Connected Preceding Vehicle Identification for Enabling Cooperative Automated Driving in Mixed Traffic

Zheng Chen[1] and B. Brian Park, M.ASCE[2]

Abstract: To enable the safe and fast formation of connected automated vehicle (CAV) platoons in real-world traffic, a preceding vehicle identification system for mixed traffic (PVIS-mixed) is proposed. PVIS-mixed utilizes the vehicle's radar measurements and global positioning system (GPS) measurements reported by surrounding connected vehicles to find the communication identity of the preceding vehicle. The design of PVIS-mixed is based on three goals: a low probability of making a wrong identification, a low probability of missing the connected preceding vehicle, and short time consumption of the identification procedure. The proposed PVIS-mixed is evaluated in highway traffic simulated by real vehicle trajectory data from the Next Generation Simulation (NGSIM) program. Evaluation results showed that the performance of PVIS-mixed is not related to the adoption rate of connected vehicles, and 1 m is found to be the required relative positioning accuracy to make 99th percentile time consumption <10 s. It was observed that the multipath bias of GPS positioning could affect the usability of CAV platooning. The possible solutions are then discussed as future work. **DOI: 10.1061/JTEPBS.0000661.** © *2022 American Society of Civil Engineers.*

Introduction

Improving safety and mobility are the two key objectives for future transportation systems. In 2018, more than 1.35 million people worldwide died from traffic crashes, which has become the leading cause of deaths of people aged 5–29 years (WHO 2018). Vehicle automation has been considered the most effective way to prevent crashes because most roadway crashes are associated with drivers' improper behaviors (Singh 2015). While fully automated vehicles (i.e., driverless vehicles) are still far from large-scale implementation due to technical difficulty and high cost, some partially automated vehicle applications, such as adaptive cruise control (ACC), have become commercially available (Bengler et al. 2014). Furthermore, significant enhancement in transportation efficiency is expected to be achieved by introducing connectivity to automated vehicles, such as vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications, which would enable automated vehicles to travel in cooperative ways.

Cooperative adaptive cruise control (CACC), enhanced from ACC, is one the most promising and prepared applications of connected automated vehicles (CAVs). It enables vehicles to stably travel as compact platoons in short headways like 0.6 s, and makes it possible to double or triple the current roadway capacity and essentially resolve congestion (Lioris et al. 2017; Shladover et al. 2012). Apart from the state of the preceding vehicle measured by range sensor, CACC vehicles also make use of the acceleration or intended acceleration of the preceding vehicle or even information from vehicles further ahead via communications (Milanés et al. 2013; Naus et al. 2010; Ploeg et al. 2011).

[1]Research Engineer, PCI Technology Group Co., Ltd., No. 2 Xincen Fourth Rd., Guangzhou 510630, China (corresponding author). ORCID: https://orcid.org/0000-0002-2344-5892. Email: zc4ac@virginia.edu

[2]Professor, Link Lab and Dept. of Engineering Systems and Environment, Univ. of Virginia, 151 Engineers Way, Charlottesville, VA 22904. ORCID: https://orcid.org/0000-0003-4597-6368. Email: bp6v@virginia.edu

Note. This manuscript was submitted on June 13, 2021; approved on December 21, 2021; published online on February 22, 2022. Discussion period open until July 22, 2022; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Transportation Engineering, Part A: Systems***, © ASCE, ISSN 2473-2907.

A variety of CACC systems have been developed in the last two decades. Based on the communication topology, existing CACC systems can be divided into three categories: predecessor-following (PF) CACC (Naus et al. 2010), which only communicates with the nearest preceding vehicle; predecessor-leader-following (PLF) CACC (Milanés et al. 2013), which communicates with both the preceding vehicle and platoon leader; and multi-predecessor-following (MPF) CACC (Levine and Athans 1966), which requires communications with multiple preceding vehicles in the platoon. All these CACC systems require the message from the nearest preceding vehicle, which directly affects the safety of the subject vehicle. Actually, when the subject vehicle detects another vehicle ahead by onboard sensors and tries to platoon with this preceding vehicle, the first thing it should do is to establish connection with this preceding vehicle [if it is a connected vehicle (CV)]. Once done, the communication IDs of other members in the platoon can be obtained from the preceding vehicle and connections with them can be further established. Therefore, identification of the preceding vehicle is a necessary step to form or join a CACC platoon.

Ideally, the subject vehicle can request all the surrounding connected vehicles to share their positions estimated by global positioning system (GPS), and then compare those self-reported positions with the sensor measurement on the actual preceding vehicle. The vehicle whose self-reported position matches that of the subject vehicle's sensor measurement should be considered as the preceding vehicle. However, this is not a trivial task on real-world multilane highways where the reported GPS positions may not be accurate enough to help distinguish the preceding vehicle from other nearby vehicles. Because the feedforward signal from the preceding vehicle plays a key role in CACC motion control (Al-Jhayyish and Schmidt 2017), the misidentification of the preceding vehicle could result in unexpected or undesirable car-following behaviors of the ego vehicle, undermining the ride quality and even threatening the safety (or activating collision avoidance maneuvers) in extreme situations.

The issue of preceding vehicle identification has neither been fully identified nor addressed in previous CACC demonstrations. In Milanés et al. (2013), Naus et al. (2010), and Ploeg et al. (2011), all the CACC vehicles were traveling in a single platoon so that they could be easily distinguished from each other. In the Grand Cooperative Driving Challenge (GCDC) 2011, two CACC platoons were competing on two adjacent lanes, but the participants were allowed to preset a blacklist in the vehicle's software to block the messages from vehicles in the adjacent lane (Geiger et al. 2012). This blacklist is obviously impossible in the real world. Most of vehicles in the GCDC were already equipped with high-performance GPS, but still had insufficient confidence in correctly identifying the preceding vehicle. A key question awaiting answer is that how accurate the GPS or sensors should be to guarantee correct and quick identification of the preceding vehicle.

A preceding vehicle identification system (PVIS) under 100% connected vehicle environment was proposed and evaluated in Chen and Park (2019). This PVIS calculated a dynamic searching area based on GPS or sensor errors, vehicle geometry, and radar measurement of the actual preceding vehicle. The searching area is determined so that the reported GPS position of the actual preceding vehicle tends to always be within it, while other irrelevant vehicles will be gradually screened out over time. However, this PVIS is not able to work in mixed traffic consisting of both connected and unconnected vehicles. Without a proper mechanism to decide whether the preceding vehicle is a connected vehicle or not, PVIS would end up misidentifying an irrelevant vehicle as the preceding vehicle if the preceding vehicle is actually unconnected. Because it is predicted that the adoption rate of vehicle connectivity will not reach 100% until 2040s in the United States (Bansal and Kockelman 2017), a more sophisticated PVIS that can cope with unconnected vehicles is needed for the implementation of CACC in the near future. To fill the research gap, this paper proposes a new preceding vehicle identification system for mixed traffic (PVIS-mixed).

The rest of the paper is organized as follows: The archetecture of PVIS-mixed is proposed in the second section. The third section describes the identification procedure in mixed traffic conditions. The fourth section presents the design of PVIS-mixed, including the design goals, derivation of the searching area, and optimization of the parameters of PVIS-mixed. In the fifth section, the performance of the proposed PVIS-mixed is evaluated with real-world vehicle trajectory data. The sixth section summarizes findings from this research and discusses the potential improvements for future research.

Architecture

The main assumptions in this study are:

- Imperfect adoption rate of connected vehicle;
- Connected vehicles are equipped with GPS;
- The user of PVIS-mixed (i.e., subject vehicle) is capable of CACC and is equipped with long-range radar sensor; and
- The effect of packet loss on PVIS-mixed is insignificant and can be easily handled according to the previous results in Chen and Park (2019).

With the preceding assumptions, the architecture of the proposed PVIS-mixed can be illustrated in Fig. 1. The core of PVIS-mixed is the identification procedure, a computer program to determine the communication ID of the preceding vehicle (if it exists). This procedure is structured with an outer loop and an inner loop (to be explained in the next section), which recursively search for connected vehicles whose locations are close enough to that of the preceding vehicle.

When the identification procedure runs in real time, PVIS-mixed acquires the data from the radar, V2V communication device, and GPS of the subject vehicle. These data indicate the relative location of the actual preceding vehicle and those of the surrounding connected vehicles.

To guarantee that the PVIS-mixed works correctly and efficiently, the key parameters in the identification procedure need to be optimized in advance. The offline parameterization can compute a table of the optimal parameters according to the design goals and the system models.

Finally, if PVIS-mixed determines that the preceding vehicle is a connected vehicle and finds its communication ID, then the subject vehicle is ready to activate CACC, or other CAV applications demanding this information.

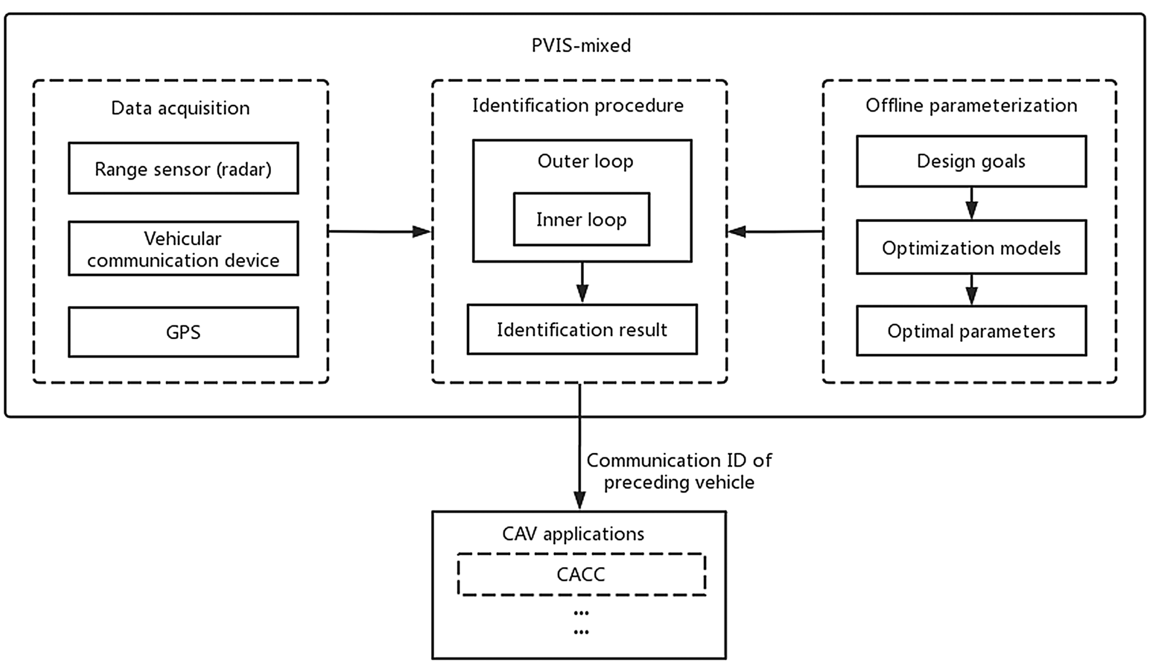

Fig. 1. Architecture of PVIS-mixed.

Identification Procedure

An intuitive way for PVIS-mixed to deal with an unconnected vehicle in mixed traffic is to determine the preceding vehicle as unconnected if all the GPS locations (with errors) reported by surrounding vehicles are continuously far from the radar-measured location of the preceding vehicle. However, such criterion should be carefully designed because the tighter criterion may lead to longer time consumption and higher unusability of CACC, i.e., not activating CACC when the preceding vehicle is actually a connected vehicle, while the looser criterion may lead to frequent misidentifications. Therefore, this study explicitly models this decision-making process and designs PVIS-mixed to optimally suppress the identification time, unusability, and misidentification rates. Accordingly, the identification procedure is proposed as shown in Fig. 2 and explained as follows:

1. Start the identification procedure when the subject vehicle detects a preceding vehicle within the preset identification range (e.g., 200 m).

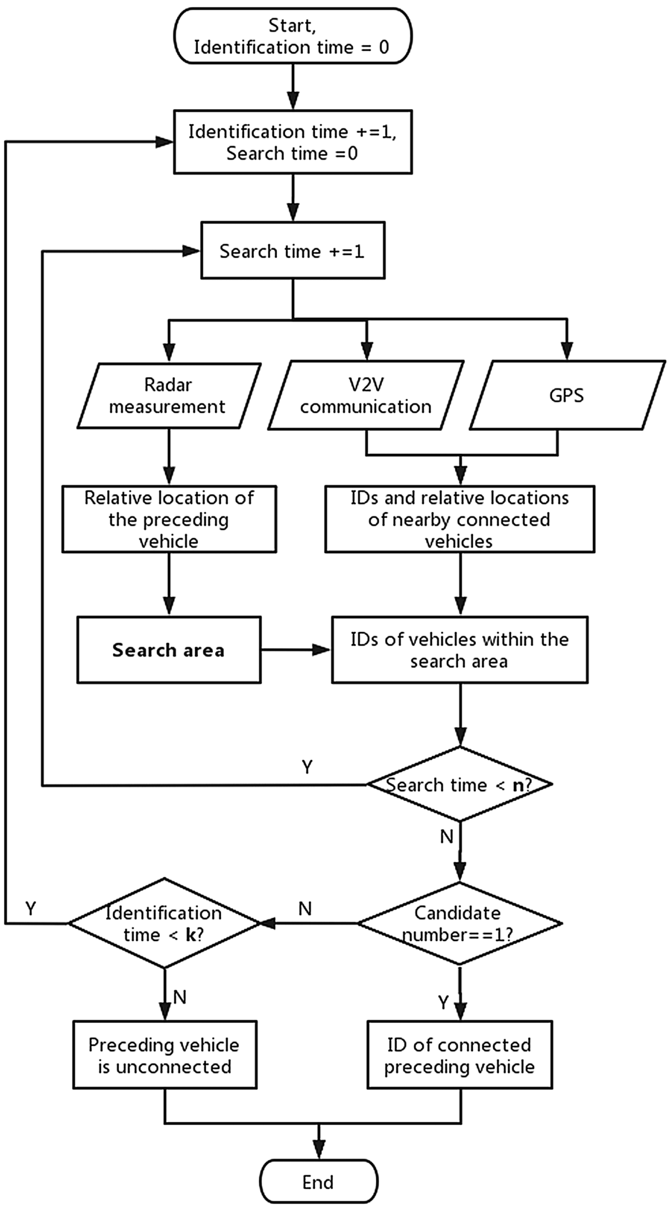

Fig. 2. Flowchart of identification procedure.

2. Measure the relative location of the preceding vehicle via radar and communicate with nearby vehicles to obtain their self-reported locations along with their communication IDs.
3. Calculate a searching area A around the radar-measured position of preceding vehicle and note the communication IDs of vehicles whose self-reported GPS locations are within A.
4. Repeat Steps 2 and 3 n consecutive times and determine the candidates for the preceding vehicle whose self-reported GPS location is within the area A in all n searches, where *search* means Steps 2–3 (i.e., the inner loop in Fig. 2). This inner loop is set up mainly for the goal of avoiding misidentification. It screens out irrelevant vehicles by utilizing the sensor measurements in multiple consecutive searches.
5. If only one candidate is found in Step 4, this candidate would then be considered as the preceding vehicle and the identification procedure ends.
6. If no candidate is found, go back to Step 1 and restart the identification, where *identification* refers to Steps 2–5 (i.e., the outer loop in Fig. 2). If the candidate is not found for k consecutive identifications, determine the preceding vehicle to be unconnected and the procedure ends. The main purpose of this outer loop is to avoid unusability when there is a connected preceding vehicle. It also limits time consumption in case the preceding vehicle is indeed unconnected.
7. If multiple candidates are found (which is unlikely), go back to Step 1 and restart the identification.

In summary, each trial of identification lasts for n time steps. If there is one and only one vehicle whose location sufficiently matches the radar-measured location of the preceding vehicle for the n consecutive time steps, then it is determined as the preceding vehicle. On the other hand, if no candidate is found in k trials of such identification, then the preceding vehicle is determined to be unconnected.

At worst, PVIS-mixed needs to consider the radar or GPS measurements during nk time steps before making the final decision, but more commonly the identification procedure is to be terminated in advance when the connected preceding vehicle is found in any of the identification trials.

Compared to the PVIS for 100% CV traffic (Chen and Park 2019), the PVIS-mixed are different mainly in three aspects:

1. Additional step. Step 6 is dedicated for PVIS-mixed. It gives a chance to stop the procedure when the preceding vehicle is unconnected. Without Step 6, PVIS (Chen and Park 2019) would keep searching until an irrelevant vehicle is misidentified as the preceding vehicle.
2. Design goals. PVIS-mixed is set to first exclude irrelevant vehicles from the candidate list after Step 4 with guaranteed probability to prevent misidentification when the actual preceding vehicle is unconnected. Second, it tries to include the actual preceding vehicle if it is connected at a guaranteed probability. Finally, the identification time consumption is taken into consideration because its maximum value is explicitly linked to the product of n and k. In contrast, the PVIS in Chen and Park (2019) only needs to guarantee the actual preceding vehicle as a candidate with certain probability because false identifications would not happen as long as the actual preceding vehicle is on the candidates list.
3. Parameterization. Because of the new parameter introduced in Step 6 and the multiple design goals, the proper searching area A, n, and k must be determined jointly. As a result, extra modeling and nonlinear optimization is required, as shown subsequently, to parameterize PVIS-mixed. As a comparison, PVIS (Chen and Park 2019) accepts arbitrary n and the searching area is simply calculated based on the desired error rate (Er) and n, i.e., $\alpha = \sqrt[n]{Er}$.

Parameter Design

Design Goals

The first goal of PVIS-mixed is to guarantee a low probability to make an error. The error is defined as the mismatch where an irrelevant vehicle is determined to be the preceding vehicle through the identification procedure. In the mixed traffic, the theoretical error rate is

$$Er = \begin{cases} (1-P_p)P_i & \text{if preceding vehicle is connected} \\ P_i & \text{if preceding vehicle is unconnected} \end{cases} \quad (1)$$

where P_p = probability that the actual preceding vehicle is identified as a candidate; and P_i = probability that an irrelevant vehicle is considered as a candidate. Because $(1-P_p)P_i \leq P_i$, it will be sufficient to design PVIS-mixed in the case that the preceding vehicle is unconnected, i.e., the first goal is to make $Er = P_i \to 0$.

Second, we need to ensure that PVIS-mixed has a small chance to miss the preceding vehicle when it is connected. The concept of unusablity rate (Ur) is defined to reflect this goal

$$Ur = 1 - P_p \quad (2)$$

Assuming an $Er \approx 0$ can be achieved, we have $Ur \approx (1-P_p)(1-E_r)$, which represents the probability that PVIS-mixed takes a connected preceding vehicle as unconnected and mistakenly gives up performing CACC. Unusablity is less serious than error because it does not cause false behavior of the vehicle. Finally, a low maximum identification time consumption is preferred for higher usability of CACC and better user experience. Because $E_r \approx 0$, Step 7 in the identification procedure actually has a rare chance to happen; thus, the maximum number of searches is nk. Assuming the update frequency of communication and GPS positioning to be 10 Hz, the maximum identification time consumption is $0.1nk$.

Searching Area

A searching area was proposed in Chen and Park (2019). To be considered as a candidate for preceding vehicle in a single search, a surrounding vehicle's GPS measurements should be close enough to the subject vehicle's radar measurements on the actual preceding vehicle

$$(x_g - x_r)^2/\delta_x^2 + (y_g - y_r)^2/\delta_y^2 < \chi^2(2, \alpha) \quad (3)$$

where y_r and x_r = radar-measured longitudinal and lateral locations of preceding vehicle; y_g and x_g = GPS-measured longitudinal and lateral locations of the surrounding vehicle; χ^2 = chi-square statistic; α = probability that the actual preceding vehicle is positioned out of the oval area defined by Eq. (3); and $\delta_x = \sqrt{(\sqrt{x_r^2 + y_r^2}\sigma_a)^2 + \sigma_x^2}$ and $\delta_y = \sqrt{\sigma_y^2 + \sigma_r^2}$, where σ_x and σ_y are the GPS standard error in longitudinal and lateral directions, respectively, and σ_a and σ_r are the standard angle error and range error of the radar sensor, respectively. With this searching area, the corresponding p_p can be computed

$$p_p = 1 - (1 - (1-\alpha)^n)^k \quad (4)$$

Error Rate Model

Because Eq. (4) has revealed the relationship between the searching area A and P_p, the next step is to relate A to the desired P_i. Obviously, P_i becomes larger when irrelevant vehicles get closer to the preceding vehicle. To guarantee Er for all cases, the most challenging situation in Fig. 3 should be considered.

In this situation, the preceding vehicle is unconnected but surrounded by connected vehicles; h is the minimum headway and w is the minimum lateral distance between vehicles. In this study, it is assumed that $h = 10$ m and $w = 2.5$ m. Around the preceding vehicle, the two nearest irrelevant vehicles are parallel to the preceding vehicle in the adjacent lanes, and two second nearest irrelevant vehicles are in the further lanes. Similarly, one can also find the third and fourth nearest irrelevant vehicles and so on. However, it is assumed that the two first nearest vehicles have dominant probability to be mismatched over other vehicles, which can be proven afterward; thus, only these two vehicles are taken into consideration when calculating p_i.

The searching area described by Eq. (3) can be adopted in real-world operation but cannot be used to determine the n and k in the design of PVIS-mixed because the area needs to change with radar measurements. For this reason, the radar error is assumed insignificant compared to GPS error. Then the boundary of area A can be reformulated with regard to the radar-measured location (0,0) of the preceding vehicle

$$x_g^2 + y_g^2 < \sigma\chi^2(2, \alpha) \quad (5)$$

where $\sigma = \sigma_x = \sigma_y$. It can be seen that area A is solely determined by parameter α when a GPS error is given.

For a surrounding irrelevant vehicle, the probability density function of its relative GPS position with respect to its real position is given in normal distribution

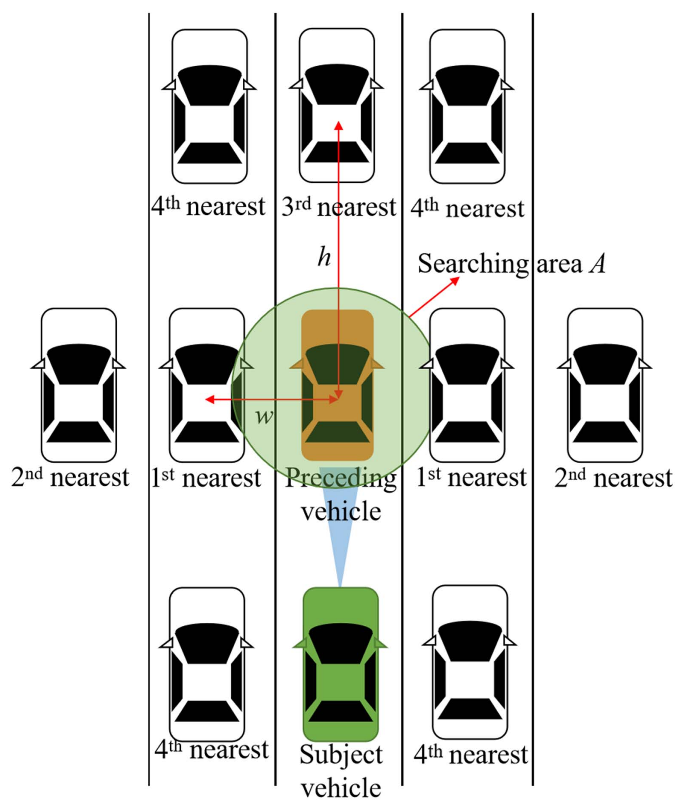

Fig. 3. Most challenging situation for PVIS-mixed.

$$f(x_g, y_g) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_g^2 + y_g^2}{2\sigma^2}\right) \tag{6}$$

And its probability to be positioned within the area A for n measurements is

$$p_n = \left(\iint_{A'} f(x, y)dxdy\right)^n \tag{7}$$

where A' = searching area in respect to the position of the irrelevant vehicle. In the view of the first nearest irrelevant vehicles, A' is the shifted A in the lateral direction by w

$$A' = \{(x, y)|(x \pm w)^2 + y^2 < \sigma\chi^2(2, \alpha)\} \tag{8}$$

At last, the probability that any of the first nearest irrelevant vehicles is determined to be the preceding vehicle, can be estimated

$$p_i \approx 2kp_n \tag{9}$$

It can be seen that p_i and p_p are both functions of α, n, and k.

Parameters Optimization

Based on Eqs. (4)–(9), an optimization problem can be set:

Minimize $J = W_a(1 - p_p(\alpha, n, k)) + 0.1W_b nk$ by changing α, n, and k

Subject to

$$p_i(\alpha, n, k) \le Er^*$$

$$0.1nk \le t_{\max}$$

$$p_p \ge P_{\min}$$

$$n, k \in N^+$$

$$0 < \alpha < 1$$

where J = cost function, a combination of weighted unusability rate (Ur) and maximum time consumption; and W_a and W_b = weights for Ur and maximum time consumption, respectively, with a higher ratio between W_a and W_b leading to higher usability of CACC but longer identification time, and vice versa; Er^* = required theoretical error rate; $t_{\max}$ = maximum acceptable time consumption; and $P_{\min}$ = minimum acceptable usability of CACC.

Because there are nonlinear and integer constraints, genetic algorithm (GA) is adopted to search for (potential) optimal parameters.

Using $W_a = 500$, $W_b = 1$, $t_{\max} = 35(s)$, and $P_{\min} = 0.95$ (users are free to choose other values according to their preferences), the optimal parameters and expected performance given $Er^* = 10^{-6} - 10^{-10}$ and different GPS error are listed in Table 1. No solution can be found when GPS standard error is greater than 1.1 m using $Er^* = 10^{-6}$ and 10^{-8}, or 1 m using $Er^* = 10^{-10}$. It can be seen that pursuing an Er closer to zero will lead to longer identification time and higher theoretical Ur.

According to n and α provided in Table 1, it can be computed that the second nearest vehicles have probability $<10^{-30}$ to be identified as the preceding vehicle, which means that they indeed have insignificant contribution to the error rate.

Table 1. Optimal parameters and expected performance of PVIS-mixed

Er^*	GPS standard error (m)	n	α	k	Maximum time consumption (s)	Theoretical, Ur (%)
10^{-6}	0.5	5	0.0026	2	1	0.02
	0.6	9	0.0031	2	1.8	0.07
	0.7	9	0.0208	4	3.6	0.09
	0.8	10	0.0506	7	7	0.18
	0.9	20	0.0184	5	10	0.29
	1	17	0.059	11	18.7	0.79
	1.1	26	0.039	10	26	1.24
10^{-8}	0.5	3	0.1254	7	2.1	0.04
	0.6	6	0.0474	5	3	0.10
	0.7	13	0.0144	4	5.2	0.09
	0.8	26	0.0062	3	7	0.34
	0.9	34	0.0089	4	13.6	0.47
	1	38	0.0155	5	22.8	0.81
	1.1	50	0.0159	7	35	1.55
10^{-10}	0.5	5	0.0211	3	1.5	0.10
	0.6	10	0.0109	3	3	0.11
	0.7	11	0.0386	6	6.6	0.19
	0.8	12	0.0822	12	14.4	0.50
	0.9	33	0.0136	5	16.5	0.63
	1	42	0.0163	6	25.2	1.54
	1.1				N/A	

Performance Evaluation

To validate the effectiveness of the proposed PVIS-mixed, the real-world vehicle trajectory data collected by the Next Generation Simulation (NGSIM) program (Alexiadis et al. 2004) were utilized to reproduce the traffic on a high-density highway segment. PVIS-mixed was then supposed to pair preceding/following vehicles under measurement errors of radar and GPS, which were generated by error models. The time consumed to make final decisions, and the resulted Er and Ur, served as the measures of effectiveness.

NGSIM Data

The NGSIM program (Alexiadis et al. 2004) was conducted by the Federal Highway Administration (FHWA) Traffic Analysis Tools Program. It recorded detailed trajectories of the vehicles on real roads using high-resolution cameras. The US Highway 101 (US 101) data set is the representative data set for highway traffic conditions. It reflects how vehicles moved over time on the 640-m-long segment of US Highway 101. While the full data set of US 101 witnessed heavy traffic during rush hour from 7:50 a.m. to 8:35 a.m., only the first 15-min period of data was used in this study. This is because in the last 30 min the highway became fully congested with frequent stop-and-go conditions, which deviated from a typical operating condition for CACC. In the first 15 min, more than 3,000 vehicles entered the highway segment, forming and reforming 2,500 pairs of preceding/following vehicles.

Sensor Error Model

Automotive Radar

Millimeter-wave radar is one of the fundamental sensors for automated vehicles. The setting of radar in the evaluations follows the Bosch long (Stuttgart, Germany)-range radar (LRR), which has been extensively used for ACC (Hasch et al. 2012). It is featured by 250-m detection range, 30° detection angle, 0.1-m distance accuracy, and 0.1° angle accuracy. The measurement errors were assumed to be white noise (Ploeg et al. 2014).

GPS

Although the most accurate real-time kinematic GPS (RTK-GPS) is capable of centimeter-level positioning, it is too expensive for large-scale implementation. Instead, the pseudorange relative positioning approach (Alam et al. 2013; Müller et al. 2014; Liu et al. 2013) is assumed to be adopted by PVIS-mixed because it can achieve better performance than normal differential GPS (DGPS) while neither expensive hardware nor support of reference station is needed. The accuracy of pseudorange relative positioning can reach 0.5–1.2 m on open highway (Müller et al. 2014; Liu et al. 2013) and 2–6.5 m in dense urban environments (Alam et al. 2013; Müller et al. 2014). Because the pseudorange DGPS's errors follow unbiased normal distribution from long-term observations (Matosevic et al. 2006), it is reasonable to assume that pseudorange relative positioning has the same long-term error pattern due to the technical similarity.

In pseudorange relative positioning, the error caused by satellite or user clock difference and the Earth's atmosphere can be mostly eliminated, so the major error source would be the multipath effect caused by roadside buildings/trees (Müller et al. 2014). Based on Giremus et al. (2007), the multipath effect can be seen as "abruptly adding biases with random magnitudes and durations to pseudorange measurements." Accordingly, an error model for pseudorange relative positioning was proposed in Chen and Park (2019) as follows:

$$e(t) \approx e_m(t) + \epsilon \tag{10}$$

$$e_m(t) = \sum_{1}^{i} \psi(t, t_{i-1}, t_i) b_i \tag{11}$$

$$\psi(t, t_{i-1}, t_i) = \begin{cases} 1 & \text{if } t_{i-1} < t \leq t_i \\ 0 & \text{else} \end{cases} \tag{12}$$

where t = time; $e(t)$ = overall positioning error in x- or y-direction; $e_m(t)$ = multipath effect error; and ε = normally distributed unmodeled error

$$\epsilon \sim N(0, \sigma_u^2) \tag{13}$$

where σ_u = standard deviation (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 $t_{i-1} - t_i$

$$b_i \sim N(0, \sigma_b^2) \tag{14}$$

$$(t_{i-1} - t_i) \sim U(t_{\min}, t_{\max}) \tag{15}$$

where σ_b = STD of positioning bias; and $t_{\min}$ and $t_{\max}$ = lower and upper boundary of the duration of the multipath effect.

While the error propagation from pseudorange measurements to position measurement is simplified, the model [Eqs. (10)–(15)] makes sense. This is because from long-term observations the overall positioning error tends to comply with an unbiased normal distribution with an STD of $\sqrt{\sigma_u^2 + \sigma_b^2}$, but from short-term observations the distribution of overall error is biased and skewed.

Eqs. (10)–(15) can only be applied to PVIS-mixed evaluation, not to the design of PVIS-mixed, because real-world vehicles can only obtain the statistics of overall error, either from long-term GPS tests or onboard sensor fusion (Hult et al. 2018).

In this paper, the STD of overall positioning errors in x- and y-directions are set the same, and the irreducible error $\epsilon = \min(\sigma_x) = 0.5$ m is assumed, so the STD of the multipath bias in Eq. (14) can be set as $\sigma_{b,x} = \sqrt{\sigma_x^2 - 0.5^2}$ and $\sigma_{b,y} = \sqrt{\sigma_y^2 - 0.5^2}$ snd duration of the multipath effect is assumed to be 10–30 s.

An exemplary snapshot of GPS positioning error in the x-direction generated by this model over 200 s and the error's frequency histogram are shown in Fig. 4 (overall error STD is set as 1 m). It can be seen that the bias of GPS measurement fluctuates over time and brings higher uncertainty to the generated positioning error than a simple normal distribution can.

Simulation

The evaluation of PVIS-mixed was conducted in MATLAB. Basically, the simulation follows the flowchart in Fig. 2. The steps of simulation are explained in detail as follows:

1. Initialize. The simulation was driven by two parameters: t and j. Parameter t means time, ranging from 0.0 to 900.0 s with a

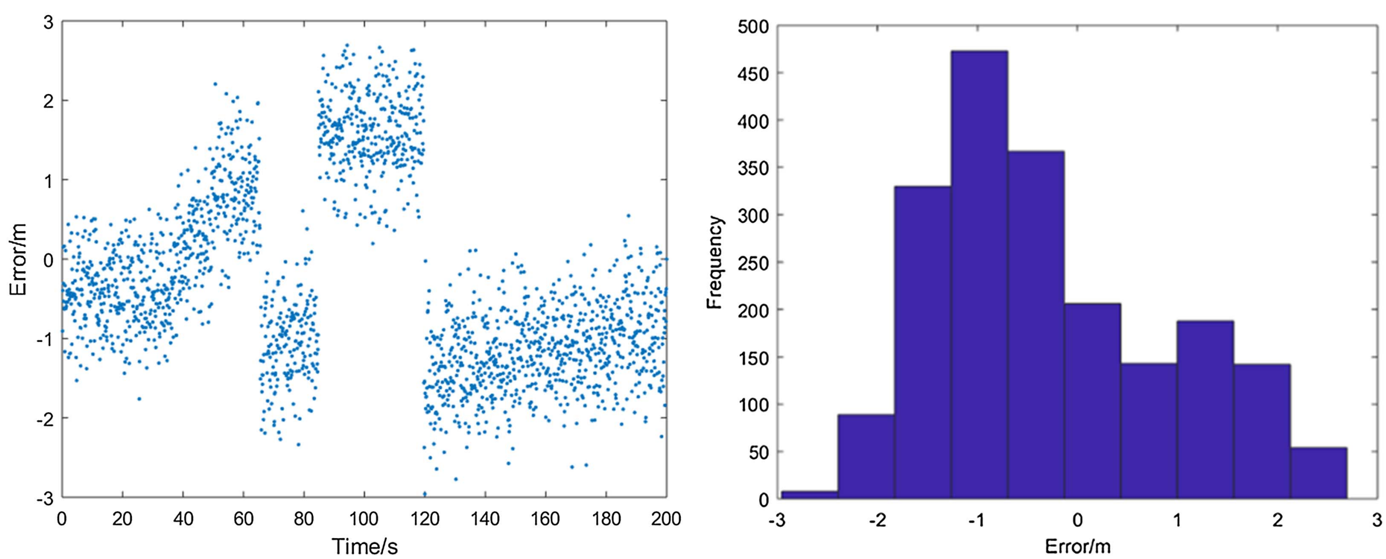

Fig. 4. Positioning error generated by the GPS error model.

resolution of 0.1 s. Parameter j means the sequence 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). Connectivity is randomly assigned to the vehicles. At the beginning of simulation, t and j were set as 0 and 1, respectively.
2. Obtain vehicle locations from NGSIM data. With an input of t and j, the locations of the jth vehicle (i.e., subject vehicle) and locations of surrounding vehicles within 200 m 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. No identification is needed if there is no vehicle ahead of the jth subject vehicle within 200 m or the current preceding vehicle has already been identified.
4. Generate sensor measurements. Radar measurements of the preceding vehicle and GPS measurements of other surrounding vehicles are generated by adding random sensor errors to the actual relative vehicle locations (which have been obtained in Step 2).
5. Update the candidate list. If this is the first search for the jth vehicle, every surrounding vehicle whose GPS location is within the searching area is added on the candidate list. For the $2-n$ searches, the former candidates that are not found at this time are removed from the candidate list. If the candidate list becomes void after checking all the surrounding vehicles, stop this identification trial immediately and restart another one if the number of identifications has not reached k.
6. Determine whether the identification procedure for the jth vehicle should be finished at this time. The procedure ends if the number of searches reaches n and one vehicle is still left on the candidate list, or the list is void but k trials of identification have been used.
7. Determine the search result. A results matrix is defined: T (9,000 rows, 3,000 columns), where each element $T(t,m)$ means the search result for the jth subject vehicle at the time t, where m is the vehicle ID of the jth subject vehicle. The default value of $T(t,j)$ is 0, indicating no identification needed for the jth subject vehicle at t. When Step 6 decides this is not the time to finish, $T(t,j)$ is to be set to 2, indicating an undecided preceding vehicle at t, and the identification should go on to $t+0.1$s; when only one candidate is found ($N=1$), check whether its vehicle ID matches that of the preceding vehicle. If positive, change $T(t,m)$ to 1, meaning a correct identification; otherwise, change $T(t,m)$ to -1, meaning an incorrect identification (i.e., an error). When no candidate is found, check the connectivity of the preceding vehicle. If it is unconnected, $T(t,j)$ is to be set to 3, indicating a correct decision on the connectivity of preceding vehicle; otherwise $T(t,j)$ is to be set to 4, indicating a failure to find the connected preceding vehicle.
8. Go through the Steps 1–7 for all subject vehicles and all time.

After obtaining the matrix T, the effectiveness of PVIS-mixed can be measured. By screening each row of T, the number of 2 before every 1 or -1 can be counted; thus, the time consumption for each final decision can be calculated. The average time consumption and 99th percentile of time consumption were found to represent the efficiency of the PVIS-mixed.

At last, the error rate and unusability rate can be computed as follows:

$$Er = \text{No. of} -1/(\text{No. of} -1 + \text{No. of } 1))$$
$$Ur = \text{No. of } 4/(\text{No. of } 4 + \text{No. of } 1))$$

Evaluation Results

PVIS-mixed was evaluated under three different adoption rates: 30%, 60%, and 90%. Following the previous study (Chen and Park 2019), $Er^* = 10^{-8}$ was used in the evaluation.The results are given in Table 2. The key findings are summarized as follows:

- The results showed no significant correlation between the performance and adoption rate of vehicle connectivity. This is because PVIS-mixed always considers the worst case as shown in Fig. 3, no matter the proportion of connected vehicles on the road.
- The time consumption increases with GPS error. The required GPS accuracy was found to be 1 m to make the 99th percentile time consumption <10 s; 0.8 m is the required accuracy to make that <5 s. These accuracy levels are realizable in a highway scenario with low-cost GPS (Müller et al. 2014; Liu et al. 2013), but difficult to be achieved in an urban environment (Alam et al. 2013; Müller et al. 2014). Nevertheless, the simulation was conducted on an urban highway with heavy traffic, and thus led to more conservative results. To some extent, the required GPS accuracy may be relaxed in lower-density cases such as rural highway or uncongested urban roadway.
- The Ur was not as low as its theoretical value. Except for accuracy = 0.5 m (where there is no multipath effect), Ur for all the other accuracy levels increase by 1%–4%. This is because the multipath effect is simulated in the evaluation but not able to be considered in the design of PVIS-mixed. Under the effect of multipath bias, the continuously large positioning error can occur to the actual preceding vehicle and make its GPS position out of the searching area during the identification procedure. Nevertheless, an $Ur < 5\%$ can still be guaranteed when GPS standard error is no greater than 1 m.
- In terms of incorrect identification, the actual Er is kept at 0% in all the runs, indicating the robustness of PVIS-mixed against incorrect identification.

To further verify the necessity of PVIS-mixed, the previous PVIS (Chen and Park 2019) was tested in the mixed traffic

Table 2. Evaluation results under different adoption rates

Adoption rate (%)	GPS standard error (m)	Actual, Er (%)	Average time consumption (s)	99th percentile time consumption (s)	Ur (%)
30	0.5	0.00	0.4	1.2	0.17
	0.6	0.00	0.7	1.8	0.87
	0.7	0.00	1.4	2.7	2.07
	0.8	0.00	2.7	5.1	3.37
	0.9	0.00	3.6	7.5	3.29
	1	0.00	4.1	8.2	4.07
	1.1	0.00	5.7	12.2	5.76
60	0.5	0.00	0.4	1	0.00
	0.6	0.00	0.7	1.7	1.02
	0.7	0.00	1.4	2.9	1.90
	0.8	0.00	2.6	4.3	2.67
	0.9	0.00	3.6	6.5	3.35
	1	0.00	4.3	9.7	3.54
	1.1	0.00	5.7	12.5	5.74
90	0.5	0.00	0.4	1.2	0.00
	0.6	0.00	0.7	1.7	1.26
	0.7	0.00	1.4	2.9	1.97
	0.8	0.00	2.7	5.2	1.74
	0.9	0.00	3.6	7.1	3.10
	1	0.00	4.2	9.4	4.97
	1.1	0.00	5.9	13.4	4.25

Table 3. Performance comparison of PVIS and PVIS-mixed under 1-m GPS error and different adoption rates

Adoption rate (%)	Method	Actual, Er (%)	Average time consumption (s)	99th percentile time consumption (s)	Ur (%)
30	PVIS	38.14	4.43	33.60	0.00
	PVIS-22.8	36.46	4.39	22.60	9.45
	PVIS-10	24.07	3.15	9.80	26.59
	PVIS-mixed	0.00	4.10	8.20	4.07
60	PVIS	23.73	2.05	20.50	0.00
	PVIS-22.8	21.82	2.05	21.60	1.08
	PVIS-10	17.93	1.60	9.80	8.12
	PVIS-mixed	0.00	4.30	9.70	3.54
90	PVIS	5.46	0.75	9.90	0.00
	PVIS-22.8	5.59	0.67	9.70	0.13
	PVIS-10	5.23	0.64	9.40	0.85
	PVIS-mixed	0.00	4.20	9.40	4.97

environment. Table 3 compares the evaluation results of PVIS and PVIS-mixed under 1-m GPS standard error and different adoption rates of connectivity.

To make the comparison even fairer, we also tested PVIS (Chen and Park 2019) enhanced with cutoff search time, after which PVIS can terminate the identification and conclude that the preceding vehicle is unconnected. Two such PVISs are defined:

1. PVIS-22.8: The time consumption was cut off at 22.8 s, which is the maximum time consumption allowed by PVIS-mixed under 1-m GPS standard error.
2. PVIS-10: The time consumption was cut off at 10 s, the general requirement for 99th percentile time consumption in Chen and Park (2019) and this study.

As expected, PVIS resulted in large numbers of misidentifications. Even when there was a cutoff time consumption, the actual Er could only be reduced slightly and at the cost of higher Ur. Also, the performance of PVIS improves with the increase of adoption rate. Nevertheless, none of PVIS, PVIS-22.8 or PVIS-10 are fundamentally suitable for mixed traffic due to the poor reliability.

Conclusions and Future Work

In this research, a preceding vehicle identification system for mixed traffic (PVIS-mixed) was developed to facilitate cooperative platooning operation under imperfect market penetration. The parameters of PVIS-mixed were optimized considering the probability to make incorrect identification, the probability to miss the actual preceding vehicle, and the time consumption of the identification procedure. The proposed PVIS-mixed was tested by simulation utilizing real vehicle trajectory data from NGSIM. To make the test more realistic, the multipath bias of GPS positioning was modeled. The results showed that the performance of PVIS-mixed was irrelevant to the adoption rate of connected vehicles, and the required GPS accuracy to make 99th percentile time consumption <10 s was 1 m, with the theoretical error rate of 10^{-8}. The results also showed that the usability of CACC was negatively affected by the multipath bias of GPS, which makes it easier to mistakenly take a connected preceding vehicle as unconnected.

An immediate improvement that can be made is to further lower the unusability rate of CACC, while the constraints on error rate and time consumption should still be met. It can be found that a critical variable that limits the efficiency of PVIS-mixed is the minimum lateral intervehicle distance w. When a constant w is used, it is either so conservative that no irrelevant vehicle is that close to the preceding vehicle in most cases, or too loose to include the worst situation. An adaptive w is a possible solution to improve the performance of PVIS-mixed. In fact, when the subject vehicle is equipped with a wider-angle radar or lidar, the subject vehicle should be capable of detecting the existence of the first nearest irrelevant vehicle and determining a more reasonable w based on how close it is to the preceding vehicle. However, higher cost on the range sensor can also be expected for its wider detection angel.

Additional future work includes extending this PVIS-mixed for more connected vehicle applications such as cooperative lane change (Dolk et al. 2017) or cooperative platooning controls dedicated for mixed traffic (Chen and Park 2020; Jin and Orosz 2018) where more surrounding vehicles need to be identified rather than the preceding vehicle. Hardware-in-the-loop simulations and real-world experiments are also necessary steps to prepare PVIS-mixed for future implementation.

Data Availability Statement

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request:

- Optimization program used to parameterize the PVIS-mixed,
- The vehicle trajectory data used for the PVIS-mixed evaluation, and
- The simulation code used for the PVIS-mixed evaluation.

Acknowledgments

This research is in part supported by the US National Science Foundation under Grant No. CMMI-2009342, and in part supported by a generous unrestricted research fund given by Toyota Motor North America R&D.

References

Alam, N., A. Kealy, and A. G. Dempster. 2013. "An ins-aided tight integration approach for relative positioning enhancement in VANETs." *IEEE Trans. Intell. Transp. Syst.* 14 (4): 1992–1996. https://doi.org/10.1109/TITS.2013.2265235.

Alexiadis, V., J. Colyar, J. Halkias, R. Hranac, and G. McHale. 2004. "The next generation simulation program." *Inst. Transp. Eng. ITE J.* 74 (8): 22–26.

Al-Jhayyish, A. M., and K. W. Schmidt. 2017. "Feedforward strategies for cooperative adaptive cruise control in heterogeneous vehicle strings." *IEEE Trans. Intell. Transp. Syst.* 19 (1): 113–122. https://doi.org/10.1109/TITS.2017.2773659.

Bansal, P., and K. M. Kockelman. 2017. "Forecasting Americans' long-term adoption of connected and autonomous vehicle technologies." *Transp. Res. Part A Policy Pract.* 95 (Jan): 49–63. https://doi.org/10.1016/j.tra.2016.10.013.

Bengler, K., K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner. 2014. "Three decades of driver assistance systems: Review and future perspectives." *IEEE Intell. Transp. Syst. Mag.* 6 (4): 6–22. https://doi.org/10.1109/MITS.2014.2336271.

Chen, Z., and B. B. Park. 2019. "Preceding vehicle identification for cooperative adaptive cruise control platoon forming." *IEEE Trans. Intell. Transp. Syst.* 21 (1): 308–320. https://doi.org/10.1109/TITS.2019.2891353.

Chen, Z., and B. B. Park. 2020. "Cooperative adaptive cruise control with unconnected vehicle in the loop." *IEEE Trans. Intell. Transp. Syst.* 1–11. https://doi.org/10.1109/TITS.2020.3041840.

Dolk, V., J. den Ouden, S. Steeghs, J. G. Devanesan, I. Badshah, A. Sudhakaran, K. Elferink, and D. Chakraborty. 2017. "Cooperative automated driving for various traffic scenarios: Experimental validation in the GCDC 2016." *IEEE Trans. Intell. Transp. Syst.* 19 (4): 1308–1321. https://doi.org/10.1109/TITS.2017.2750079.

Geiger, A., M. Lauer, F. Moosmann, B. Ranft, H. Rapp, C. Stiller, and J. Ziegler. 2012. "Team AnnieWAY's entry to the 2011 Grand Cooperative Driving Challenge." *IEEE Trans. Intell. Transp. Syst.* 13 (3): 1008–1017. https://doi.org/10.1109/TITS.2012.2189882.

Giremus, A., J.-Y. Tourneret, and V. Calmettes. 2007. "A particle filtering approach for joint detection/estimation of multipath effects on GPS measurements." *IEEE Trans. Signal Process.* 55 (4): 1275–1285. https://doi.org/10.1109/TSP.2006.888895.

Hasch, J., E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt. 2012. "Millimeter-wave technology for automotive radar sensors in the 77 GHz frequency band." *IEEE Trans. Microwave Theory Tech.* 60 (3): 845–860. https://doi.org/10.1109/TMTT.2011.2178427.

Hult, R., F. E. Sancar, M. Jalalmaab, A. Vijayan, A. Severinson, M. Di Vaio, P. Falcone, B. Fidan, and S. Santini. 2018. "Design and experimental validation of a cooperative driving control architecture for the grand cooperative driving challenge 2016." *IEEE Trans. Intell. Transp. Syst.* 19 (4): 1290–1301. https://doi.org/10.1109/TITS.2017.2750083.

Jin, I. G., and G. Orosz. 2018. "Connected cruise control among human-driven vehicles: Experiment-based parameter estimation and optimal control design." *Transp. Res. Part C Emerging Technol.* 95 (Oct): 445–459. https://doi.org/10.1016/j.trc.2018.07.021.

Levine, W., and M. Athans. 1966. "On the optimal error regulation of a string of moving vehicles." *IEEE Trans. Autom. Control* 11 (3): 355–361. https://doi.org/10.1109/TAC.1966.1098376.

Lioris, J., R. Pedarsani, F. Y. Tascikaraoglu, and P. Varaiya. 2017. "Platoons of connected vehicles can double throughput in urban roads." *Transp. Res. Part C Emerging Technol.* 77 (Apr): 292–305. https://doi.org/10.1016/j.trc.2017.01.023.

Liu, K., H. B. Lim, E. Frazzoli, H. Ji, and V. C. Lee. 2013. "Improving positioning accuracy using GPS pseudorange measurements for cooperative vehicular localization." *IEEE Trans. Veh. Technol.* 63 (6): 2544–2556. https://doi.org/10.1109/TVT.2013.2296071.

Matosevic, M., Z. Salcic, and S. Berber. 2006. "A comparison of accuracy using a GPS and a low-cost DGPS." *IEEE Trans. Instrum. Meas.* 55 (5): 1677–1683. https://doi.org/10.1109/TIM.2006.880918.

Milanés, V., S. E. Shladover, J. Spring, C. Nowakowski, H. Kawazoe, and M. Nakamura. 2013. "Cooperative adaptive cruise control in real traffic situations." *IEEE Trans. Intell. Transp. Syst.* 15 (1): 296–305. https://doi.org/10.1109/TITS.2013.2278494.

Müller, F., E. M. Diaz, B. Kloiber, and T. Strang. 2014. "Bayesian cooperative relative vehicle positioning using pseudorange differences." In *Proc., IEEE/ION PLANS 2014*, 434–444. New York: IEEE.

Naus, G. J., R. P. Vugts, J. Ploeg, M. J. van De Molengraft, and M. Steinbuch. 2010. "String-stable CACC design and experimental validation: A frequency-domain approach." *IEEE Trans. Veh. Technol.* 59 (9): 4268–4279. https://doi.org/10.1109/TVT.2010.2076320.

Ploeg, J., B. T. Scheepers, E. Van Nunen, N. Van de Wouw, and H. Nijmeijer. 2011. "Design and experimental evaluation of cooperative adaptive cruise control." In *Proc., 2011 14th Int. IEEE Conf. on Intelligent Transportation Systems (ITSC)*, 260–265. New York: IEEE.

Ploeg, J., E. Semsar-Kazerooni, G. Lijster, N. van de Wouw, and H. Nijmeijer. 2014. "Graceful degradation of cooperative adaptive cruise control." *IEEE Trans. Intell. Transp. Syst.* 16 (1): 488–497. https://doi.org/10.1109/TITS.2014.2349498.

Shladover, S. E., D. Su, and X.-Y. Lu. 2012. "Impacts of cooperative adaptive cruise control on freeway traffic flow." *Transp. Res. Rec.* 2324 (1): 63–70. https://doi.org/10.3141/2324-08.

Singh, S. 2015. *Critical reasons for crashes investigated in the national motor vehicle crash causation survey.* Washington, DC: USDOT, National Highway Traffic Safety Administration, National Center for Statistics and Analysis.

WHO (World Health Organization). 2018. *Global status report on road safety 2018: Summary.* Geneva: WHO.