive the evaluation of PIVSs are introduced. Section 3 presents mathematical models of the three PVISs. In Section 4, these models along with real world vehicle trajectory data are used to evaluate the performance of PVISs. Finally, conclusions drawn from simulation results and limitations of the study are presented in Section 5.

II. Methodology

A. Identification Procedure

There are several main assumptions in this study:

a. 100% connected vehicle environment, i.e., all the vehicles are equipped with V2V communication device that can transmit and receive messages;
b. Vehicles are traveling in parallel with lane lines;
c. The PVIS user (subject vehicle) is assumed to be automated vehicle and equipped with radar.

When a preceding vehicle is detected by the sensors of subject vehicle, the subject vehicle is supposed to compare the preceding vehicle's status estimated by sensors (denoted by S) with the status of target vehicles (i.e. surrounding un-platooned vehicles and the last vehicles in platoons) obtained by communication (denoted by C). If for i^{th} target vehicle $C_i = S$, then this vehicle should be considered a candidate for preceding vehicle. Ideally, there are no two vehicles with the exact same status. Thus, there should always be only one target vehicle being candidate for preceding vehicle. However, due to the limited resolution and accuracy

TABLE I
CONFIGURATION OF LOCATION-BASED PVIS AND DISTANCE-BASED PVIS

PVIS type	Location-based	Distance-based
Required equipment	GPS, DSRC, radar	UWB, radar
Representative of S and C	Preceding vehicle's location relative to subject vehicle	The distance between preceding vehicle and subject vehicle
Source of S	Radar	Radar
Source of C	GPS	UWB
Communication	DSRC	UWB

of sensors in real world, the condition of being a candidate has to be degraded to:

$$|C_i - S| < \in \tag{1}$$

where $\in$ is a preset threshold so that the correct preceding vehicle has a low probability to be missed or mismatched (i.e., low failure rate), and yet screens out as many irrelevant vehicles as possible. Considering the hazard consequences that false identifications of preceding vehicle may lead to, the proposed PVIS' theoretical failure rate is set to be 1 per hundred million identification attempts (i.e., 10^{-8} probability to make mistake per attempt). This number is chosen to consider the fatality rate of human driving, that is 1.09 per hundred million miles [17].

It should be noted that 10^{-8} only represents the probability that the actual preceding vehicle does not satisfy (1). It is still possible that other target vehicles have similar status with the preceding vehicle thus satisfy (1) and become candidates for the preceding vehicle. In this case, the subject vehicle cannot immediately distinguish the preceding vehicle from them. It has to take more attempts in the following time until status of all "other vehicles" become significantly different from S and only one candidate (i.e., the actual preceding vehicle) satisfies (1). Obviously, a desirable preceding-vehicle identification system should screen out all non-preceding vehicles as quickly as possible.

Although preceding vehicle's status can be represented by a variety of measures captured by sensors and/or communication, not all of them are suitable for vehicle identification. For example, one should not set (1) with S and C being vehicle speeds. This is because there are too many vehicles traveling in a similar speed, thus make it difficult to distinguish preceding vehicle from others.

As noted, three preceding vehicle identification systems (PVIS), namely, location-based PVIS, distance-based PVIS, and integrated PVIS combining location and distance measurements are proposed and will be evaluated. As shown in Table I, location-based PVIS measures the relative location of preceding vehicle (i.e., S) by radar. Simultaneously, it estimates the subject vehicle's location by GPS and communicates with all target vehicles to obtain their GPS information, thus decides the relative location of those vehicles (i.e., C). It is assumed that the V2V communication in location-based PVIS is enabled by dedicated short range communication (DSRC), which complies with IEEE 802.11p at the physical (PHY) and medium access control (MAC) layers [18].

On the other hand, distance-based PVIS identifies the preceding vehicle based on inter-vehicle distance. This PVIS is taking advantage of the rise of accurate wireless ranging technology such as Ultra-wideband (UWB) communication [19]. Distance-based PVIS is supposed to measure the distance from the preceding vehicle (i.e., S) by radar, and communicates all target vehicles via UWB to estimate the distance from those vehicles (i.e., C). The recent progress in 5th generation wireless systems (5G) shows that millimeter-wave 5G has an excellent capability of ranging [20]–[22]. Therefore, the communication and distance measurement between vehicles may also be realized by 5G in the future. Integrated PVIS uses both relative location from GPS and inter-vehicle distance based on UWB or 5G.

These three PVISs will be modeled and simulated in a highway scenario using real vehicle trajectory data available from NGSIM. The simulation scenario is chosen to be a rush-hour highway with fairly dense vehicles. This is to investigate whether these PIVSs can accurately and quickly distinguish the preceding vehicle from other target vehicles even in fairly congested environment. The time consumed by PVIS to make correct identifications (i.e., time consumption, for short) and the proportion of incorrect identifications are used to measure the effectiveness of PVISs.

B. NGSIM Data

To evaluate the performance of PVISs, the natural vehicle trajectory data from Next Generation Simulation (NGSIM) program is used to reconstruct the real-world traffic condition.

NGSIM program [23] was launched by FHWA's Traffic Analysis Tools Program. It used high-resolution cameras to record trajectory of the vehicles on the real roads. The US Highway 101 (US 101) dataset was one dataset that reflected highway traffic condition. It recorded the precise trajectory of vehicles in all 6 lanes within the 640-meter long study area, resulting in detailed relative positions of vehicles that can be used for PVIS evaluation.

A total of 45 minutes of data are available in the full dataset of US101, segmented into three 15-minute periods: 7:50 a.m. to 8:05 a.m.; 8:05 a.m. to 8:20 a.m.; and 8:20 a.m. to 8:35 a.m. These periods represent the buildup of congestion, the transition between uncongested and congested conditions, and full congestion during the rush hour. Considering that CACC can greatly decrease the stop-and-go conditions in heavy traffic [24], only the first 15-minute period data was used. This is because most vehicles kept moving and with only a few stop-and-go conditions during the first 15-minute period, which is when PVIS can effectively help implementing CACC to improve mobility.

More than 3000 vehicles passed the study area, and 2500 pairs of preceding/following vehicles (with longitudinal distance $<$100m) formed in this period. Pairs of preceding/following vehicles that have already formed before entering the study area are excluded. The detailed trajectories of these

preceding/following vehicles as well as their surrounding vehicles are used to generate the situations where preceding vehicle identification is needed to initialize a CACC platoon in connected vehicle environment.

C. Sensors and Communications

To make the PVIS models and performance evaluation more realistic, sensor errors and communication imperfections need to be adequately assumed.

1) Automotive Radar Error: Millimeter-wave radar sensor has been equipped for active safety and driver assistance systems. High-accuracy radar sensors are usually working on the frequency band of 76-77 GHz. A typical commercial radar sensor is Bosch long range radar (LRR) which is commonly used for ACC [25]. It can detect the location of objects in front within 250m and 15° on both left and right sides, with distance accuracy of 0.1m, and angle accuracy of 0.1°. The settings of radar in PVIS evaluations are following this radar sensor's parameters.

2) GPS Positioning Error: Global Positioning System (GPS), one of Global Navigation Satellite Systems (GNSSs), is the most widely-used vehicle positioning tool. A typical error of plain GPS is 2∼5m on open road and up to 10∼15m in urban environment [26], which is insufficient for safety-critical vehicular applications. To improve the positioning accuracy, many GPS augmentation methods are proposed. The conventional one is differential GPS (DGPS). DGPS uses a reference station to broadcast the correction messages which contain their observations of GPS signal and the known fixed positions so that the surrounding receiver can use the message to correct their positions. This process eliminates a large portion of positioning errors from ionosphere, troposphere, GPS satellite orbit errors, and satellite clock drifts, because GPS receivers in the vicinity are affected by these sources of errors in a similar way [27]. According to the content of correction message, there are two types of DGPS: real-time kinematic (RTK)-GPS and pseudo-range DGPS.

RTK-GPS base station broadcasts its observation of the carrier wave of GPS signal and its location. Receivers can correct their positions with centi-meter accuracy using the correction message and their own observations of carrier wave. However, the complexity of observing and computing carrier phase makes RTK-GPS expensive [26] and less likely for large-scale deployment.

Pseudo-range DGPS is at low cost because it requires neither complicated signal observation nor processing. Pseudo-range refers to the measured distance between a satellite and a GPS signal receiver. A receiver uses pseudo-ranges from 4 different satellites to determine its position. In practice, DGPS reference stations can easily calculate the differences between the measured satellite pseudo-ranges and actual distances (based on their known positions). The surrounding receivers just need to correct their pseudo-ranges by the same amount, and positioning error can be reduced to sub-meter level in ideal situation [28].

By exchanging GPS/DGPS measurement data, relative position between vehicles can be obtained. There have been plenty of research efforts focused on the performance of vehicle relative positioning by low-cost GPS/DGPS. Two relative positioning methods under urban environment are compared in [29] via field experiments. The first one is taking the difference between two vehicles' absolute positions obtained by pseudo-range DGPS, which led to 10.2m root mean square error (RMSE). The second one is to exchange two vehicles' pseudo-range measurements and directly calculate their relative position. In this case, the subject vehicle itself plays as a "dynamic" base station, and no extra DGPS infrastructure is needed. The second method achieved 6.5m RMSE. De Ponte Müller *et al.* [30] tested the performance of baseline vector measurement based on the pseudo-range exchange. The mean error of baseline length was 1.17m in highway scenario, and 2.16m in urban scenario. A weighted least squares double difference method proposed in [31] to detect baseline by sharing of GPS pseudo-range measurements. The real-world tests showed 0.57m mean distance error in open field, 4.06m on the roadside with trees, and 2.31m on the roadside with buildings.

In addition, it's found that the distribution of pseudo-range DGPS's errors in three dimensions (X, Y, and Z) are separately complying with zero-mean normal distributions from long-term observations [32]. It is then reasonable to assume that pseudo-range relative positioning has this same characteristic with pseudo-range DGPS due to their similarity.

To sum up, a normal distribution with standard deviations (STD) varying from 0.5m to 6.5m can roughly cover the long-term positioning errors of the pseudo-range relative positioning approach in different scenarios. This pseudo-range relative positioning approach is assumed to be used in the PVISs because it only requires plain GPS receivers and limited amount of data transmission but has better performance than low-cost DGPS. Due to pseudo-range (double) differencing, the positioning error caused by satellite clock error, user clock error, ephemeris, tropospheric and ionospheric are canceled out [30], thus the main error term would be multipath error which is caused by roadside obstacles. According to [33], the multipath effect can be considered as abruptly adding biases of random magnitudes and durations to pseudo-range measurements. Therefore, relative positioning error in this research is modeled by:

$$\begin{aligned} e(t) &\approx e_m(t) + \varepsilon \\ e_m(t) &= \sum\nolimits_1^i \mathrm{F}(t, t_{i-1}, t_i)\, b_i \\ \mathrm{F}(t, t_{i-1}, t_i) &= \begin{cases} 1 & t_{i-1} < t \le t_i \\ 0 & else \end{cases} \end{aligned} \tag{2}$$

where t is time, $e(t)$ is the overall positioning error, $e_m(t)$ is the multipath effect error, and ε is normally distributed unmodeled error:

$$\varepsilon \sim N(0, \sigma_u^2)$$

where σ_u is the STD of the unmodeled error and also the minimum STD of overall error when there is no multipath effect.

Multipath effect error $e_m(t)$ is considered as random bias that has normally distributed magnitude b_i and uniformly

distributed duration $t_{i-1} - t_i$:

$$b_i \sim N(0, \sigma_b^2)$$
$$(t_{i-1} - t_i) \sim U(t_{min}, t_{max})$$

where σ_b is the STD of positioning bias, and t_{min} and t_{max} are lower and upper boundary of the duration of the multipath effect.

While the error propagation from pseudo-range measurements to position measurement is simplified, the model (2) makes sense that from long-term observations the overall positioning error tends to comply with an unbiased normal distribution with a STD of $\sqrt{\sigma_u^2 + \sigma_b^2}$, but from short-term observations the overall error is biased and skewed which makes its RMSE fluctuate over time.

It should be noted that (2) can only be used for PVIS evaluation, not for the design of PVIS, because in real world the only information the vehicles can know is the statistics of overall error, either from long-term field tests or onboard sensor fusion [34].

3) UWB Ranging Error: Ultra-wideband (UWB) is an emerging technology for transmitting information with signal bandwidth exceeds 500 MHz or 20% of the arithmetic center frequency (defined by U.S. Federal Communications Commission). In 2006, Institute of Electrical and Electronics Engineers (IEEE) established the IEEE 802.15.4a standard for UWB. Due to the large bandwidth, UWB signals have extremely high time resolution, which renders UWB the ability to measure the distance between objects accurately even in multipath-effect environment. UWB has been widely used for indoor and outdoor positioning [35], [36]. While UWB can work in both line-of-sight (LOS) and non-line-of-sight (NLOS) scenarios, preceding vehicle identification is a typical LOS scenario because antennas are mounted on the top of vehicles and the subject vehicle is adjacent to the preceding vehicle.

Ranging capability of an IEEE 802.15.4a compliant UWB prototype from Decawave is tested in [37]. It shows outstanding LOS accuracy both in open field (ranging error <0.1m) and multipath environment (ranging error <0.2m). Kristem *et al.* [38] investigated the ranging accuracy of UWB in dense urban environments with a pair of Skycross Omni-directional antennas. Experiments showed that RMSE varied from 0.12m to 0.14m for LOS measurements. Performance comparison of two latest commercial IEEE 802.15.4a UWB modules, BeSpoon UM100 and Decawave DW1000, are conducted in [39]. The outdoor LOS tests showed the errors of two modules complied with zero-mean normal distributions. The STD are 0.11m and 0.055m respectively.

Because UWB is highly resistant to multipath effect in LOS scenarios [35], the multipath bias of UWB is assumed insignificant, and the ranging error of state-of-art UWB modules is simply modeled by an unbiased normal distribution with STD from 0.05m to 0.2m, to cover all the aforementioned errors of different UWB products in different environments.

4) Transmission Latency and Packet Drop: There are two types of communication devices in this study: DSRC for GPS-based PVIS, and UWB for UWB-based and integrated PVIS.

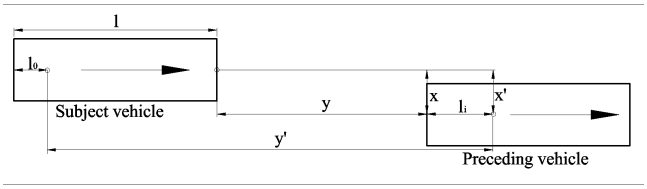

Fig. 1. Variables that need to be measured or obtained by GPS-based PVIS.

It has been found that DSRC has neglectable transmission latency (<5ms), but considerable amount of packet drops may happen [40], [41], depending on multiple factors such as node density, data rate, and operating environments. A minimal requirement on DSRC specified by [42] is that packet delivery rate should not be lower than 90%. Therefore, to consider the worst case, a packet loss rate of 10% is assumed in the evaluation of GPS-based PVIS.

As for UWB, both the low transmission latency (<10ms) and high packet delivery rate (>99.95%) can be achieved when communicating with up to 25 terminals [43], [44], so the effects of transmission latency and packet drop of UWB are assumed insignificant in this study.

III. Mathematical Models

A. GPS-Based Identification

It is assumed that:

1) Radar and GPS antenna are installed on the longitudinal centerline of each vehicle;
2) The subject vehicle knows the distance between its GPS antenna and vehicle tail, which is denoted by l_0, and the vehicle length, which is denoted by l; i^{th} target vehicle knows the distance between its GPS antenna and vehicle tail, which is denoted by l_i;
3) Subject vehicle can obtain surrounding target vehicle's GPS information and l_i through DSRC;

We first consider the situation without packet drop. Based on the measurement of radar on subject vehicle, the tail of preceding vehicle can be located:

$$y_r = y + e_r;$$
$$x_r = x + y \cdot e_a;$$

where y_r/x_r are the radar-measured longitudinal/lateral location of preceding vehicle's tail relative to subject vehicle's radar; y/x are the actual longitudinal/lateral location of preceding vehicle's tail relative to subject vehicle's radar, as shown in Fig. 1; e_r and e_a are radar's distance error and angle error respectively.

Based on GPS's measurement, GPS antenna of i^{th} target vehicle can be located:

$$y'_g = y' + e_y;$$
$$x'_g = x' + e_x;$$

where y'_g/x'_g are the GPS-measured longitudinal/lateral distance between the GPS antennas of subject vehicle and i^{th} vehicle; y'/x' are the actual longitudinal/lateral distance between two the vehicles' GPS antennas as shown in Fig. 1;

e_x and e_y are V2V relative positioning error in lateral direction and that in longitudinal direction respectively.

The location of i^{th} vehicle's tail relative to the radar can also be estimated based on GPS measurement and vehicle geometry:

$$y_g = y'_g - l + l_0 - l_i = y' + e_y - l + l_0 - l_i;$$
$$x_g = x'_g = x' + e_x;$$

y_g and x_g are location of i^{th} vehicle's tail relative to the radar estimated based on GPS measurement.

If i^{th} vehicle is just the preceding vehicle, then:

$$y = y' - l + l_0 - l_i;$$
$$x = x';$$

Thus, the location of i^{th} vehicle's tail relative to radar based on GPS's measurement can be transformed to:

$$y_g = y + e_y;$$
$$x_g = x + e_x;$$

The y_g and x_g for preceding vehicle should be close to y_r and x_r, thus the condition of being a candidate for preceding vehicle would be:

$$\sqrt{(x_g - x_r)^2 + (y_g - y_r)^2} \approx \sqrt{(e_x - y_r \cdot e_a)^2 + (e_y - e_r)^2} < \text{threshold} \quad (3)$$

The "$\approx$" is used in (3) because the actual y is not available in practice thus it's approximately replaced with y_r.

To find a proper threshold, we assume the sensors' errors comply with normal distribution:

$$e_x \sim N(0, \sigma_x^2)$$
$$e_y \sim N(0, \sigma_y^2)$$
$$e_r \sim N(0, \sigma_r^2)$$
$$e_a \sim N(0, \sigma_a^2)$$

where σ_x, σ_y, σ_r and σ_a are those errors' standard deviations.

As a result:

$$(x_g - x_r) \sim N(0, \delta_x^2)$$

where $\delta_x = \sqrt{(y_r \cdot e_a)^2 + {e_x}^2}$;

$$(y_g - y_r) \sim N(0, \delta_y^2)$$

where $\delta_y^2 = \sqrt{{e_y}^2 + {e_r}^2}$.

Then $\left(\frac{x_g - x_r}{\delta_x}\right)^2 + \left(\frac{y_g - y_r}{\delta_y}\right)^2$, as a squared sum of two independent standard normal variables, complies with Chi-square distribution with 2 degrees of freedom.

If the subject vehicle requires an identification failure rate of α, and takes only one measurement to determine candidate(s) for preceding vehicle, condition (3) would become:

$$\left(\frac{x_g - x_r}{\delta_x}\right)^2 + \left(\frac{y_g - y_r}{\delta_y}\right)^2 < \chi^2(2, \alpha) \quad (4)$$

Which can be then transformed back to:

$$\sqrt{(x_g - x_r)^2 + (y_g - y_r)^2} < \sqrt{\frac{\delta_x^2 \cdot \delta_y^2 \cdot \chi^2(2, \alpha) - \left(\delta_x^2 - \delta_y^2\right) \cdot (y_g - y_r)^2}{\delta_y^2}}$$

where $\chi^2(2, \alpha)$ is the Chi-square statistic for 2 degrees of freedom and P-value of α.

If the subject vehicle takes multiple steps to identify preceding vehicle, that is, taking n times of measurement consecutively to determine candidate(s) for preceding vehicle, (4) would change to:

$$\left(\frac{x_g - x_r}{\delta_x}\right)^2 + \left(\frac{y_g - y_r}{\delta_y}\right)^2 < \chi^2\left(2, \sqrt[n]{\alpha}\right) \quad (5)$$

In this case, only vehicles that satisfy condition (5) for n consecutive times would be considered candidates for preceding vehicle.

Now, we consider the situation where the messages from target vehicles have 10% chance to be randomly dropped. It would be dangerous to determine the candidates for preceding vehicle when packet drop occurs since the dropped packet may come from the actual preceding vehicle. However, under 10% packet loss rate it is also infeasible to let PVIS just wait until next time frame to make decision. This is because PVIS may work in high-density traffic with a lot of target vehicles, and in most of time PVIS may only receive a portion of the messages from the target vehicles. As a result, PVIS would be doing nothing but wait for long time.

One adjustment is to make decision on the vehicles whose messages are dropped based on the previous candidate list. To be more specific, when a vehicle's message is dropped but it has been determined to be a candidate at last time frame, it would still be considered a candidate at this time frame, and vice versa. An additional requirement is that the actual preceding vehicle must be on the candidate list at the very beginning. To make this happen, a preparation phase of 10 time frames is added prior to the identification procedure. During these time frames, PVIS only collects ID of surrounding vehicles but doesn't make any decision. It then takes all the collected vehicle ID onto the initial candidate list. It can be computed that the actual preceding vehicle has only 10^{-10} probability to be excluded from the initial candidate list.

Afterwards, the actual preceding vehicle can remain on candidate list with guaranteed failure rate even when its message is dropped, while the irrelevant vehicles that has been screened out will not keep reappearing on the candidate list due to packet losses.

B. UWB-Based Identification

It is assumed that subject vehicle can measure the distance to every i^{th} target vehicle and obtain its l_i by UWB communications.

Based on radar's measurement of subject vehicle, the distance to the tail of preceding vehicle is:

$$d_r = d + e_r;$$

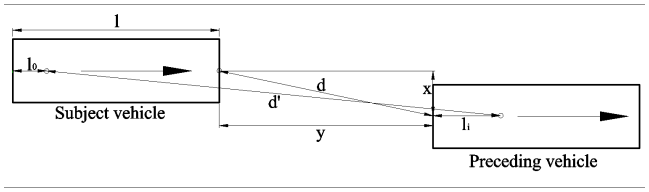

Fig. 2. Variables that need to be measured or obtained by UWB-based PVIS.

where d_r is the measured distance between the preceding vehicle's tail and subject vehicle's radar; d is the actual distance between the preceding vehicle's tail and subject vehicle's radar, as shown in Fig. 2; e_r is radar's distance error.

Based on UWB's measurement, the distance to i^{th} vehicle's UWB antenna can be obtained:

$$d'_u = d' + e_u;$$

where d'_u is the measured distance between the preceding vehicle's UWB antenna and subject vehicle's UWB antenna; d' is the actual distance between the preceding vehicle's UWB antenna and subject vehicle's UWB antenna, as shown in Fig. 2; e_u is the ranging error of UWB.

If i^{th} vehicle is the preceding vehicle, then:

$$d'^2 - ((l - l_0 + l_i) + y)^2 = d^2 - y^2;$$

Thus:

$$d' = \sqrt{d^2 + (l - l_0 + l_i)^2 + 2y\,(l - l_0 + l_i)};$$

Therefore, given the vehicle geometry, the distance between the preceding vehicle's UWB antenna and subject vehicle's UWB antenna can also be estimated by radar:

$$d'_r = \sqrt{{d_r}^2 + (l - l_0 + l_i)^2 + 2y_r\,(l - l_0 + l_i)}$$

where d'_r is the distance between the preceding vehicle's UWB antenna and subject vehicle's UWB antenna estimated by radar.

The error of d'_r can be derived by error-propagation formula:

$$e(d'_r) = \frac{\partial d'_r}{\partial d_r} \cdot e_r + \frac{\partial d'_r}{\partial y_r} \cdot e_r = (\frac{d_r}{d'_r} + \frac{l - l_0 + l_i}{d'_r}) \cdot e_r$$

And thus:

$$d'_r \approx d' + e(d'_r) = d' + (\frac{d_r}{d'_r} + \frac{l - l_0 + l_i}{d'_r}) \cdot e_r$$

The condition of being a candidate for preceding vehicle can be set:

$$|d'_u - d'_r| = |e_u - (\frac{d_r}{d'_r} + \frac{l - l_0 + l_i}{d'_r}) \cdot e_r| < \text{threshold} \quad (6)$$

To find a proper threshold, we assume the error of UWB ranging comply with normal distribution:

$$e_u \sim N(0, \sigma_u^2)$$

where σ_u, is the UWB ranging error's standard deviations.

Then, $e_u - (\frac{d_r}{d'_r} + \frac{l - l_0 + l_i}{d'_r}) \cdot e_r$, as a combination of two independent normal variables, complies with a normal distribution whose standard deviation is:

$$\sigma = \sqrt{\sigma_u^2 + (\frac{d_r}{d'_r} + \frac{l - l_0 + l_i}{d'_r})^2 \cdot \sigma_r^2}$$

If the subject vehicle pursues a failure rate of α, and takes only one time of measurement to determine candidate(s) for preceding vehicle, condition (6) would become:

$$|d'_u - d'_r| < Z_{\alpha/2} \cdot \sigma \quad (7)$$

where $Z_{\alpha/2}$ is the Z statistic for and two-tail Z test with P-value of α.

If the subject vehicle takes n times of measurement consecutively to determine candidate(s) for preceding vehicle, (7) would change to:

$$|d'_u - d'_r| < Z_{\sqrt[n]{\alpha}/2} \cdot \sigma \quad (8)$$

Candidates for preceding vehicle need to satisfy condition (8) for n consecutive times.

C. Integration of UWB and GPS

Given that vehicles are possible to be equipped with both GPS and UWB, combining both measurements may help identify preceding vehicle better. In this scenario, the candidate(s) for preceding vehicle should satisfy both of conditions below:

$$\begin{cases} |d'_u - d'_r| < Z_{k\cdot\alpha/2} \cdot \sigma \\ \left(\frac{x_g - x_r}{\delta_x}\right)^2 + \left(\frac{y_g - y_r}{\delta_y}\right)^2 < \chi^2\,(2, (1-k)\cdot\alpha) \end{cases} \quad (9)$$

where k is a weight factor that $0 < k < 1$; it indicates how much the system relies on UWB over GPS. α is still the failure rate of the system, which can be proved by:

$$1 - (1 - k \cdot \alpha) \cdot (1 - (\alpha - k \cdot \alpha)) \approx k \cdot \alpha + (\alpha - k \cdot \alpha) = \alpha$$

If the subject vehicle takes n times of consecutive measurements to determine candidate(s) for preceding vehicle, (9) would change to:

$$\begin{cases} |d'_u - d'_r| < Z_{\sqrt[n]{\beta}/2} \cdot \sigma \\ \left(\frac{x_g - x_r}{\delta_x}\right)^2 + \left(\frac{y_g - y_r}{\delta_y}\right)^2 < \chi^2\,(2, \sqrt[n]{\alpha - \beta}) \end{cases} \quad (10)$$

where $\beta = k \cdot \alpha$.

IV. Performance Evaluation

A. Simulation Design

The proposed preceding vehicle identification systems were modeled and tested in MATLAB with NGSIM data. The simulation followed the steps shown in Fig. 3 and is described as below:

1) Initialization. The simulation is driven by two parameters: t and j. Parameter t means time, ranging from 0 to 900.0 seconds with a resolution of 0.1 second, because NGSIM data was recorded at this resolution. Communication and GPS positioning are also assumed

to be conducted every 0.1 second. At any given t, parameter j means the number of the vehicle on the road, ranging from 1 (which denotes the vehicle closest to the entrance) to j_{max} (which denotes the vehicle farthest to the entrance). At the beginning of simulation, t and j were set 0 and 1 respectively;

2) Obtain vehicle status from NGSIM data. NGSIM data was imported to MATLAB as a matrix of 1,048,575 rows and 8 columns. Each row represents a vehicle's status at given t, consisting of Vehicle ID, Time, Lateral coordinate, Longitudinal coordinate, Vehicle length, Vehicle width, Vehicle velocity, and Lane number. Lateral coordinate and Longitudinal coordinate were measured from the entrance of the road segment. All the rows were sorted by time from the smallest to the largest; rows with the same time were sorted by Longitudinal coordinate from the smallest to the largest. With an input of t and j, the state of j^{th} vehicle (i.e. subject vehicle) and states of surrounding vehicles within 200m (in all lanes) are returned. The number of surrounding vehicles was denoted by i, ranging from 1 (denoting the vehicle closest to the subject vehicle) to i_{max} (denoting the vehicle farthest to the subject vehicle within 200 m);
3) Check necessity of identification. First, check the existence of preceding vehicles. If there are surrounding vehicles that are ahead of the j^{th} subject vehicle within 50m, the nearest one would be denoted the preceding vehicle, and the ID of preceding vehicle would be recorded; if none, stop the PVIS simulation for the j^{th} subject vehicle, jump to the $(j+1)^{th}$ subject vehicle. Second, check whether the preceding vehicle has already been identified in a previous time step. If the current preceding vehicle's ID is the same with a preceding vehicle which has been identified before, the j^{th} vehicle would have no need to search for the preceding vehicle again, unless the preceding vehicle changes (i.e., the current preceding vehicle has a different ID from the previous preceding vehicle). In real world case, the change of preceding vehicle can be detected by radar;
4) Generate sensor errors and packet drops. To simulate the harshest case, it's assumed that all the surrounding vehicles are un-platooned and thus all the surrounding vehicles are target vehicle that need to be examined. The processes in dashed frames are only for GPS-based PVIS in which packet drop is considered. The parameter t_i represent the time consumed by subject vehicle i for current identification. Only when the preparation phase (i.e. 1.0s) is finished and packet is not dropped, would the location of the target vehicle be used to generate its GPS measurement. Radar measurement of preceding vehicle, UWB and/or GPS measurements of other surrounding vehicles were generated by adding random sensor errors to the actual distance or location difference between vehicles. UWB and radar errors were assumed to comply with normal distributions, and GPS error was generated by model (2). The antennas of GPS and UWB are assumed to be at the geometric center of vehicles;
5) Determine candidate(s) for the preceding vehicle. As described in (4)∼(10), the preceding vehicle can be identified by matching radar measurement with GPS measurement, with UWB measurement, or with both of GPS and UWB measurements, depending on the type of PVIS being evaluated. If the difference between the measurement of preceding vehicle and measurement of i^{th} surrounding vehicle was smaller than the threshold, i^{th} vehicle is considered as the preceding vehicle. A theoretical failure rate of 10^{-8} was used in calculating the dynamic thresholds in equations (4∼9). Every surrounding vehicle was checked in this step, and the total number of candidate(s) was counted as N.
6) Determine the identification result. In a n-step PVIS (n can be 1), if one surrounding vehicle has been the only candidate for $n \times 0.1$ seconds, the subject vehicle identifies it as the preceding vehicle. A Results Matrix was defined: T (9000 rows, 3000 columns). When $n = 1$ in step 3, each element T(t, m) means the result of one-step identification for the j^{th} subject vehicle at the time t, where m is the Vehicle ID of the j^{th} subject vehicle. The result of multi-step identification can also be derived from T matrix later. As shown by its size, matrix T can record at most 3000 vehicles' identification result in $9000 \times 0.1 = 900$ seconds. The default value of T(t, j) is 0, indicating no identification needed for the j^{th} subject vehicle at the time t. When no candidate is found ($N = 0$) or multiple candidates are found ($N > 1$), T(t, m) would be changed to 2, indicating that the preceding vehicle cannot be decided at this step, and additional round of identification is needed; when only one candidate is found ($N = 1$), check whether its vehicle ID is the same with that of preceding vehicle. If yes, change T(t, m) to 1, indicating a correct identification; otherwise, change T(t, m) to -1, indicating an incorrect identification;
7) Repeat the processes 1 through 6 for every subject vehicle at every time.

Based on obtained matrix T, the performance of PVIS can be revealed. If one-step PVIS is evaluated, the performance measures are calculated as follows:

Average time consumption = (time step) × (number of attempts)/(number of correct identifications) = 0.1s× (number of "1" + number of "−1" + number of "2")/(number of "1");

Failure rate = (number of false identification)/(number of attempts) = (number of "−1")/(number of "1" + number of "−1" + number of "2");

It can be found that failure rate as a performance measure can lead to misleading conclusion, because when making same number of false identifications ("−1"), PIVS that consumes more time to make decisions (more "2") will be in favor instead of PIVS that has better identification speed (fewer "2"). Therefore, to eliminate this unfairness, failure rate was replaced by effective failure rate (EFR):

EFR = (number of false identifications)/ (number of identifications) = (number of "−1")/ (number of "1" + number of "−1");

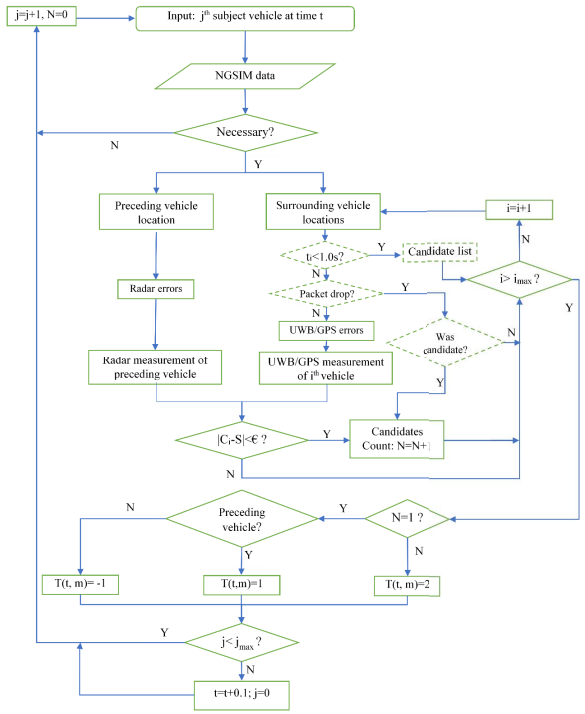

Fig. 3. Flowchart of the PVIS simulation.

Status	T ……m^th column……	T' ……m^th column……
No preceding vehicle	0	0
No preceding vehicle	0	0
One candidate once (correct)	1	2
Multiple candidate	2	2
One candidate once (incorrect)	-1	2
One candidate once (correct)	1	2
One candidate twice (correct)	1	1
No need	0	0
No need	0	0
No need	0	0

Fig. 4. Converting matrix T to T'.

Since m^{th} column of matrix T recorded the identification result for an individual vehicle m, the time consumption of every individual identification can also be calculated:

1) Find the k^{th} "1" in m^{th} column of matrix T;
2) (Time consumption of k^{th} identification for vehicle m) = 0.1+ (the total number of successive "2" above the k^{th} "1");
3) Calculate time consumption of every identification for every vehicle, and sort them from smallest to greatest, the 95th percentile and 99th percentile of time consumption can be obtained.

If a n-step PVIS is being evaluated, matrix T needs to be converted to matrix T' (an example of 2-step PVIS is shown in Fig. 4):

4) In every column of matrix T, find every n successive "1", change all but last one of them to "2";
5) In every column of matrix T, find every n successive "−1", change all but last one of them to "2";
6) Change all the other untouched "1" and "−1" to "2";

TABLE II
PERFORMANCE OF ONE-STEP PVIS WITH GPS ERROR BETWEEN 0.5m AND 6.5m

GPS error (m)	Mean (s)	95th percentile (s)	99th percentile (s)	EFR
0.5	1.2	2	2.7	0
1	2.7(1.8)	9(5.0)	16.5(9.1)	0(0)
2	22.8(1.9)	75.2(5.3)	146.1(10.0)	0(0.49)
3	100+(-)	100+(-)	100+(-)	0(-)
4	100+(-)	100+(-)	100+(-)	0(-)
5	100+(-)	100+(-)	100+(-)	0(-)
6.5	100+(-)	100+(-)	100+(-)	0(-)

Note: the numbers in brackets are the results when higher theoretical failure rate is set to achieve a 99th percentile of time consumption less than 10s, and (-) means unachievable.

With matrix T', one can go over the same procedure as for one-step PVIS, to calculate failure rate, average time consumption, 95th percentile and 99th percentile of time consumption.

B. Evaluation Results

GPS-based PVIS, UWB-based PVIS and integrated PVIS are evaluated in terms of failure rate and time consumption measures.

1) GPS-Based PVIS: Performance of one-step PVIS is first investigated. The GPS error STD in both direction (σ_x and σ_y) were assumed to be equal and chosen from 0.5m to 6.5m which basically covers the possible relative positioning errors of low-cost GPS in different scenarios. Assuming the irreducible error $\varepsilon = Min(\sigma_x) = 0.5m$, the corresponding STD of the multipath bias in (2) in both directions can be set:

$$\sigma_{b,x} = \sqrt{\sigma_x^2 - 0.5^2}$$
$$\sigma_{b,y} = \sqrt{\sigma_y^2 - 0.5^2}$$

The duration of each bias is assumed to be 0∼30s.

The theoretical failure rate is set 10^{-8}, and the packet loss rate is set 10%. The simulation results including the mean, 95^{th} percentile, 99^{th} percentile of time consumption, and effective failure rate (EFR) are listed in table II.

It can be seen that the identifications can be quickly and accurately done by the PVIS when GPS error is 0.5m, while the time consumptions went unacceptably high when the GPS error $\geqslant$ 2m. In these cases, only less than 50% subject vehicles identified its preceding vehicle before the preceding vehicle left away. It should be noted that the EFR remains 0 no matter how large the GPS error is, which means the dynamic threshold described in the section *Mathematical Models* can avoid incorrect identification even when the GPS is of poor accuracy.

On the other hand, the numbers in brackets show that the rapidity of PVIS can be improved by setting lower thresholds (i.e., higher theoretical failure rates) when GPS error = 2m,

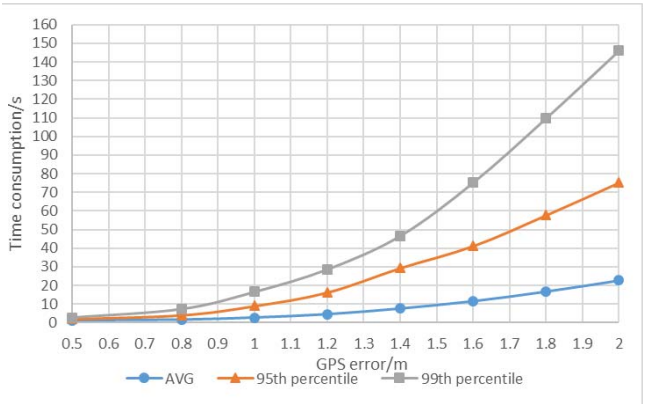

Fig. 5. One-step identification time consumption with GPS error 0.5∼2m.

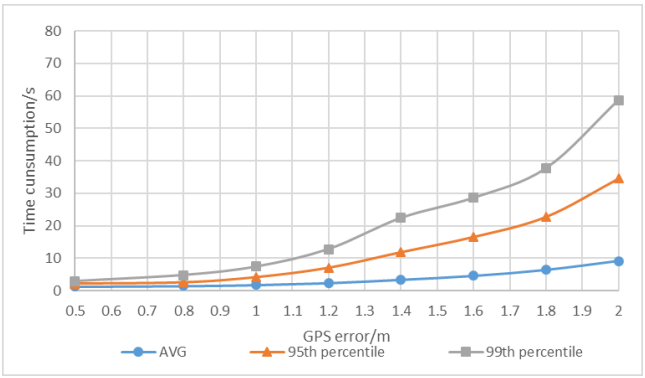

Fig. 6. Two-step identification time consumption with GPS error 0.5∼2m.

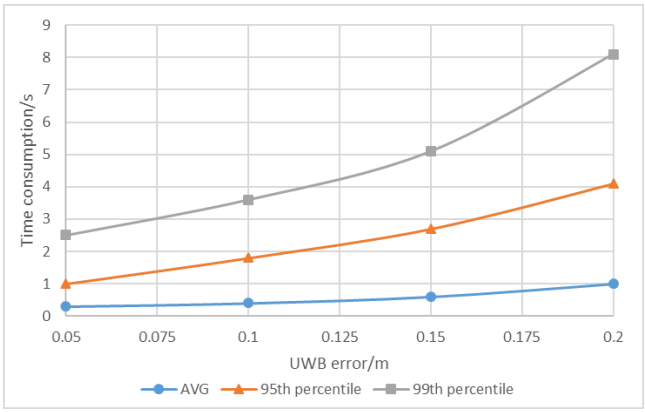

Fig. 7. One-step identification time consumption with UWB ranging error 0.05∼0.2m.

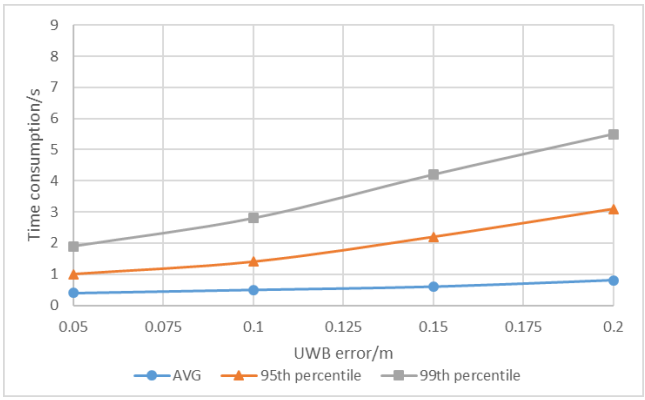

Fig. 8. Two-step identification time consumption with UWB ranging error 0.05∼0.2m.

but the identification accuracy is greatly impaired. In summary, the GPS-based PVIS is not feasible when the GPS error $\geqslant$ 2m.

Since it has been shown that GPS-based PVIS works well when GPS error $=$ 0.5m and cannot work when GPS error $\geqslant$ 2m, its performance with GPS error between 0.5m and 2m were further investigated. The one-step and two-step PVIS were both evaluated, and results are shown in Fig. 5 and Fig. 6. It can be seen that the two-step configuration can greatly improve the performance: time consumption of two-step PVIS is about half that of one-step PVIS regardless of GPS error. However, the time consumption of both PVISs can vary greatly even with a small change in GPS error. In this sense, although the two-step PVIS performs much better than one-step PVIS, it cannot proportionally lower down the requirement of GPS accuracy to achieve a certain acceptable performance. For example, if users required the waiting time to be less than 10s in 99% cases, then it can be found that the required GPS accuracy for one-step PVIS is 0.85m, while that of two-step PVIS is 1.1m, only 28% greater.

In addition, when the packet loss is set 0, the required GPS accuracy for one-step and two-step PVIS are found to be 0.9m and 1.15m respectively, which indicates that the 10% packet loss rate has not significantly impaired the feasibility of PVIS, because the GPS accuracy does not need to be improved much compared with no-packet-drop case.

We also noticed a disadvantage of two-step PVIS that the false identifications happened when GPS error was large enough. While the EFR remained 0 when GPS error $\leq$1.6m, it became 0.003 and 0.005 when GPS error was 1.8m and 2m respectively. This is resulted from the multipath bias in GPS positioning which can make the position error continuously higher than expected. In detail, when a high bias error is present, the actual preceding vehicle has $\sqrt{\alpha}$ chance to be missed by PVIS for the first-round search but has a probability larger than $\sqrt{\alpha}$ to be missed again in the second-round search (thus the total failure rate $>\alpha$) because the bias will last for a while. However, this issue won't happen to one-step PVIS because each search is independent and the failure rate α is strictly kept for each search. Thus, in terms of robustness, one-step configuration is even better than two-step configuration for GPS-based PVIS.

2) UWB-Based PVIS: The UWB-based PVIS with a ranging accuracy between 0.05m to 0.2m, covering the possible error of commercialized UWB modules under different conditions, was evaluated. The results are shown in Fig. 7 and Fig. 8. EFR was 0 for both one-step and two-step PVIS because UWB is highly-resistant to multipath effect thus bias error was assumed insignificant. Similar to GPS-based PVIS, the two-step verification greatly improved the performance of UWB-based PVIS.

In general, UWB-based PVIS has better performances than GPS-based PVIS. Even with the ranging error of 0.2m, the average time consumption of UWB-based PVIS was

TABLE III
IMPACT OF k ON TIME CONSUMPTION PERFORMANCE

k	AVG (s)	95th percentile (s)	99th percentile (s)	EFR
0.0001	0.28	0.7	1.6	0
0.001	0.28	0.7	1.5	0
0.01	0.27	0.7	1.4	0
0.1	0.27	0.7	1.4	0
0.3	0.26	0.6	1.3	0
0.5	0.26	0.6	1.3	0
0.7	0.26	0.6	1.4	0
0.9	0.26	0.6	1.4	0
0.99	0.27	0.7	1.4	0
0.999	0.28	0.7	1.5	0
0.9999	0.28	0.7	1.5	0

below 1s, and 99th percentile of time consumption was 5.5s for two-step identification. Only when GPS error is less than 0.9m can this performance be realized by GPS-based PVIS. Moreover, 0.9m positioning accuracy of GPS is more difficult to achieve (especially in dense urban environment) than 0.2m ranging error of UWB, as shown in the section *Methodology - Sensors*.

3) Integrated PVIS: With a typical GPS error of 2m and UWB error of 0.1m, a two-step integrated PVIS was evaluated. EFR remained 0 for all the runs. A unique factor in the integrated PVIS model is k, representing the reliance on UWB over GPS. The performance of the PVIS with different k is shown in Table III. It can be seen that k did not make much difference on the performance until it was very close to 0 or 1. The measures of time consumption actually remained exactly the same when k varied between 0.3 and 0.7. In this range of k, the integrated PVIS maintained the best performance. The average time consumption was 0.26s, and with 99% confidence the preceding vehicle could be identified within 1.3s, which is much lower that of UWB-based PVIS (2.8s when ranging error is 0.1m). Although a GPS-based PVIS with 2m positioning accuracy did not work well itself according to Table II, it can help greatly reduce the time consumption when integrated to a UWB-based PVIS. On the other hand, it has been known that large GPS error can lead to incorrect identification in two-step PVIS, so a higher k (i.e., to reduce the reliance on GPS) will improve the robustness of PVIS when a larger GPS error is present.

Since the results showed promising performance of integrated PVIS, it was natural to explore its tolerance to lower-accuracy sensors, which may lead to lower cost of the vehicle. The 99th percentile time consumption of integrated PVIS with different GPS and UWB errors are shown in Fig. 9. To guarantee a failure rate of zero, k was set 0.999 when the GPS error was greater than 2m, and k = 0.5 otherwise. It was found that the integration of UWB and GPS can greatly lower down the requirement for their accuracy. For example, when a 4m-accuracy GPS is integrated, the accuracy of UWB can be relaxed to as large as 0.4m to guarantee 99th percentile of time consumption <10s; when a 3m-accuracy GPS is integrated, the accuracy of UWB can be relaxed to 0.5m.

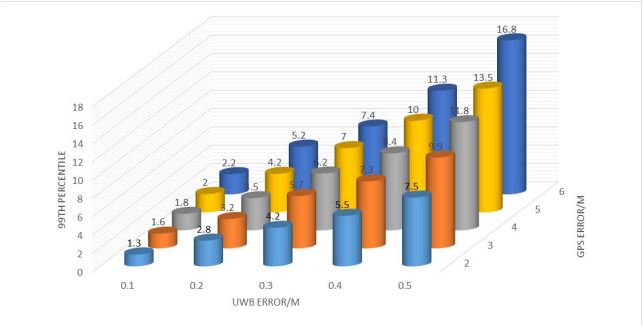

Fig. 9. 99th percentile time consumption of integrated PVIS with different sensor errors.

V. CONCLUSIONS AND FUTURE RESEARCH

In this paper, a procedure of identifying the preceding vehicle to form CACC under fully connected vehicle environment is proposed. Three preceding vehicle identification systems (PVISs), namely, GPS-based PVIS, UWB-based PVIS, and integrated PVIS combining location and distance measurements are developed with mathematical models. These models explicitly consider sensor errors, vehicle geometry, acceptable failure rate, and number of steps the PVIS takes to make decisions. A dynamic threshold for determining preceding vehicle is designed for each PVIS to pursue a certain theoretical failure rate. For integrated PVIS, an additional weight factor k is proposed to describe the PVIS's reliance on UWB over GPS.

To evaluate the performance of proposed PVISs, simulations are conducted based on real vehicle trajectory data from NGSIM, which can reflect how vehicles' relative positions change in a high-density, near-congested segment of highway. In this challenging environment, the feasibility and potential of the three PVISs with different configurations are evaluated. The failure rate, average time consumption, 95th and 99th percentiles of time consumption of preceding vehicle identification serve as performance measures. Followings are some key findings from the simulations conducted in this research:

1) GPS-based PVIS did not work well when GPS error ≥2m. To ensure an acceptable 99th percentile of identification time shorter than 10s without incorrect identification under packet loss rate of 10%, the required GPS relative positioning accuracy was 0.85m for one-step PVIS and 1.1m for two-step PVIS. The required accuracies could be slightly relaxed to 0.9m and 1.15m when there was no packet loss. These accuracies are realizable only in highway scenarios under moderate multipath effect, thus it's challenging for low-cost GPS-based PVIS to achieve an acceptable performance in urban environment;
2) UWB-based PVIS had better performances than GPS-based PVIS. Even with the ranging error of 0.2m, the average identification time of UWB-based PVIS was below 1s, and 99th percentile of time consumption

was below 6s. Because this ranging accuracy of UWB has been proven achievable in multipath environment, UWB-based PVIS is more suitable for urban scenario;

3) Under the theoretical failure rate of 10^{-8}, no incorrect identification was observed in all simulations except for two-steps GPS-based PVIS which is sensitive to the multipath bias in GPS positioning. In other word, one-step GPS-based PVIS is more robust than the two-step one, although it has worse performance in time consumptions. It's also observed that the incorrect identification can be eliminated when GPS error is smaller than 1.6m in two-step GPS-based PVIS.
4) The integrated PVIS of UWB and GPS showed promising performance and the potential to lower down the requirement for sensor accuracies. The shortest time consumption of the integrated PVIS could be realized as long as the weight factor $k = 0.3 \sim 0.7$, but higher k is needed for the robustness of PVIS under large GPS bias.

It should also be noted that all the findings are subject to the following limitations of the study:

1) The NGSIM data used for performance evaluation were collected from high-density urban highway during rush hour, thus have led to more conservative results. The obtained requirements for sensor accuracy might be unnecessarily high in lower-density cases such as rural highway or uncongested urban roadway;
2) The performance evaluations are done based on an assumption that each subject vehicle knows the exact error STD of their sensors. Specifically, the error STDs in the performance evaluation were the same with the error STDs used in the dynamic threshold (4)∼(10). In real world, sensors' performance might change with the operating condition, but this change may not be perceived or accurately estimated by the vehicle itself. Over-estimation of error might impair the efficiency of PVIS, while under-estimation of error might increase the probability of false identification.

Immediate future research is to enhance the proposed PVIS for mix traffic conditions instead of 100% connected vehicle environment. The preceding vehicle must be determined more carefully due to the existence of unconnected vehicle. Imagine a scenario that the preceding vehicle in front of the subject vehicle is not connected, while an irrelevant vehicle beside the preceding vehicle is spreading its message. Since current PVIS assumes preceding vehicle's connectivity and uses a relatively "loose" threshold to make sure not to miss the preceding vehicle, there would be a chance that the irrelevant vehicle meets the threshold and becomes the only candidate for preceding vehicle due to the absence of message from the actual preceding vehicle. This concern could be addressed by assuring the requirement for PVIS that it must screen out all irrelevant vehicles (instead of "as many irrelevant vehicles as possible" stated in section Methodology) within the identification procedure, when working in the mix traffic.

Other possible future research ideas include:

1) Investigating the proposed PVIS's performance in uncongested highways and urban roads;
2) Developing a prototype integrated PVIS device with real radar, GPS and UWB sensors, and test its performance in real traffic;
3) Considering other cooperative automated driving applications such as cooperative merging [45].

REFERENCES

[1] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner, "Three decades of driver assistance systems: Review and future perspectives," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 4, pp. 6–22, Oct. 2014.

[2] H. Liu, X. Kan, S. E. Shladover, X. Y. Lu, and R. E. Ferlis, "Modeling impacts of cooperative adaptive cruise control on mixed traffic flow in multi-lane freeway facilities," *Transp. Res. C, Emerg. Technol.*, vol. 95, pp. 261–279, Oct. 2018.

[3] J. Lioris, R. Pedarsani, F. Y. Tascikaraoglu, and P. Varaiya, "Platoons of connected vehicles can double throughput in urban roads," *Transp. Res. C, Emerg. Technol.*, vol. 77, pp. 292–305, Apr. 2017.

[4] E. Talavera, A. Díaz-Álvarez, F. Jiménez, and J. Naranjo, "Impact on congestion and fuel consumption of a cooperative adaptive cruise control system with lane-level position estimation," *Energies*, vol. 11, no. 1, p. 194, 2018.

[5] S. Shladover, X.-Y. Lu, and S. Yang, "Cooperative adaptive cruise control (CACC) for partially automated truck platooning," UC Berkeley, Berkeley, CA, USA, Final Rep., 2018. [Online]. Available: https://escholarship.org/uc/item/260060w4

[6] K. C. Dey *et al.*, "A review of communication, driver characteristics, and controls aspects of cooperative adaptive cruise control (CACC)," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 491–509, Feb. 2016.

[7] R. Rajamani and S. E. Shladover, "An experimental comparative study of autonomous and co-operative vehicle-follower control systems," *Transp. Res. C, Emerg. Technol.*, vol. 9, no. 1, pp. 15–31, 2001.

[8] J. Ploeg, B. T. M. Scheepers, E. van Nunen, N. van de Wouw, and H. Nijmeijer, "Design and experimental evaluation of cooperative adaptive cruise control," in *Proc. Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2011, pp. 260–265.

[9] G. J. L. Naus, R. P. A. Vugts, J. Ploeg, M. J. G. van de Molengraft, and M. Steinbuch, "String-stable CACC design and experimental validation: A frequency-domain approach," *IEEE Trans. Veh. Technol.*, vol. 59, no. 9, pp. 4268–4279, Nov. 2010.

[10] V. Milanés, S. E. Shladover, and J. Spring, "Cooperative adaptive cruise control in real traffic situations," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 1, pp. 296–305, Feb. 2014.

[11] E. V. Nunen, M. R. J. A. E. Kwakkernaat, J. Ploeg, and B. D. Netten, "Cooperative competition for future mobility," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1018–1025, Sep. 2012.

[12] L. Guvenc *et al.*, "Cooperative adaptive cruise control implementation of team Mekar at the grand cooperative driving challenge," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1062–1074, Sep. 2012.

[13] J. Ploeg, S. Shladover, H. Nijmeijer, and N. van de Wouw, "Introduction to the special issue on the 2011 grand cooperative driving challenge," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 989–993, Sep. 2012.

[14] J. Ploeg, E. Semsar-Kazerooni, G. Lijster, N. van de Wouw, and H. Nijmeijer, "Graceful degradation of CACC performance subject to unreliable wireless communication," in *Proc. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2013, pp. 1210–1216.

[15] A. Geiger *et al.*, "Team AnnieWAY's entry to the 2011 grand cooperative driving challenge," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 3, pp. 1008–1017, Sep. 2012.

[16] J. Levinson *et al.*, "Towards fully autonomous driving: Systems and algorithms," in *Proc. IEEE Intell. Vehicles Symp. IV*, Jun. 2011, pp. 163–168.

[17] N. Kalra and S. M. Paddock, "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability," *Transp. Res. A, Policy Pract.*, vol. 94, pp. 182–193, Dec. 2016.

[18] J. B. Kenney, "Dedicated short-range communications (DSRC) standards in the United States," *Proc. IEEE*, vol. 99, no. 7, pp. 1162–1182, Jul. 2011.

[19] R. J. Fontana, "Recent system applications of short-pulse ultra-wideband (UWB) technology," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 9, pp. 2087–2104, Sep. 2004.

[20] X. Cui, T. A. Gulliver, J. Li, and H. Zhang, "Vehicle positioning using 5G millimeter-wave systems," *IEEE Access*, vol. 4, pp. 6964–6973, 2016.
[21] D. Kumar, J. Saloranta, G. Destino, and A. Tölli, "On trade-off between 5G positioning and mmWave communication in a multi-user scenario," in *Proc. 8th Int. Conf. Seamless Indoor-Outdoor Localization (ICL-GNSS)*, Jun. 2018, pp. 1–5.
[22] H. Wymeersch, G. Seco-granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmWave positioning for vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, Dec. 2017.
[23] V. Alexiadis, J. Colyar, J. Halkias, R. Hranac, and G. McHale, "The next generation simulation program," *Inst. Transp. Eng. J.*, vol. 74, no. 8, pp. 22–26, 2004.
[24] J. Ma, F. Zhou, and M. J. Demetsky, "Evaluating mobility and sustainability benefits of cooperative adaptive cruise control using agent-based modeling approach," in *Proc. IEEE Syst. Inf. Design Symp. (SIEDS)*, Apr. 2012, pp. 74–78.
[25] J. Hatch, A. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band," *IEEE Trans. Microw. Theory Techn.*, vol. 60, no. 3, pp. 845–860, Mar. 2012.
[26] J. G. Zato *et al.*, "Limitations of positioning systems for developing digital maps and locating vehicles according to the specifications of future driver assistance systems," *IET Intell. Transp. Syst.*, vol. 5, no. 1, pp. 60–69, 2011.
[27] G. J. Morgan-Owen and G. T. Johnston, "Differential GPS positioning," *Electron. Commun. Eng. J.*, vol. 7, no. 1, pp. 11–21, 1995.
[28] L. S. Monteiro, T. Moore, and C. Hill, "What is the accuracy of DGPS," *J. Navigat.*, vol. 58, no. 2, pp. 207–225, 2005.
[29] N. Alam, A. Kealy, and A. G. Dempster, "An INS-aided tight integration approach for relative positioning enhancement in VANETs," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1992–1996, Dec. 2013.
[30] F. de Ponte Müller, E. M. Diaz, B. Kloiber, and T. Strang, "Bayesian cooperative relative vehicle positioning using pseudorange differences," in *Proc. Rec. IEEE PLANS, Position Locat. Navig. Symp.*, May 2014, pp. 434–444.
[31] K. Liu, H. B. Lim, E. Frazzoli, H. Ji, and V. C. S. Lee, "Improving positioning accuracy using GPS pseudorange measurements for cooperative vehicular localization," *IEEE Trans. Veh. Technol.*, vol. 63, no. 6, pp. 2544–2556, Jul. 2014.
[32] M. Matosevic, Z. Salcic, and S. Berber, "A comparison of accuracy using a GPS and a low-cost DGPS," *IEEE Trans. Instrum. Meas.*, vol. 55, no. 5, pp. 1677–1683, Oct. 2006.
[33] A. Giremus, J. Tourneret, and V. Calmettes, "A particle filtering approach for joint detection/estimation of multipath effects on GPS measurements," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1275–1285, Sep. 2007.
[34] R. Hult *et al.*, "Design and experimental validation of a cooperative driving control architecture for the grand cooperative driving challenge 2016," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1290–1301, Apr. 2018.
[35] A. Alarifi *et al.*, "Ultra wideband indoor positioning technologies: Analysis and recent advances," *Sensors*, vol. 16, no. 5, pp. 1–36, 2016.
[36] L. Cao, Z. Chen, L. Yan, Q. Qin, and R. Zhang, "A proposed vision and vehicle-to-infrastructure communication-based vehicle positioning approach," *J. Intell. Transp. Syst. Technol. Planning, Oper.*, vol. 21, no. 2, pp. 123–135, 2017.
[37] T. Ye, M. Walsh, P. Haigh, J. Barton, A. Mathewson, and B. O. Flynn, "Experimental impulse radio IEEE 802.15.4a UWB based wireless sensor localization technology: Characterization, reliability and ranging," in *Proc. 22nd IET Irish Signals Syst. Conf.* Jun. 2011, pp. 1–7.
[38] V. Kristem, S. Niranjayan, S. Sangodoyin, and A. F. Molisch, "Experimental determination of UWB ranging errors in an outdoor environment," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 4838–4843.
[39] A. R. J. Ruiz and F. S. Granja, "Comparing ubisense, bespoon, and decawave UWB location systems: Indoor performance analysis," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 8, pp. 2106–2117, Aug. 2017.
[40] M. I. Hassan, H. L. Vu, and T. Sakurai, "Performance analysis of the IEEE 802.11 MAC protocol for DSRC safety applications," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3882–3896, Oct. 2011.
[41] F. Bai and H. Krishnan, "Reliability analysis of DSRC wireless communication for vehicle safety applications," in *Proc. IEEE Intell. Transp. Syst. Conf.*, Sep. 2006, pp. 355–362.
[42] *Standard Specification for Telecommunications and Information Exchange between Roadside and Vehicle Systems-5.9 GHz Band Dedicated Short Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, Standard ASTM E.2213-03, ASTM International, 2003, pp. 17–18, vol. 4.
[43] R. Reinhold and R. Kays, "Improvement of IEEE 802.15.4a IR-UWB for time-critical industrial wireless sensor networks," in *Proc. IFIP Wireless Days*, Nov. 2013, pp. 1–4.
[44] M.-G. Di Benedetto, L. De Nardis, M. Junk, and G. Giancola, "(UWB)2: Uncoordinated, wireless, baseborn medium access for UWB communication networks," *Mobile Netw. Appl.*, vol. 10, no. 5, pp. 663–674, 2005.
[45] V. Dolk *et al.*, "Cooperative automated driving for various traffic scenarios: Experimental validation in the GCDC 2016," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1308–1321, Apr. 2018.

Zheng Chen received the B.S. and M.S. degrees from Hunan University, China, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Department of Engineering Systems and Environment, University of Virginia, VA, USA. His research interests are dynamics and control of connected vehicle systems.

Byungkyu Brian Park (M'05–SM'18) is currently an Associate Professor with the Engineering Systems and Environment Department, University of Virginia. He has published over 150 journal and conference papers in the areas of transportation system operations and managements, and intelligent transportation systems. His research interests include cyber-physical system for transportation, stochastic optimization, connected and automated vehicle safety assessment, microscopic simulation model application, and transportation system sustainability.