

An Efficient Distributed Multivehicle Cooperative Tracking Framework via Multicast

Sijiang Li, *Graduate Student Member, IEEE*, Dongliang Duan[✉], *Senior Member, IEEE*,
Jianan Zhang[✉], *Member, IEEE*, Xiang Cheng[✉], *Fellow, IEEE*, and Liuqing Yang[✉], *Fellow, IEEE*

Abstract—To support various upper applications of intelligent vehicles ranging from driving assistance to automated planning and control, accurate localization, and tracking are the fundamental tasks. Given the limited versatility and efficiency of traditional single-vehicle multisensor and multivehicle multisensor localization and tracking solutions, this article presents an efficient distributed multivehicle cooperative tracking framework via multicast. Once the self-positioning data is locally fused with assistance from roadside units, each vehicle shares the local-fusion results with surrounding vehicles through multicast and observes surrounding vehicles with on-board sensing equipment. The vehicles can then jointly feed the local-fusion results, received multicast information, and observation results into a global filter to obtain accurate and robust cooperative tracking. By leveraging multicast, the communication load is reduced, which promotes the efficiency of communication resource utilization. By optimizing the data fusion procedure, the error caused by error correlation is eliminated and the sensitivity to non-ideal conditions, including packet loss, interruption, time-varying cooperative vehicles, etc., is reduced, which improves the versatility of the framework in real-world applications. Furthermore, several practical issues, such as random communication delay, packet loss, communication load, and localization robustness are also involved. To verify the effect of the framework, both theoretical analyses and simulation results are presented to show the accuracy and robustness of our proposed cooperative tracking framework.

Index Terms—Autonomous driving, cooperative tracking, intelligent transportation systems (ITSs), Internet of Vehicles (IoV), multicast.

Manuscript received 15 May 2023; revised 3 July 2023; accepted 17 July 2023. Date of publication 2 August 2023; date of current version 25 December 2023. This work was supported in part by the Ministry National Key Research and Development Project under Grant 2020AAA0108101; in part by the National Natural Science Foundation of China under Grant 62125101 and Grant 62341101; and in part by the Guangzhou Municipal Science and Technology Project under Grant 2023A03J0011. (Corresponding author: Xiang Cheng.)

Sijiang Li, Jianan Zhang, and Xiang Cheng are with the State Key Laboratory of Advanced Optical Communication Systems and Networks, School of Electronics, Peking University, Beijing 100871, China (e-mail: pkulsj@pku.edu.cn; zhangjianan@pku.edu.cn; xiangcheng@pku.edu.cn).

Dongliang Duan is with the Department of Electrical and Computer Engineering, University of Wyoming, Laramie, WY 82071 USA (e-mail: dduan@uwyo.edu).

Liuqing Yang is with the Internet of Things Thrust and Intelligent Transportation Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou 510000, China, and also with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, SAR, China (e-mail: lqyang@ust.hk).

Digital Object Identifier 10.1109/IJOT.2023.3301055

I. INTRODUCTION

AS AN essential part of the intelligent transportation system (ITS), intelligent vehicles, and autonomous driving technology have attracted more and more attention in recent years [1]. There are mainly five modules for autonomous driving, namely, localization, perception, planning, control, and system management [2]. Accurate localization and tracking of vehicles are fundamental for autonomous driving, as they provide precise information about the position, attitude (i.e., localization), and continuous trajectory of vehicles (i.e., tracking). This information is essential for various aspects of autonomous driving, ranging from environment perception to vehicle control. The key performance metrics of interest in localization and tracking are accuracy and robustness [3]. In terms of accuracy, a root mean squared error (RMSE) of less than 5 m is needed to justify which road the vehicle is in, less than 1.5 m is needed for lane, and less than 1 m is needed for the detailed position in the lane. For active control of the vehicle, the RMSE should be less than 0.1 m [4]. In terms of robustness, the accuracy must be maintained at the desired level in all driving conditions, including complex dynamic urban environments and various harsh weather conditions, such as rain and snow. In summary, a subdecimeter-level localization and tracking scheme that is robust in complex road environments is desired for autonomous driving [5].

In recent years, numerous methods have been proposed to provide accurate localization [6] using various technologies, such as global positioning system (GPS), inertial measurement unit (IMU), and simultaneous localization and mapping (SLAM) based on light detection and ranging (LiDAR), camera, etc. However, it is challenging to tradeoff accuracy, robustness, and cost for single-vehicle localization using only one or more on-board devices. With the development of Internet of Vehicles (IoV) and 5G communication techniques, wireless interconnections with high quality of services, such as high bandwidth and low latency among vehicles are possible in the near future, which supports the ITS to have stronger cooperative intelligence capability [7], [8], [9], [10], [11]. In such a context, multivehicle cooperative localization methods have been proposed to address the limitations of single-vehicle localization. In our previous work [12], we proposed a multisensor multivehicle (MSMV) cooperative localization and tracking framework, which takes full advantage of various on-board equipment, vehicle-vehicle cooperation, and vehicle-road cooperation. However, this framework also has issues, such as direct vehicle-to-vehicle (V2V) communication load and multivehicle performance loss.

To leverage emerging communication technologies and take full advantage of cooperative tracking, this article presents a multicast-based distributed cooperative tracking framework. Similar to [12], it continues to use a variety of sensors and RSUs, as well as the vehicles' dynamic model for the tracking process. However, it improves the scheme of the information interactions and the structure of the fusing algorithm to further improve performance and efficiency. Specifically, the two-layer filtering structure is replaced by a new local-fusion global-filtering framework. In the local-fusion phase, vehicles fuse self-positioning data from GPS, IMU, etc., with the auxiliary positioning data from RSUs to integrate vehicle-road cooperation into the vehicle-vehicle cooperation framework. Then, they share the locally fused data with surrounding vehicles through multicast communications. With malicious user detection algorithms like [13], trusted cooperators can be selected among surrounding vehicles as multicast members. In the meantime, each vehicle takes observations on surrounding vehicles to obtain relative states with them using the on-board sensing systems, including LiDAR, camera, etc. Finally, each vehicle feeds data from different sources into the global-filtering algorithm independently to generate cooperative localization and tracking results, which constitutes a distributed framework. Delay compensation is also utilized to ensure localization accuracy.

Compared with [12], the main contributions of this article lie in the optimization of communication content and data fusion procedure, which leads to gains in communication load and versatility. Specifically, the communication content enables the utilization of multicast to reduce communication load. The data fusion procedure can handle a time-varying number of cooperative vehicles, keep robust in lossy communication environments, and eliminate the asymptotic localization error as the number of vehicles increases. This adaptability makes the framework applicable to the real IOV scenarios. The contributions of this article are in the following aspects.

- 1) The optimization of the communication content facilitates the utilization of multicast to reduce communication load and avoids complex V2V direct communication, which improves the efficiency of utilizing communication resources.
- 2) The optimization of the data fusion procedure reduces the number of filters and enhances the resistance to packet loss, random delay, communication collision, interruption, and time-varying cooperative vehicles, which improves the scenario versatility.
- 3) The asymptotic error caused by error correlation is analyzed and eliminated. The performance of the framework is shown in analytical forms and simulation results correspondingly, which shows the effectiveness of cooperative tracking.

The remainder of this article is organized as follows. The related work on localization and tracking are presented in Section II. The system model and problem formulation are presented in Section III. The multicast-based cooperative tracking algorithm and theoretical analysis are presented in Section IV. Then, some numerical simulations

are given in Section V to evaluate the performance of the framework. Finally, conclusions and ongoing research issues are highlighted in Section VI.

II. RELATED WORK

To realize localization and tracking, a traditional method is the dead reckoning derived from navigation, which uses accelerators and gyroscopes in the IMU to conduct low-cost localization and tracking [14], [15]. Due to its reliance on accurate initial positioning and susceptibility to accumulation error, using only the IMU for tracking is not feasible. However, the IMU's insensitivity to environmental factors makes it a valuable assistant to real-time absolute localization methods such as GPS. Furthermore, some technologies rely on a priori map for localization. With a high-precision map of the area, the vehicle can match the map using preset laser or other forms of labels, thus providing stable and high-precision positioning [16], [17]. However, due to the high cost of building and updating the map, it is not practical for open areas.

One of the most widely used low-cost outdoor positioning methods is GNSS, including GPS, GLONASS, BeiDou, Galileo, etc. However, the accuracy of traditional GNSS ranges from several meters to even tens of meters in open outdoors [18], providing a reference only at the road level. Additionally, GNSS faces accuracy reduction and signal loss issues in closed environments, such as tunnels and underground areas. New methods derived from traditional GNSS have been developed, with real-time kinematic GPS being the most advanced. To improve the accuracy, new methods derived from traditional GNSS have been developed, with real-time kinematic GPS being the most advanced. It uses the phase correlation between base station signals and GPS signals to improve accuracy, achieving centimeter-level accuracy in fixed solutions [19] with a good signal connection. However, it suffers from ten times more terminal cost compared to regular GPS, limited coverage of ground stations, and higher sensitivity to occlusion environments. Consequently, GNSS alone cannot meet the precision and robustness requirements for localization and tracking in autonomous driving, necessitating the use of other sensors in combination.

Intelligent vehicles are typically equipped with various sensing devices, such as LiDAR [20], [21], camera [22], [23], millimeter-wave radar [24], etc. SLAM using these sensors [25] enables positioning without the need for a prior map. By installing a variety of high-performance sensors on a single vehicle, information fusion can be used to achieve more accurate localization and tracking through filtering algorithms, such as the Kalman filter [26], [27] and particle filter [28], [29]. However, utilizing high-precision sensors on a single vehicle can be costly in practice. Additionally, the performance of sensors on a single vehicle tends to be correlated. For instance, in challenging conditions like long dark tunnels, multiple sensors may be affected simultaneously. GPS localization can be lost due to satellite signal occlusion, LiDAR point clouds may be difficult to be matched due to the lack of structural features

on smooth walls, and camera images may be unclear in dim light. In such cases, the additional cost of multisensor single-vehicle localization and tracking may not result in significant performance gains.

Since the challenges of high costs and adverse conditions are difficult to overcome in both the single-sensor single-vehicle and multisensor single-vehicle strategies, an alternative approach is to leverage multivehicle cooperation. Some early studies like [30] and [31] have proposed the concept of cooperative localization in wireless sensor networks. Through Bayesian inference, recent research has optimized the location of multiple nodes by measuring the angle of arrival [32] and transmission range [33] using wireless signals.

However, this approach only takes into account quasi-stationary nodes and does not account for the mobility dynamics of vehicles. In a vehicular environment, by utilizing the multipath channel between the base station and vehicles [34], as well as between vehicles themselves [35], vehicles can achieve radio-based cooperative SLAM. In [36] and [37], GPS raw data at a low level are exchanged to improve GPS localization by exploiting multivehicle trajectory correlation. Notably, in such methods, the GPS information and cooperative information of multiple vehicles are closely integrated. However, this tightly coupled scheme is only suitable for specific device scenarios.

As the loosely coupled scheme is more general to scenarios with heterogeneous vehicle sensor types, [38], [39] reduce the coupling between self-localization and relative position correlation to improve applicability. Using dedicated short-range communication signal processing, information, such as time of arrival, received signal strength, and Doppler shift are extracted to calculate relative localization. Then, filtering and joint optimization methods are used, respectively, to optimize localization. However, it should be noted that since relative localization is obtained through wireless signals, these methods are inapplicable to mainstream intelligent vehicles equipped with sensors, such as LiDAR and cameras.

Utilizing on-board sensing devices, [40] has proposed a cooperative LiDAR-SLAM framework, which shares LiDAR point clouds to achieve relative pose registration between vehicles and robust tracking. However, this framework does not incorporate vehicle-road cooperation, and the transmission of multiline LiDAR point clouds imposes a significant communication load. In [12], we proposed a general architecture that enables the fusion of information from various sources, such as surrounding vehicles, road side units (RSUs), and different sensors, such as GPS, IMU, LiDAR, camera, etc. This architecture utilizes a two-layer filtering algorithm and shows advantages in terms of cost, accuracy, generality, and robustness. However, in practical applications, communication interruptions can cause instabilities in the algorithm due to changing cooperation scales. Additionally, the direct V2V communication load is high, and the two-layer algorithm proposed in [12] has high complexity and performance loss caused by error correlation. These limit the versatility and efficiency of the algorithm.

As multicast has been considered a resource-efficient way to transmit the same content to multiple users, the 3GPP 5G

Release 17 supports multicast and broadcast services over the existing 5G framework. Research efforts have focused on achieving point-to-multipoint communication with low latency and high reliability. For example, in [41], unmanned aerial vehicles (UAVs) are employed to assist in the dissemination of data to a group of vehicles. One notable advantage of multicast over broadcast is its ability to choose receivers, which better ensures data security. Therefore, multicast is suitable for intelligent vehicles in terms of efficiency and security.

III. SYSTEM MODEL

As stated in our previous work [12], the dynamic process of vehicles and observation can be described by a first-order hidden Markov model [28]. It is assumed that vehicles are equipped with IMU, GPS, and one or more sensing devices to possess a fundamental localization and sensing capability. In Section III-A, the dynamic process is modeled as a state transfer function and the self-positioning and sensing of the vehicles are modeled as observations. In Section III-B, RSUs' auxiliary positioning is introduced to extend the framework with vehicle-road cooperation. Furthermore, the effect of communication delay is modeled in Section III-C, along with compensation methods. The key models are briefly described as follows and more details can be found in [12].

A. State Transfer and Observation Model

We are interested in localizing and tracking an object in the 2-D plane using Cartesian coordinates x and y . Angle component can be added into state equations and filtering algorithms, if we consider vehicles' orientation [42]. For a vehicle V_i , its mobility can be described by a system state transfer function

$$\mathbf{x}_i[k] = \mathbf{A}\mathbf{x}_i[k-1] + \mathbf{B}\mathbf{u}_i[k] + \mathbf{w}_i[k] \quad (1)$$

and

$$\mathbf{x}_i = \begin{pmatrix} x_i \\ \dot{x}_i \\ y_i \\ \dot{y}_i \end{pmatrix}, \mathbf{u}_i = \begin{pmatrix} a_{i,x} \\ a_{i,y} \end{pmatrix}, \mathbf{w}_i = \begin{pmatrix} w_{x_i} \\ w_{\dot{x}_i} \\ w_{y_i} \\ w_{\dot{y}_i} \end{pmatrix}$$

$$\mathbf{A} = \begin{pmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathbf{B}\mathbf{u} = \begin{pmatrix} \frac{\Delta t^2}{2} & 0 \\ \Delta t & 0 \\ 0 & \frac{\Delta t^2}{2} \\ 0 & \Delta t \end{pmatrix}$$

where $\mathbf{x}_i$ is the state vector, containing positions and velocities in a Cartesian coordinate; $a_{i,x}$ and $a_{i,y}$ are the acceleration of the vehicle, which are generated from the control system of a vehicle and measured by IMU; $\mathbf{w}_i$ is the process noise, which is the additive white Gaussian noise (AWGN) with a covariance matrix $\mathbf{Q}_i$; The value of $\mathbf{Q}_i$ is approximately proportional to the time step length Δt ; matrices $\mathbf{A}$ and $\mathbf{B}\mathbf{u}$ are obtained by the physical dynamics. To discretize the continuous motion of the vehicle, the typical value of Δt is 0.1 s, and k is the discrete time index. The GPS and sensors report data also at a rate of 10 Hz.

The observation data of V_i consist of two parts: 1) the measurement provided by sensors, such as GPS and wheel-speed

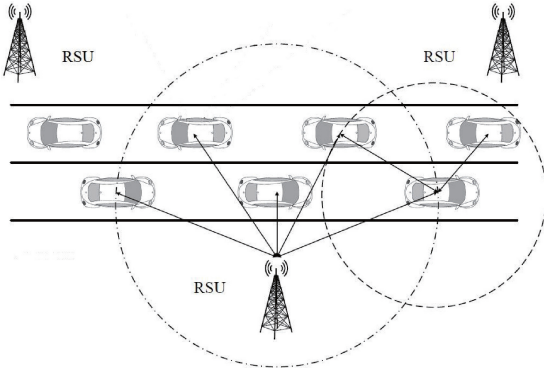


Fig. 1. IoV with RSUs [12].

sensors with information only related to its position and velocity, denoted as $\mathbf{z}_i$ and 2) the measurement provided by the sensing systems related to both its own and surrounding V_j 's states, denoted as $\mathbf{z}_{i \rightarrow j}$.

For $\mathbf{z}_i$, we have

$$\mathbf{z}_i[k] = \mathbf{H}_i \mathbf{x}_i[k] + \mathbf{v}_i[k]. \quad (2)$$

For $\mathbf{z}_{i \rightarrow j}$, we have

$$\mathbf{z}_{i \rightarrow j}[k] = \mathbf{H}_{i \rightarrow j} \mathbf{x}_{i \rightarrow j}[k] + \mathbf{v}_{i \rightarrow j}[k] \quad (3)$$

where $\mathbf{v}_i$ and $\mathbf{v}_{i \rightarrow j}$ are measurement noises, $\mathbf{H}_i$ and $\mathbf{H}_{i \rightarrow j}$ are measurement matrices determined by properties of device, and $\mathbf{x}_{i \rightarrow j}[k] = \mathbf{x}_j[k] - \mathbf{x}_i[k]$ is the relative state between V_j and V_i . Without loss of generality, we assume that $\mathbf{H}_i$ and $\mathbf{H}_{i \rightarrow j}$ are identity matrices and the noises are AWGN, and the covariance matrices $\mathbf{R}_i$ and $\mathbf{R}_{i \rightarrow j}$ are diagonal.

B. RSU Auxiliary Positioning Model

In ITS, RSUs serve as road facilities that provide auxiliary service to vehicles within their communication range through on-board units, as shown in Fig. 1. In general, RSUs are considered auxiliary facilities with a precise prior location. In some traditional positioning methods, vehicles can be located using IMU without relying on GPS, thanks to RSUs and wireless positioning [43]. With the development of sensing devices, RSUs now possess perception capability.

Unlike moving vehicles, RSUs are stationary and therefore can provide more accurate positioning services. Additionally, RSUs are relatively limited in number, making the investment costs for implementing them in ITS moderate. Therefore, RSUs can serve as powerful participants in the cooperative tracking process. Particularly in enclosure spaces, such as tunnels and underground, where GPS suffers from signal loss, it is easier to implement RSU coverage to ensure positioning accuracy.

The m th RSU's positioning for V_i can be described as follows:

$$\begin{aligned} \mathbf{z}_{r_m, i}[k] &= \mathbf{z}_{r_m \rightarrow i}[k] + \mathbf{x}_{r_m}[k] \\ &= \mathbf{H}_{r_m \rightarrow i} \mathbf{x}_{r_m \rightarrow i}[k] + \mathbf{v}_{r_m \rightarrow i}[k] + \mathbf{x}_{r_m}[k] \\ &= \mathbf{x}_i[k] + \mathbf{v}_{r_m \rightarrow i}[k] \end{aligned} \quad (4)$$

where $\mathbf{z}_{r_m, i}[k]$ is the measurement of V_i by the m th RSU, consisting of relative measurement $\mathbf{z}_{r_m \rightarrow i}[k]$, and absolute state vector of RSU $\mathbf{x}_{r_m}[k]$. The error of $\mathbf{x}_{r_m}[k]$ can be ignored since the RSU is fixed and its location is known a priori. The measurement follows a similar format as (2), with the difference only in the error part. Thus, the measurement from the RSU can be considered as a special self-positioning for V_i . Owing to the high precision of RSU, the variance of $\mathbf{v}_{r_m \rightarrow i}[k]$ is much smaller than that in (2).

C. Compensation for Communication Delays

Similar to our previous work [12], we assume that there is a finite communication delay in any data-sharing process between two intelligent agents [44], including V2V and vehicle-to-road (V2R) communications. The delay can be measured by the timestamps reported by the sender [45]. In the current state-of-the-art IoV, the delay can be controlled within 0.1 s [45]. Neglecting the delay in localization can lead to drift in the results, particularly when the vehicle is traveling at high speeds and the communication quality is poor. This can undermine the advantages of cooperation and even introduce risks. During the cooperation, at time instant k , the self-positioning data $\mathbf{z}_j[k - k_{\tau_{j \rightarrow i}}]$ received by V_i from another vehicle V_j corresponds to $\mathbf{x}_j[k - k_{\tau_{j \rightarrow i}}]$, where $k_{\tau_{j \rightarrow i}} = \tau_{j \rightarrow i} / \Delta t$ is normalized delay. Due to V_j 's motion in $\tau_{j \rightarrow i}$, positioning drift can be introduced, and the magnitude of the drift increases with V_j 's speed. Similarly, auxiliary positioning data $\mathbf{z}_{r_m, i}[k - k_{\tau_{r_m \rightarrow i}}]$ received from the m th RSU corresponds to $\mathbf{x}_i[k - k_{\tau_{r_m \rightarrow i}}]$ and drifted due to V_i 's motion.

To ensure system performance, we utilize a compensation method that employs the state transfer function as (1) to estimate $\mathbf{z}_j[k]$ based on $\mathbf{z}_j[k - k_{\tau_{j \rightarrow i}}]$

$$\begin{aligned} \hat{\mathbf{z}}_j[k] &= \mathbf{H}_j \mathbf{A}|_{\tau_{j \rightarrow i}} \mathbf{H}_j^{-1} \mathbf{z}_j[k - k_{\tau_{j \rightarrow i}}] \\ &\quad + \mathbf{H}_j \mathbf{B}|_{\tau_{j \rightarrow i}} \mathbf{u}_j[k - k_{\tau_{j \rightarrow i}}] \end{aligned} \quad (5)$$

where $\mathbf{A}|_{\tau_{j \rightarrow i}}$ and $\mathbf{B}|_{\tau_{j \rightarrow i}}$ represent dynamic matrices with Δt replaced by $\tau_{j \rightarrow i}$, which can also be derived from physical dynamics. We assume $\mathbf{u}_j[k - k_{\tau_{j \rightarrow i}}] = \mathbf{u}_j[k]$ considering the smoothness of vehicle control in a short period.

The effect of delay compensation is briefly discussed as follows. As an estimate of the state $\mathbf{x}_j[k]$, $\hat{\mathbf{z}}_j[k]$ can be rewritten by substituting (1) and (2) into (5)

$$\hat{\mathbf{z}}_j[k] = \mathbf{x}_j[k] + \mathbf{A}|_{\tau_{j \rightarrow i}} \mathbf{v}_j[k - k_{\tau_{j \rightarrow i}}] - \mathbf{w}_j|_{\tau_{j \rightarrow i}}[k]. \quad (6)$$

The error consists of two parts: 1) the measurement error $\mathbf{v}_j[k - k_{\tau_{j \rightarrow i}}]$ which is amplified by $\mathbf{A}|_{\tau_{j \rightarrow i}}$ after the prediction using the dynamic model and 2) the error of the process noise $\mathbf{w}_j|_{\tau_{j \rightarrow i}}[k]$ with covariance $\mathbf{Q}|_{\tau_{j \rightarrow i}}$ corresponding to $\tau_{j \rightarrow i}$. Compared with the noise $\mathbf{v}_j[k]$ in the case without delay, even though compensation is applied, the variance of the error unavoidably increases. However, since the error remains zero-mean, the compensation effectively reduces the position drift.

Correspondingly, we can utilize similar delay compensation strategies in the m th RSU's positioning for V_i as follows:

$$\begin{aligned}\hat{\mathbf{z}}_{r_m,i}[k] &= \mathbf{A}|_{\tau_{r_m \rightarrow i}} \mathbf{z}_{r_m,i}[k - k_{\tau_{r_m \rightarrow i}}] \\ &\quad + \mathbf{B}|_{\tau_{r_m \rightarrow i}} \mathbf{u}_i[k - k_{\tau_{r_m \rightarrow i}}] \\ &= \mathbf{x}_i[k] + \mathbf{A}|_{\tau_{r_m \rightarrow i}} \mathbf{v}_{r_m \rightarrow i}[k - k_{\tau_{r_m \rightarrow i}}] \\ &\quad - \mathbf{w}_i|_{\tau_{r_m \rightarrow i}}[k].\end{aligned}\quad (7)$$

IV. DISTRIBUTED COOPERATIVE TRACKING FRAMEWORK VIA MULTICAST

A diagram of our proposed multicast-based cooperative mobility tracking framework is shown in Fig. 2, which contains local-fusion and global-filtering. The ego-vehicle is denoted as V_s and there are N other vehicles ($V_1, V_2, \dots, V_N$) cooperating in the IoV, which can be observed by V_s to measure the relative states $\mathbf{z}_{s \rightarrow 1}, \mathbf{z}_{s \rightarrow 2}, \dots, \mathbf{z}_{s \rightarrow N}$. Using V_s 's on-board sensor data, including commonly used traditional localization devices, such as IMU and GPS, V_s can obtain a self-measurement of its state $\mathbf{z}_s$. This self-measurement is then locally fused with measurements from the RSUs to obtain the fused self-measurement $\bar{\mathbf{z}}_s$ in Section IV-A. Then, V_s can send $\bar{\mathbf{z}}_s$ to other vehicles and receive their local-fusion results $\bar{\mathbf{z}}_1, \bar{\mathbf{z}}_2, \dots, \bar{\mathbf{z}}_N$ through multicast in Section IV-B. By matching and subtracting, local-fusion results of other vehicles can be combined with V_s 's observation toward them to generate a set of V_s 's state measurements in Section IV-C. Together with V_s 's locally fused self-measurement, they are fed into a global filtering algorithm for localization and tracking in Section IV-D. Such a process is done by each vehicle independently, which forms a distributed framework. The key variables in the proposed cooperative tracking algorithm are listed in Table I.

A. Local-Fusion and Delay Compensation

1) *Fusion of Self-Positioning and RSU*: Before sharing data with $V_1, V_2, \dots, V_N$, the ego-vehicle V_s will locally fuse its self-measurements from GPS, IMU, etc., with the auxiliary measurements from RSUs. Suppose that at time k , there are M RSUs within V_s 's communication range. The fused data involves $\mathbf{z}_s[k], \mathbf{z}_{r_1,s}[k], \mathbf{z}_{r_2,s}[k], \dots, \mathbf{z}_{r_M,s}[k]$, as indicated by (2) and (4). All of the mentioned data can be treated as Gaussian random vectors with a mean value equal to the true state $\mathbf{x}_s[k]$. Therefore, the local-fusion can be formulated as a data fusion problem under a linear Gaussian system, and the local-fusion result $\bar{\mathbf{z}}_s$ can be denoted as a linear combination

$$\bar{\mathbf{z}}_s = \sum_{m=1}^M \mathbf{A}_i \mathbf{z}_{r_m,s} + \mathbf{A}_s \mathbf{z}_s \quad (8)$$

where $\mathbf{A}_i$ and $\mathbf{A}_s$ are weights of RSUs and self measurement which satisfy $\sum_{m=1}^M \mathbf{A}_m + \mathbf{A}_s = \mathbf{I}$. The time parameter k is omitted because all the vectors are corresponding to the same k . The maximum likelihood estimate of $\bar{\mathbf{z}}_s$ is equivalent to minimizing the variance due to the Gaussian noise assumption. Therefore, the local fusion problem can be solved by utilizing the Lagrange multiplier method, which is identical to

the global filter in [12]. The optimal weights can be obtained as follows:

$$\begin{aligned}\mathbf{A}_s &= \mathbf{R}_s^{-1} \left(\sum_{m=1}^M \mathbf{R}_{r_m,s}^{-1} + \mathbf{R}_s^{-1} \right)^{-1} \\ \mathbf{A}_i &= \mathbf{R}_{r_m,s}^{-1} \left(\sum_{i=1}^M \mathbf{R}_{r_m,s}^{-1} + \mathbf{R}_s^{-1} \right)^{-1}.\end{aligned}\quad (9)$$

The optimized covariance matrix $\bar{\mathbf{R}}_s$ can be denoted as follows:

$$\bar{\mathbf{R}}_s = \left(\sum_{m=1}^M \mathbf{R}_{r_m,s}^{-1} + \mathbf{R}_s^{-1} \right)^{-1}. \quad (10)$$

2) *Delay Compensation in Local Fusion*: According to the analysis in Section III-C, when data from RSUs are transmitted to V_s , there might be a communication delay $\tau_{r_m \rightarrow s}$. This delay introduces a time misalignment, and as a result, we can no longer omit the time parameter k . Hence, (8) needs to be modified as follows:

$$\bar{\mathbf{z}}_s[k] = \sum_{m=1}^M \mathbf{A}_i \mathbf{z}_{r_m,s}[k - k_{\tau_{r_m \rightarrow s}}] + \mathbf{A}_s \mathbf{z}_s[k] \quad (11)$$

which results in bias in estimating $\mathbf{x}_s[k]$. It is essential to compensate for the communication delay in RSUs' measurements before the local-fusion process. This compensation can be done using the method in Section III-C

$$\begin{aligned}\hat{\mathbf{z}}_{r_m,s}[k] &= \mathbf{A}|_{\tau_{r_m \rightarrow s}} \mathbf{z}_{r_m,s}[k - k_{\tau_{r_m \rightarrow s}}] \\ &\quad + \mathbf{B}|_{\tau_{r_m \rightarrow s}} \mathbf{u}_s[k]\end{aligned}\quad (12)$$

and the covariance after the compensation is

$$\hat{\mathbf{R}}_{r_m,s} = \mathbf{A}|_{\tau_{r_m \rightarrow s}} \mathbf{R}_{r_m,s} \mathbf{A}^T|_{\tau_{r_m \rightarrow s}} + \mathbf{Q}|_{\tau_{r_m \rightarrow s}}. \quad (13)$$

B. Multicast and Observation

1) *Intervehicle Data Exchange and Observation*: After performing the local-fusion and delay compensation process, the ego vehicle V_s and other cooperating vehicles $V_1, V_2, \dots, V_N$ have obtained their locally fused results $\bar{\mathbf{z}}_s$ and $\bar{\mathbf{z}}_i$. Now the local-fusion data need to be shared among vehicles via the IoV. For V_s , at time k , a data package containing the local-fusion result $\bar{\mathbf{z}}_s$, covariance $\bar{\mathbf{R}}_s$, acceleration $\mathbf{u}_s$, and timestamp t_s is multicast to surrounding trusted vehicles $V_1, V_2, \dots, V_N$. Meanwhile, packages from surrounding vehicles are collected by the ego vehicle V_s to obtain $\bar{\mathbf{z}}_i, \bar{\mathbf{R}}_i, \mathbf{u}_i, t_i$.

To optimize the localization of V_s through the multicast information from V_i , V_s needs to obtain observations on them, which are expressed as measurement vectors $\mathbf{z}_{s \rightarrow i}[k]$ according to (3). By utilizing the positions multicast by the surrounding vehicles, the sensing results obtained from multiple sensors can be matched and fused, resulting in unique $\mathbf{z}_{s \rightarrow i}[k]$ vectors for each V_i .

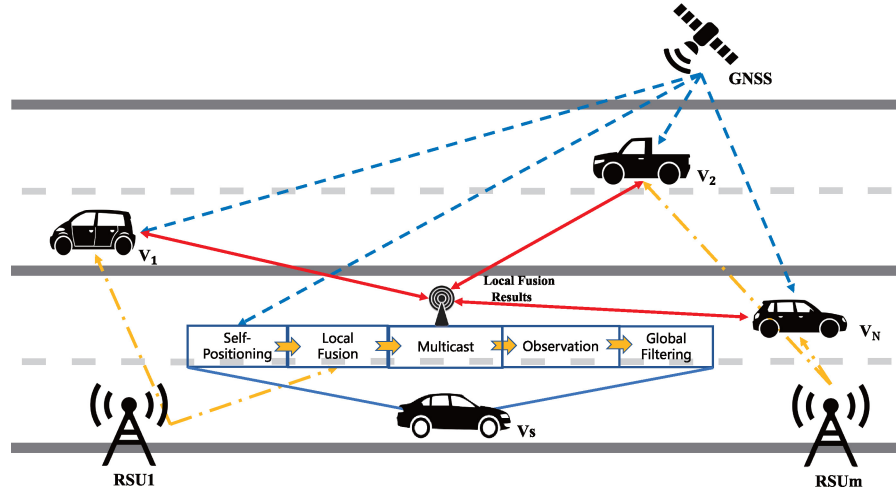


Fig. 2. Schematic of the multicast-based distributed cooperative tracking framework.

TABLE I
DEFINITIONS OF COOPERATIVE MOBILITY TRACKING KEY VARIABLES

Symbol	Definition
A, B_u	Dynamic model matrices
$k, k_{\tau_i \rightarrow j}, k_{\tau_{r_m} \rightarrow j}$	Discretized time and delay
u_i	Control vector of V_i
$x_i, x_{i \rightarrow j}$	The state of V_i and the state of V_j relative to i
$x_{r_m \rightarrow j}$	The state of V_j relative to RSU m
$z_i, z_{i \rightarrow j}$	Measurement of the state of V_i and the state of V_j relative to i
$z_{r_m \rightarrow j}$	Measurement of the state of V_j relative to RSU m
$\tilde{z}_i$	Local-fusion result of V_i 's state
$\tilde{\tilde{z}}_i$	Local-fusion result of V_i 's state after delay compensation
$z_{i,j}$	Measurement of V_j 's state utilizing V_i
$\tilde{\tilde{z}}_{i,j}$	Measurement of V_j 's state utilizing V_i after delay compensation
$z_{r_m,j}$	Measurement of V_j 's state utilizing RSU m
$\tilde{\tilde{z}}_{r_m,j}$	Measurement of V_j 's state utilizing RSU m after delay compensation
$z_{i,g}$	Global observation vector of V_i 's state
$\hat{x}_i$	Global filtering result of V_i

2) *Communication Load Analysis*: In our previous work [12], the ego vehicle V_s needs to estimate the states of surrounding vehicles, and then send them to the corresponding vehicles. Other vehicles' estimations of V_s 's state are also received by V_s for cooperative tracking. In such a framework, different data are sent to different vehicles, so point-to-point V2V links among vehicles are needed. When the number of cooperating vehicles is large, there will be a large communication load on direct V2V communication.

The communication load involved in cooperative tracking is analyzed quantitatively as follows. A state vector contains four numbers; the covariance matrix is 4×4 , and since it is a sparse block diagonal matrix, it can be denoted by six numbers; the acceleration vector contains two numbers; the timestamps contain two numbers. This results in a total of 14 numbers for each package involved in the data-sharing process. Assuming that the data type is float, then each number takes up 8 bytes. Given a package transmission rate of 10 Hz which is the typical data rate for vehicle on-board sensors, the data rate would be about $R = 1.5$ kB/s.

When there are N vehicles around V_s , in the previous work [12], V_s needs to establish N direct V2V links with surrounding vehicles. As different data are transmitted over different links, the data rate of sending and receiving are both NR , and the total load is $2NR$. If no direct V2V link is available, the base station is required to collect and distribute all the packages. As each of the $N+1$ vehicles sends and receives data both at a rate of NR , the base station has a total communication load of $2(N+1)NR$.

For the multicast-based method proposed in this article, V_s only needs to make a one-to- N multicast to trusted vehicles. The data being sent remains consistent regardless of the number of vehicles, meaning that the data sending rate is always R at V_s . Considering a receiving rate of NR , the total load is $(N+1)R$, which is 50% smaller than that in [12]. Moreover, if the multicast is also supported by the base station, each of the $N+1$ vehicles transmits data to the base station at a rate of R , and the base station multicasts at a rate of $(N+1)R$. As a result, the total communication load is $2(N+1)R$, which grows with N slower than that in [12].

C. Data Fusion and Delay Compensation

1) *Integration of Observations on the Ego-State via Other Vehicles:* In contrast to the method proposed in [12], where observations on V_s by $V_1, V_2, \dots, V_N$ are directly sent to V_s , the multicast-based method requires V_s to integrate the data received from other vehicles with the information obtained from its onboard sensors to construct its self-state observations.

At time instant k , when V_s receives the locally fused self-positioning data $\bar{z}_i$ from V_i , it can obtain the observation on its self-state via V_i as follows:

$$\begin{aligned} z_{i,s}[k] &= \bar{z}_i[k] - z_{s \rightarrow i}[k] \\ &= \mathbf{x}_i[k] + \bar{\mathbf{v}}_i[k] - \mathbf{H}_{s \rightarrow i} \mathbf{x}_{s \rightarrow i}[k] - \mathbf{v}_{s \rightarrow i}[k] \end{aligned} \quad (14)$$

where $\bar{\mathbf{v}}_i[k]$ is the error of V_i 's locally fused self-positioning and its covariance is $\bar{\mathbf{R}}_i$, according to (13). Substituting $\mathbf{x}_{s \rightarrow i}[k] = \mathbf{x}_i[k] - \mathbf{x}_s[k]$, (14) can be rewritten as follows:

$$z_{i,s}[k] = \mathbf{x}_s[k] + \bar{\mathbf{v}}_i[k] - \mathbf{v}_{s \rightarrow i}[k]. \quad (15)$$

The random errors $\bar{\mathbf{v}}_i[k]$ and $\mathbf{v}_{s \rightarrow i}[k]$ are independent and both AWGN, so the covariance of the error part in $z_{i,s}[k]$ is

$$\mathbf{R}_{i,s} = \bar{\mathbf{R}}_i + \mathbf{R}_{s \rightarrow i} \quad (16)$$

which contains the uncertainty of both the locally fused self-positioning of V_i and V_s 's observation on V_i .

From surrounding vehicles $V_1, V_2, \dots, V_N$, the ego vehicle V_s can obtain the observations on its self state $z_{1,s}[k], z_{2,s}[k], \dots, z_{N,s}[k]$, all of which are Gaussian vectors with a mean of V_s 's true state and independent errors. Together with V_s 's local-fusion result, these observations on the state of V_s by surrounding vehicles can be organized into a long global observation vector

$$\mathbf{z}_{s,g} = [\bar{z}_s \ z_{1,s} \ z_{2,s} \ \dots \ z_{N,s}]^T \quad (17)$$

where each subvector is an observation of V_s 's state. Correspondingly, since subvectors are independent, the covariance is a block diagonal matrix

$$\mathbf{R}_{s,g} = \begin{bmatrix} \bar{\mathbf{R}}_s & 0 & \dots & 0 \\ 0 & \mathbf{R}_{1,s} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{R}_{N,s} \end{bmatrix}. \quad (18)$$

2) *Time Alignment of Data:* Multicast collisions are usually managed by interframe interval and waiting mechanisms. As a result, there might be a random delay $\tau_{i \rightarrow s}$ when V_s receives data from V_i . Assuming that each vehicle attempts to multicast its local-fusion data at time k and the actual arrival time is $k + k_{\tau_{i \rightarrow s}}$. Then, V_s matches and subtracts them with the intervehicle observation results at $k + k_{\tau_{i \rightarrow s}}$. So, (14) can be rewritten as follows:

$$\begin{aligned} z_{i,s}[k + k_{\tau_{i \rightarrow s}}] &= \mathbf{x}_i[k] + \bar{\mathbf{v}}_i[k] - \mathbf{H}_{s \rightarrow i} \mathbf{x}_{s \rightarrow i}[k + k_{\tau_{i \rightarrow s}}] \\ &\quad - \mathbf{v}_{s \rightarrow i}[k + k_{\tau_{i \rightarrow s}}]. \end{aligned} \quad (19)$$

Because of the motion of V_i , $\mathbf{x}_i[k] \neq \mathbf{x}_i[k + k_{\tau_{i \rightarrow s}}]$ introduces bias in the result. Thus, $\mathbf{z}_{s,g}$ should be rewritten as follows:

$$\mathbf{z}_{s,g}[k] = \begin{bmatrix} \bar{z}_s[k], \\ z_{1,s}[k + k_{\tau_{1 \rightarrow s}}], \\ z_{2,s}[k + k_{\tau_{2 \rightarrow s}}], \\ \dots \\ z_{N,s}[k + k_{\tau_{N \rightarrow s}}] \end{bmatrix}. \quad (20)$$

Assume that delay $\tau_{i \rightarrow s}$ from different vehicles are bounded by $\tau_m = \max_{1 \leq i \leq N} \{\tau_{i \rightarrow s}\}$. By the time instant $k + k_{\tau_m}$, V_s has received all the packages and the relative state between V_s and V_i can be obtained from sensing devices. According to the dynamic model, the delay can be compensated as follows:

$$\hat{\mathbf{z}}_i[k + k_{\tau_m}] = \mathbf{A}|_{\tau_m} \bar{z}_i[k] + \mathbf{B}|_{\tau_m} \mathbf{u}_i[k] \quad (21)$$

where the acceleration vector $\mathbf{u}_i[k]$ of V_i can be easily obtained from the data package. Based on the compensation, (14) can be rewritten as follows:

$$\hat{z}_{i,s}[k + k_{\tau_m}] = \hat{\mathbf{z}}_i[k + k_{\tau_m}] - z_{s \rightarrow i}[k + k_{\tau_m}] \quad (22)$$

and the compensated covariance is

$$\begin{aligned} \hat{\mathbf{R}}_{i,s}[k + k_{\tau_m}] &= \mathbf{A}|_{\tau_m} \bar{\mathbf{R}}_i[k] \mathbf{A}^T|_{\tau_m} \\ &\quad + \mathbf{Q}|_{\tau_m} + \mathbf{R}_{s \rightarrow i}[k + k_{\tau_m}]. \end{aligned} \quad (23)$$

After compensating for the delays from V_1 to V_N , the global observation vector $\mathbf{z}_{s,g}$, can be constructed as follows:

$$\mathbf{z}_{s,g} = [\hat{z}_s \ \hat{z}_{1,s} \ \hat{z}_{2,s} \ \dots \ \hat{z}_{N,s}]^T \quad (24)$$

and the covariance matrix is

$$\mathbf{R}_{s,g} = \begin{bmatrix} \hat{\mathbf{R}}_s & 0 & \dots & 0 \\ 0 & \hat{\mathbf{R}}_{1,s} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\mathbf{R}}_{N,s} \end{bmatrix}. \quad (25)$$

Comparing with (17) and (18), time corresponding to $\mathbf{z}_{s,g}$ is $k + k_{\tau_m}$, and the exact value of τ_m is determined by the communication condition.

In communication environments with packet loss due to multicast collisions, if the data from V_i at a particular time instant k fail to be received by V_s , V_s will temporally exclude V_i from the list of cooperative vehicles until subsequent data are received. This situation can be modeled as a temporary reduction in the number of cooperators and the dimension reduction of the global observation vector $\mathbf{z}_{s,g}$. In the worst case of communication interruption, only the local-fusion result of V_s is available and $\mathbf{z}_{s,g} = \hat{\mathbf{z}}_s$.

D. Global Dynamic Filtering

1) *Algorithm Description:* After the steps above, we can achieve "tracking" beyond "localization" by introducing the dynamic model to fusion and smooth multisource data. In this article, we apply the Kalman filter to this linear model for simplicity and conduct theoretical performance analysis.

Based on the system model in Section III-A and the Kalman prediction equations [26], the prediction process can be described as follows:

$$\begin{aligned}\hat{\mathbf{x}}'[k+1] &= \mathbf{A}\hat{\mathbf{x}}[k] + \mathbf{B}\mathbf{u}[k+1] \\ \mathbf{P}'[k+1] &= \mathbf{A}\mathbf{P}[k]\mathbf{A}^T + \mathbf{Q}\end{aligned}\quad (26)$$

where $\hat{\mathbf{x}}'[k+1]$ is the prediction of time $k+1$ based on the state estimation $\hat{\mathbf{x}}[k]$ at time k and $\mathbf{P}'[k+1]$ denotes the prediction variance.

In our framework, the global observation vector $\mathbf{z}_{s,g}$ can be considered as a sensor that can give N independent direct measurements of the state of V_s simultaneously. The covariance is block diagonal, as described in (18). Following (2), the measurement matrix $\mathbf{H}$ in the measurement equation of the dynamic model is

$$\mathbf{H}_{4N \times 4} = [\mathbf{I}_{4 \times 4} \ \mathbf{I}_{4 \times 4} \ \cdots \ \mathbf{I}_{4 \times 4}]^T. \quad (27)$$

Consequently, the linear combination of the observation and the prediction according to the Kalman update equations can be written as follows:

$$\hat{\mathbf{x}}[k+1] = \hat{\mathbf{x}}'[k+1] + \mathbf{K}(\mathbf{z}_{s,g}[k+1] - \mathbf{H}\hat{\mathbf{x}}'[k+1]) \quad (28)$$

where $\mathbf{K}$ is the Kalman Gain which is a $4 \times 4N$ linear combination coefficient matrix of $\mathbf{z}_{s,g}$ and $\hat{\mathbf{x}}'_{k+1}$ for our model. The Kalman Gain can be calculated by

$$\mathbf{K} = \mathbf{P}'[k+1]\mathbf{H}^T(\mathbf{H}\mathbf{P}'[k+1]\mathbf{H}^T + \mathbf{R}_{s,g})^{-1}. \quad (29)$$

The covariance $\mathbf{P}_{k+1}$ of $\hat{\mathbf{x}}_{k+1}$ can be obtained by the Kalman variance updating equation

$$\mathbf{P}[k+1] = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}'[k+1]. \quad (30)$$

In summary, the Kalman filter operates based on (26)–(30) to continuously track the state of the ego vehicle V_s . Due to the property of Kalman filter, the variance will eventually converge, resulting in continuous and effective tracking. The algorithm is summarized in Algorithm 1.

2) *Comparison With Our Previous Work [12]*: Compared to the approach proposed in [12, Algorithm 2], the framework presented in this article optimizes the fusion procedure to attain zero asymptotic error, decrease the number of Kalman filters, and enhance flexibility in complex scenarios. The corresponding analysis is shown as follows. In [12], the proposed approach adopts a “local filtering and global filtering” two-layer structure. First, V_s observes surrounding vehicles $V_1, V_2, \dots, V_N$ and gets the measurement of their states based on relative observation $\mathbf{z}_{s \rightarrow i}$ and self-positioning $\mathbf{z}_s$. Subsequently, V_s sends $\mathbf{z}_{s \rightarrow i} + \mathbf{z}_s$ to surrounding vehicles and receives measurements $\mathbf{z}_{i \rightarrow s} + \mathbf{z}_i$ from them. Then, $\mathbf{z}_s$ and $\mathbf{z}_{i \rightarrow s} + \mathbf{z}_i$ from other vehicles are fed into $N+1$ local Kalman filters separately as shown in [12, Algorithm 1]. Finally, a global filter based on maximum likelihood estimation is employed to fuse the outputs of local filters.

This method utilizes information from multiple vehicles to optimize location and tracking. However, in the final step, the optimality of the global filtering depends on the assumption of error independence among the outputs of each local filter. However, due to the repeated Kalman filtering on the same vehicle trajectory, correlated process noise is introduced into the system. A detailed analysis can be found in Appendix A. The effect of correlated noise is illustrated in [12, Fig. 7].

Algorithm 1 Multicast-Based Cooperative Mobility Tracking Flow

```

1: Initialize: the estimation value  $\hat{\mathbf{x}}_0$ , and its covariance  $\mathbf{P}_0$ ;
2: for  $k = 1:T$  do
3:   Self-positioning: Obtain  $\mathbf{z}_s$  and its covariance  $\mathbf{R}_s$  from GPS,
     acceleration  $\mathbf{u}_s$  from IMU;
4:   for  $m = 1:M$  do
5:     RSU auxiliary: Receive  $\mathbf{z}_{r_m,s}$  and its covariance  $\mathbf{R}_{r_m,s}$ 
       from  $m$ th RSU and perform delay compensation;
6:   end for
7:   Local Fusion: Fuse  $\mathbf{z}_s$  and  $\mathbf{z}_{r_m,s}$  ( $m = 1, 2, \dots, M$ ) locally;
8:   Data sending: Multicast data package that consists of
      $\bar{\mathbf{z}}_s, \bar{\mathbf{R}}_s, \mathbf{u}_s$  and timestamp;
9:   for  $i = 1:N$  do
10:    Data receiving: Receive  $\bar{\mathbf{z}}_i, \bar{\mathbf{R}}_i$  and  $\mathbf{u}_i$  from  $V_i$  and
      perform delay compensation;
11:    Sensing: Observe  $V_i$  to obtain  $\mathbf{z}_{s \rightarrow i}$  and calculate  $\mathbf{z}_{i,s} =$ 
       $\bar{\mathbf{z}}_i - \mathbf{z}_{s \rightarrow i}$ ;
12:    end for
13:    Fusion: Align  $\bar{\mathbf{z}}_s$  and  $\mathbf{z}_{i,s}$  ( $i = 1, 2, \dots, N$ ) and integrate into
       $\mathbf{z}_{s,g}$  and  $\mathbf{R}_{s,g}$ ;
14:    Global Filtering:  $(\hat{\mathbf{x}}[k], \mathbf{P}[k]) = KF(\mathbf{z}_{s,g}, \mathbf{R}_{s,g}, \mathbf{u}_s, \hat{\mathbf{x}}[k-1], \mathbf{P}[k-1])$ ;
15:  end for

```

From the figure, it can be observed that as the number of vehicles increases, there is diminishing marginal utility, and further increase in the number of vehicles leads to the convergence of the RMSE to a nonzero value, in contrast to the zero asymptotic error achieved under the assumption of independent noise. Due to this performance bottleneck, the method proposed in [12] is not suitable for large-scale cooperation or high-precision sensor fusion.

In this article, according to (28), observation from different sources and prediction are jointly fused by global optimal weights. In other words, there is only one global filter for V_s 's tracking. In addition to avoiding the cumulative error, it also reduces the calculation and storage overhead caused by extra filters. Moreover, due to packet loss or even connection loss, the number of cooperative vehicles can be time varying. In the MSMV framework in [12], when some vehicle V_i is disconnected, the local filter corresponding to it cannot continue to work. When a vehicle joins the cooperation, a new filter needs to be allocated. Hence, the framework is sensitive to a time-varying number of cooperative vehicles. However, in this article, when some vehicles are disconnected, only the dimension of the integrated data vector changes and the filter still stably works, which endows the system with stronger stability and flexibility.

E. Theoretical Performance Analysis

In this section, we conduct a theoretical performance analysis of the proposed algorithm. According to Section IV-D1, based on linear assumption, the global filter can fuse N independent measurement vectors and a prediction vector to reach the unique global optimum. In the dynamic model, N different vectors can be treated equally as N direct observations on the vehicle's state so they can be replaced by a single measurement vector $\mathbf{z}_g$ with covariance $\mathbf{R}_g$. In other words, the

TABLE II
SIMULATION PARAMETERS

Discrete time step	0.1 [s]
Duration of simulation	20 [s]
Self-positioning accuracy	0.7 [m]/[m/s]
Sensing accuracy	0.3 [m]/[m/s]
RSU accuracy	0.15 [m]/[m/s]
Process noise	0.05 [m]/[m/s]

global filter is equivalent to a two-layer structure where measurements from various sources are combined linearly into z_g before being fused with the prediction. It means that $z_{s,g}$ and $R_{s,g}$ in (28) to (30) can be substituted with z_g and R_g . The specific expressions for them can be found in Appendix B. Based on the aforementioned substitutions, we establish the following result.

Theorem 1: The steady-state variance P will be $A^{-1}(P' - Q)(A^T)^{-1}$, where P' is the solution of $AP'A^T - P' - AP'[P' + R_g]^{-1}P'A^T + Q = 0$

Proof: See Appendix B. ■

Theoretical analysis suggests that the accuracy of the cooperative tracking framework depends on the vehicle number in the system, the road environment, and the sensors' performance. In practice, since the error is not strictly AWGN as in (1) and the performance of the sensor may not remain stationary in complex road environments. Therefore, the actual performance fluctuates around this theoretical value.

V. SIMULATION RESULTS

In this section, multiple simulations are performed to verify the performance of the cooperative tracking framework in various scenarios. Considering the actual accuracy of sensors and the requirements of localization, simulation parameters are listed in Table II. Given the high positioning accuracy requirements of intelligent vehicles, the sensor parameters chosen for the simulations possess relatively high performance capabilities while keeping costs within reasonable limits.

The simulations are organized as follows.

- 1) The performance of the algorithm under ideal communications, including the trajectory and RMSE, is shown in Section V-A.
- 2) The impact of delay, the effectiveness of delay compensation, and the impact of packet loss are mentioned in Section V-B.
- 3) The robustness of the algorithm in adverse conditions, including disturbed self-positioning and intervehicle observation failure, is shown in Section V-C.
- 4) The extension of RSU range due to vehicle cooperation is discussed in Section V-D.

A. Accuracy of the Algorithm Under Ideal Communications

First, the performance of cooperative localization and tracking under ideal communication is shown. To intuitively reflect the performance improvement, the trajectory of a simple two-vehicle scenario is presented in Fig. 3(a).

The x -axis represents the horizontal direction, while the y -axis represents the vertical direction, illustrating the trajectory

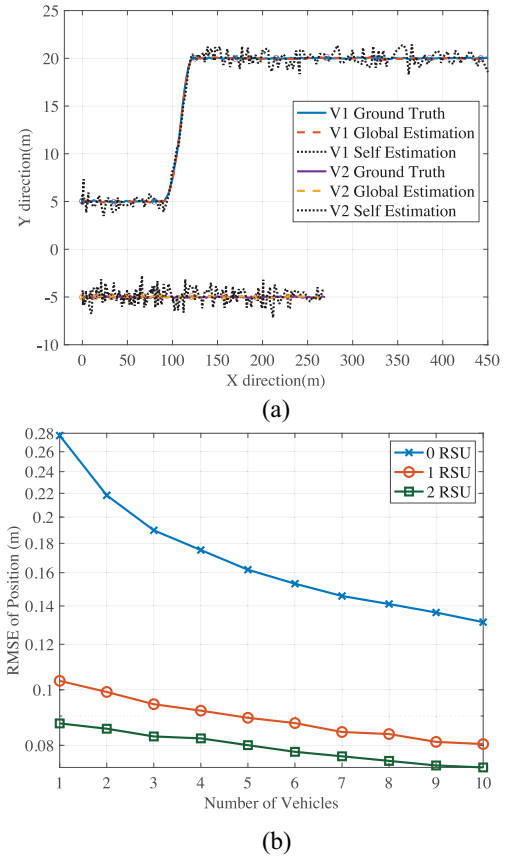


Fig. 3. Intuitive and quantitative result of cooperative tracking performance. (a) Trajectory of two vehicles. (b) Position RMSE of multiple vehicles.

directly. There are two vehicles and an RSU connected in the V2X network. One vehicle remains driving straight and the other changes lane in the y direction. RSU continuously provides auxiliary positioning for both vehicles. The dotted lines represent the self-positioning of the vehicles, which have relatively large errors when compared to the ground truth shown in solid lines. The dashed lines indicate the global estimates of the trajectory. From the figure, it is evident that cooperative localization and tracking effectively reduce estimation errors and result in a smoother trajectory. A comparison of tracking performance in systems with different numbers of vehicles and RSUs is presented by RMSE in Fig. 3(b). We take V_1 as the ego vehicle and evaluate its RMSE. The trajectory trend of other vehicles can be simulated similarly to V_2 , resembling the traffic flow moving in the same direction on the road.

Fig. 3(b) illustrates that a single vehicle without RSU has an RMSE of 0.277 m. This level of accuracy only allows for lane-level localization assistance and is insufficient to support active control of intelligent vehicles. Without the assistance of RSU, when the number of vehicles in cooperation reaches 5 and 10, RMSEs are 0.162 and 0.131, respectively, which are 41.5% and 52.7% smaller than the single-vehicle case. Therefore, the performance is notably improved through multivehicle cooperation. However, since the fused information among vehicles includes both self-positioning and observation errors, the RMSE of the cooperative localization still falls short of achieving centimeter-level precision even with

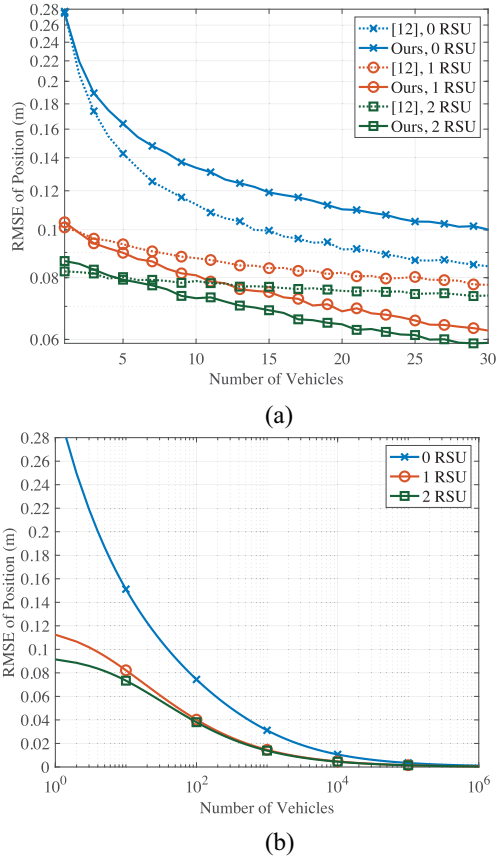


Fig. 4. Effect of eliminating error correlations. (a) Simulation RMSE comparison with [12]. (b) Theoretical asymptotic RMSE.

10 vehicles. Moreover, considering the diminishing marginal utility, the performance gains from having 1 to 5 vehicles are much more substantial than the gains obtained from 5 to 10 vehicles. This suggests that additional vehicles would yield only limited improvement while increasing the communication and computing load. Hence, it becomes necessary to enhance accuracy by leveraging the assistance of an RSU.

When one RSU is incorporated into the framework and the numbers of vehicles are 1, 5, and 10, the RMSE are, respectively, 0.104, 0.089, and 0.080. Comparing these results with the same number of vehicles without an RSU, we observe an improvement of 62.5%, 45.1%, and 38.9%. With an increase in the number of RSUs to 2, the percentage improvements are 68.5%, 50.6%, and 44.3%. The significant enhancement in tracking performance can be attributed to the assistance provided by RSUs, as they do not have self-positioning errors. From the data shown in Fig. 3(b), we see that the centimeter-level precision becomes easily attainable through cooperative tracking with either one vehicle and two RSUs or two vehicles with one RSU.

According to our analysis, due to the presence of process noise w , data provided by different vehicles in our previous work [12] may consist of correlated errors. So, the performance gains will be saturated as the number of vehicles increases. A comparison is shown in Fig. 4(a). It can be seen that due to the two-layer filters involved, the method in [12] does have an overall advantage of around 15% in

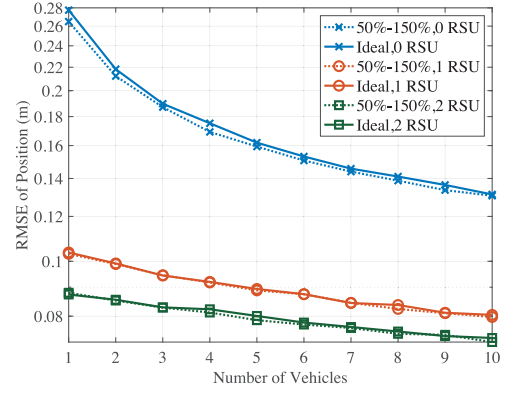


Fig. 5. RMSE of cooperative localization in random accuracy case with floating variance.

the case where there is no RSU. However, when introducing high-precision RSU and more vehicles in the cooperation, the saturation in the performance gains shows up as illustrated in the circled and squared dash curves. However, the framework in this article always maintains a performance gain with the number of cooperating vehicles, especially in the cases with RSUs, as shown in the solid curves. The zero asymptotic error when the vehicle number $N \rightarrow \infty$ is also shown in Fig. 4(b). As the simulation is intractable when N is large, only theoretical calculation is presented.

Simulations with random accuracy are also conducted to account for the variable performance of localization and sensing modules on intelligent vehicles in practical scenarios. In Fig. 5, the variance of position and velocity measurement in self-positioning, positioning from RSUs, and intervehicle positioning are independently sampled from the uniform distributions from 50% to 150% of the values in Table II. The randomness results in different RMSE values compared to the ideal case depicted in Fig. 3(b). However, despite these variations, the performance gains obtained from cooperation remain approximately the same. This demonstrates the practical applicability of the framework in practical sensor scenarios.

B. Effects of Communication Delay, Compensation, and Packet Loss

The simulations above are based on ideal communications, but in practice, there would always be communication delays and packet loss. According to the previous analysis, the communication delay may reduce the cooperation gains and even make the performance worse than single-vehicle tracking. In our simulation, the delay follows a uniform distribution from 5 to 35 ms, and the impact of delay is introduced by adding drift due to vehicle motion according to the model in Section III-C.

As shown in Fig. 6(a), in the case of no RSU, as the number of vehicles increases, more delayed data are involved in the global filtering of the ego vehicle and the performance deteriorates as shown in the gap between the solid and the dotted crossed curves. Specifically, in the cases of five and ten vehicles, compared with the ideal case without delay, the RMSE is increased by 31.3% and 57.9%, respectively. Furthermore, the

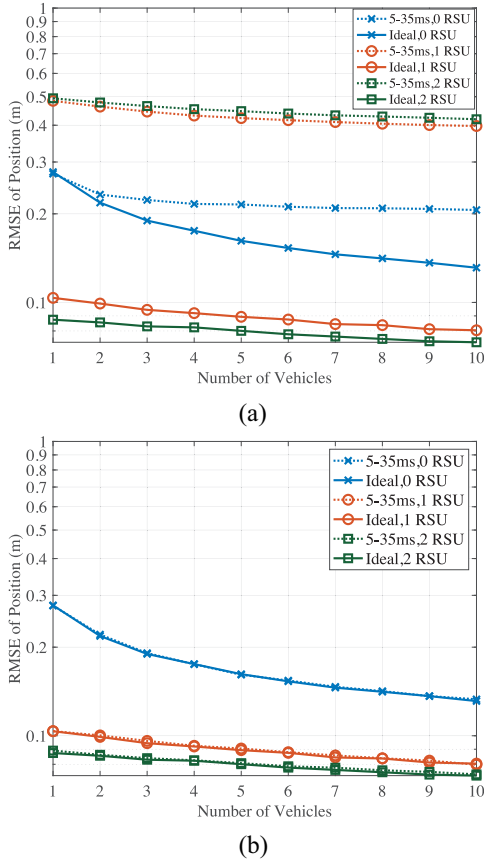


Fig. 6. Performance loss and compensation effects in 5–35 ms delay and slow-moving surrounding vehicles. (a) Without compensation. (b) With compensation.

misleading effects of delayed data from RSU are even more significant. The RMSE in the case of one RSU is increased by 376% to 394% as compared with the ideal case without delay. Indeed, the higher weight of RSU data in the fusion process can make the overall performance more sensitive to communication delays.

As the position drift is caused by the vehicles' motion during the delay, the vehicles' velocity may significantly influence the RMSE. In Fig. 7(a), the speed of surrounding vehicles is 29 m/s, which is larger than 9 m/s in Fig. 6(a). Otherwise, the speed of the ego vehicle is 6.5 m/s, which is smaller than 24.6 m/s in Fig. 6(a). It causes a smaller motion of the ego vehicle and a larger motion of the surrounding vehicles during the delay. When only RSU data are involved, the localization drift is only determined by the motion of the ego vehicle. So, the impact of RSU in Fig. 7(a) is not as much as that in Fig. 6(a). When the number of surrounding vehicles increases, their motion gradually influences the global filtering of the ego vehicles. So, the impact of more surrounding vehicles in Fig. 7(a) is much larger than that in Fig. 6(a).

The result of delay compensation is shown in Figs. 6(b) and 7(b). After compensation, RMSE is only increased by no more than 1% compared to the case without delay and the additional error solely comes from a random error in compensation. As the delay is relatively short as compared with the sampling period of the dynamic model (5–35 ms delay versus 100-ms sampling period), the effect of compensation is quite

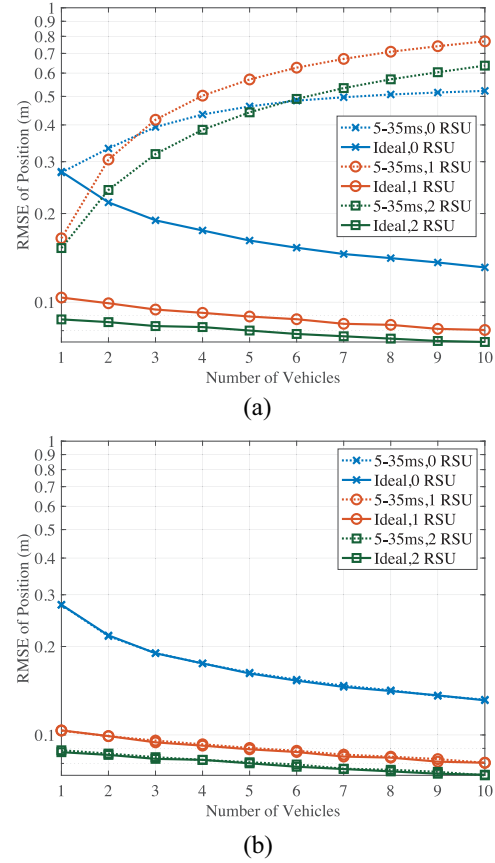


Fig. 7. Performance loss and compensation effects in 5–35 ms delay and fast-moving surrounding vehicles. (a) Without compensation. (b) With compensation.

well. Therefore, the proposed delay compensation scheme is effective in mitigating the effects of delay and maintaining the overall performance.

In the case of packet loss, data from the surrounding vehicles are randomly disabled with a certain probability to simulate packet loss. The simulation result is shown in Fig. 8. When the packet loss rate is 10%, the overall RMSE increases by about 3% to 5%, from zero-RSU case to two-RSU case. The blue curve also indicates that the RMSE in ten-vehicle cooperation with 10% packet loss is the same as that in the nine-vehicle case with ideal communication. This indicates that the packet loss problem in our proposed framework is equivalent to the temporal reduction of cooperators.

C. Robustness of the Cooperative Tracking Algorithm Under Adverse Conditions

In addition to accuracy, robustness is also a crucial factor in achieving stable localization and tracking in a cooperative system, especially when GPS deteriorates or even fails due to signal loss. When self-positioning is disturbed, positioning accuracy could be maintained at an acceptable level thanks to the cooperation with surrounding vehicles and RSUs as shown in Fig. 9(a). Suppose that the ego vehicle's self-positioning variance is ten times larger than normal due to harsh conditions, while the mutual communication and observation with surrounding vehicles can be performed normally.

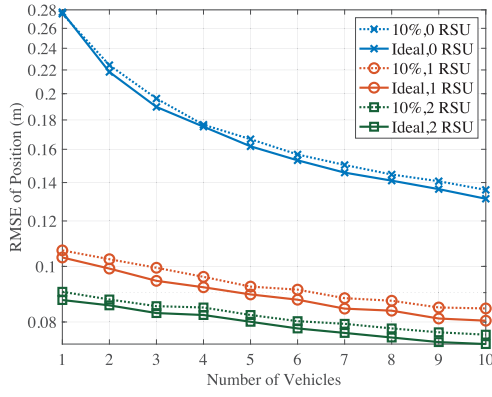


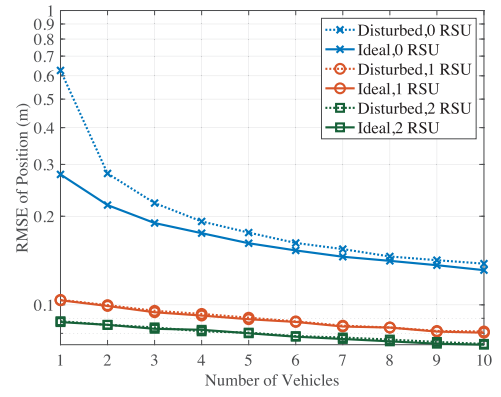
Fig. 8. Effect of 10% packet loss rate on cooperative localization performance.

Through cooperation with 0–9 vehicles and 0–2 RSUs, the robustness of tracking is significantly improved. In the single-vehicle scenario, when the GPS is disturbed (as shown in the blue dashed curves), RMSE increased from 0.277 to 0.626, representing an increase of 126%. This substantial increase indicates the unsatisfactory robustness of a single vehicle. However, through cooperation, in the case of five and ten vehicles, the RMSE increases due to the deterioration of the GPS signal are only 8% and 5.3%, respectively. When there is RSU, due to the high precision of RSUs' self-positioning, the RMSE increase is only about 1%. It shows again that the RSU plays an important role in improving vehicle localization and tracking performance.

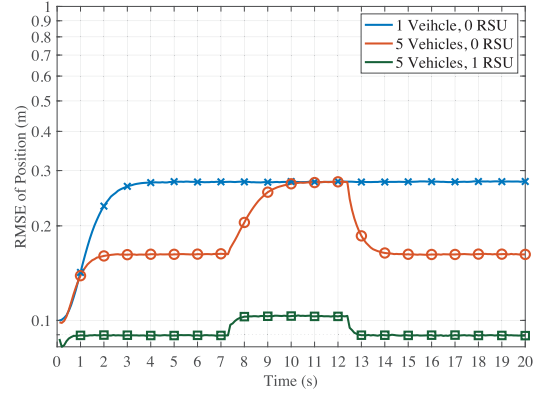
Intervehicle observation is crucial for obtaining relative states with surrounding vehicles and achieving vehicle–vehicle cooperation. However, temporary relative observation failures can occur due to factors, such as heavy occlusion or sensor failures. Therefore, it is essential to evaluate the tracking performance under such observation failures to ensure robustness. As shown in Fig. 9(b), in the road section where an RSU is deployed, five vehicles drive for 20 s. Suppose that between 7.5 and 12.5 s, the ego vehicle experiences a failure in observing the surrounding four vehicles, leading to a lack of cooperation during that period. However, the five-vehicle cooperative tracking RMSE (shown as the orange circled curve) still keeps 1.41% smaller than the single-vehicle tracking (shown as the blue crossed curve) at 10 s, which is 2.5 s after the observation fails. Once the relative observation is restored, the ego vehicle promptly utilizes the observation data and the RMSE recovers in approximately 1.5 s. Furthermore, when one RSU is available to assist (as shown in the green squared curve), even during the period of 7.5 and 12.5 s, the RMSE remains 62.7% smaller than the single-vehicle case. In general, if a vehicle experiences relative observation failure, it can still maintain a certain performance gain through self-positioning and assistance from RSUs, while the tracking of other vehicles in the framework remains unaffected.

D. Extension of RSU Range With Cooperation

From the simulations above, we have seen the importance of RSUs in improving tracking accuracy and robustness.



(a)



(b)

Fig. 9. Improvement of robustness due to cooperative tracking. (a) RMSE in the adverse self-positioning scenario. (b) RMSE fluctuation in temporary intervehicle observation failure.

However, in practical transportation systems, the communication and perception ranges of RSUs, as well as the number of RSUs are limited. It is impractical to cover the entire transportation system by RSUs. Therefore, the RSU cannot replace single-vehicle localization or vehicle–vehicle cooperation. On the other hand, the gains of RSU's assistance can be transferred among vehicles by V2V multicast as depicted in Fig. 10.

In the simulation, due to the limited coverage of the RSU, only V_1 (shown as the red circled curve) can obtain assistance from it. However, when V_1 shares the high-precision localization result, it can be a better localization reference for other vehicles. Even without the direct help of the RSU, the cooperative localization accuracy of other vehicles (shown as the green squared curve) can be higher than the no-RSU control group (shown as the blue crossed curve). In other words, V2R cooperation can be extended by V2V cooperation.

When there is only one other vehicle apart from V_1 in the system, compared with the control group, the RMSE of it after cooperation is reduced by 30%, although it is 49.3% larger than the RMSE of V_1 . Therefore, the gains of the RSU are equivalently extended by V_1 . However, when there are more vehicles, e.g., in the ten vehicles case, the RMSE of other vehicles is only reduced by 7.67% compared with the control group. This indicates the capability of vehicles to extend RSU gains is limited, causing the blue crossed curve and green

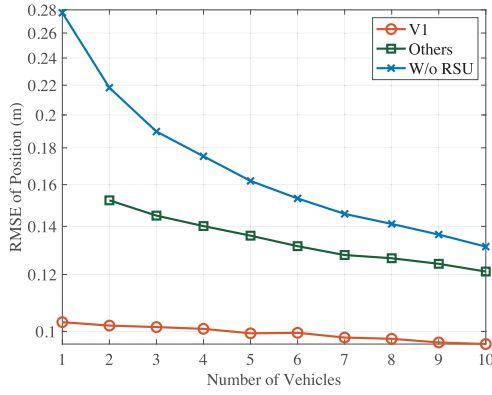


Fig. 10. Equivalent extension of RSU range due to intervehicle cooperation.

squared curve to converge as the number of vehicles increases. This demonstrates that, even with the range extension effect of vehicle-vehicle cooperation, the coverage provided by RSUs remains significant for a multivehicle system.

VI. CONCLUSION AND FUTURE WORK

This article presents an efficient distributed cooperative tracking framework via multicast to enhance the localization and tracking performance of intelligent vehicles while reducing the communication load of the IoV. Compared with the previous work, our framework simplifies the communication content and optimizes the data fusion process. By utilizing multicast, we can reduce the communication load and eliminate the need for complex direct V2V communication. Through the two-layer structure involving local fusion and global filtering, the proposed framework reduces the number of filters and has zero asymptotic error. This framework is designed to be adaptable to practical scenarios with limited communication resources, random communication delay, packet loss, and a varying number of cooperative vehicles. This article also provides theoretical analysis and simulation results on the localization performance and shows the applicability of the framework in the practical communication environment.

Based on the research in this article, future research can explore the following areas. In real-world environments, there are noncooperative objects, such as traditional vehicles and pedestrians. The perception of such targets can also be improved through intervehicle cooperation to address the problem of blinding areas and achieve perception beyond the vision range. Furthermore, with the support of high-precision tracking and perception, optimal path planning can also be made to solve the problem of traffic congestion and accidents caused by the wrong decisions of drivers and improve the safety and efficiency of ITS.

APPENDIX A

ERROR CORRELATION ISSUE IN [12]

According to [12, Algorithm 2], after V_s receive measurements from V_i at time instant $k+1$, the measurement $z_{i \rightarrow s} + z_i$ is fed into a local Kalman filter as [12, Algorithm 1]

$$\begin{aligned}\hat{x}_i[k+1] &= \hat{x}'_i[k+1] \\ &\quad + K(z_{i \rightarrow s}[k+1] + z_i[k+1] - H\hat{x}'_i[k+1]) \\ \hat{x}'_i[k+1] &= A\hat{x}_g[k] + Bu_s[k+1]\end{aligned}\quad (31)$$

where $\hat{x}'_i[k+1]$ is the prediction based on the previous output $\hat{x}_g[k]$ of global filter. The prediction is fused with measurement $z_{i \rightarrow s} + z_i$ by the Kalman Gain. The $\hat{x}_i[k+1]$ is then fed into the global filter with the outputs of other local filters.

As the output of each local filter is an estimation with Gaussian error. Global filter in [12] is a data fusion problem in a linear Gaussian system. Assuming that the errors in the outputs of all the local filters are independent, there is an optimal linear combination to minimize the variance of the global estimation $\mathbf{x}_g$.

However, due to the state transfer (1), errors in different local filters' outputs are partially correlated. Taking local filter for V_i as an example, error of $\hat{x}_i[k+1]$ comes from the measurement $z_{i \rightarrow s}[k+1] + z_i[k+1]$ and the prediction $\hat{x}'_i[k+1]$. The error in the observation is caused by the inaccuracies of the sensors

$$\begin{aligned}z_{i \rightarrow s}[k+1] + z_i[k+1] &= \mathbf{x}_s[k+1] \\ &\quad + \mathbf{v}_{i \rightarrow s}[k+1] + \mathbf{v}_i[k+1]\end{aligned}\quad (32)$$

where $\mathbf{v}_{i \rightarrow s}[k+1]$ and $\mathbf{v}_i[k+1]$ are errors in relative observation and V_i 's self-positioning. As for the prediction $\hat{x}'_i[k+1]$, according to (1), we can obtain

$$\begin{aligned}\hat{x}'_i[k+1] &= A\hat{x}_g[k] + Bu_s[k+1] \\ &= A(\mathbf{x}_s[k] + \mathbf{v}_g[k]) + Bu_s[k+1] \\ &= A\mathbf{x}_s[k] + Bu_s[k+1] + A\mathbf{v}_g[k] \\ &= \mathbf{x}_s[k+1] - \mathbf{w}_s[k+1] + A\mathbf{v}_g[k]\end{aligned}\quad (33)$$

where $\mathbf{v}_g[k]$ is the error in the global estimation of time instant k . Bring (32) and (33) into (31), $\hat{x}_i[k+1]$ can be denoted as follows:

$$\begin{aligned}\hat{x}_i[k+1] &= \mathbf{x}_s[k+1] + (\mathbf{I} - K)(-\mathbf{w}_s[k+1] + A\mathbf{v}_g[k]) \\ &\quad + K(\mathbf{v}_{i \rightarrow s}[k+1] + \mathbf{v}_i[k+1]).\end{aligned}\quad (34)$$

Since observations from different vehicles are independent of each other, $\mathbf{v}_{i \rightarrow s}[k+1] + \mathbf{v}_i[k+1]$ from different V_i is naturally independent. On the contrary, $-\mathbf{w}_s[k+1] + A\mathbf{v}_g[k]$ is derived from V_s itself and therefore remains the same for different local filters.

The correlated part $-\mathbf{w}_s[k+1] + A\mathbf{v}_g[k]$ will be brought into the output of global filter $\mathbf{x}_g[k+1]$, which will be used for local filters at $k+2$. The value of this part of the error remains constant as the number of vehicles N increases and converges to a nonzero asymptotic error as N approaches infinity. In a typical road environment, the process noise is usually much smaller than the sensor noise. However, as the number of cooperative vehicles gradually increases and the global localization accuracy gradually improves, this portion of the noise can no longer be ignored.

APPENDIX B

PROOF OF THEOREM 1

Suppose V_s is surrounded by N trusted vehicles V_i . The variance of self-positioning for V_s is $\mathbf{R}_s$, while for V_i , it

is R_i . The variance of relative observation from V_s to V_i is $R_{s \rightarrow i}$. Additionally, M RSUs provide V_s and V_i with auxiliary positioning, each having variances $R_{r,s}$ and $R_{r,i}$, respectively. We denote the communication delay as τ and the variance of the process noise as Q . As the filtering algorithm rapidly converges and sensor performance remains stable over a short period, the above-mentioned variances can be approximated as unchanged until the filtering reaches a steady state, and the time index is omitted.

Considering the delay compensation in Section IV-A2, the locally fused position variance is denoted as follows:

$$\begin{aligned}\bar{R}_s &= \left(R_s^{-1} + \sum_{j=1}^M (A|_{\tau} R_{r,j,s} A^T|_{\tau} + Q|_{\tau})^{-1} \right)^{-1} \\ \bar{R}_i &= \left(R_i^{-1} + \sum_{j=1}^M (A|_{\tau} R_{r,j,i} A^T|_{\tau} + Q|_{\tau})^{-1} \right)^{-1}.\end{aligned}\quad (35)$$

After multicast and data fusion including delay compensation in Section IV-C2, V_s gets multiple observations of its own states, denoted as $\hat{z}_s$, $\bar{z}_{i,s}$ in (24) and the variances are

$$\begin{aligned}\hat{R}_s &= A|_{\tau} \bar{R}_s A^T|_{\tau} + Q|_{\tau} \\ \hat{R}_{i,s} &= A|_{\tau} \bar{R}_i A^T|_{\tau} + Q|_{\tau} + R_{s \rightarrow i}.\end{aligned}\quad (36)$$

The equivalent measurement z_g can be denoted as $A_s \hat{z}_s + \sum_{i=1}^N A_i \bar{z}_{i,s}$, where A_i and A_s can be determined by the method of global filter in [12]. The variance of z_g is

$$R_g = \left(\hat{R}_s^{-1} + \sum_{i=1}^N \hat{R}_{i,s}^{-1} \right)^{-1} \quad (37)$$

and the multiple-observation Kalman filter is equivalent to a single-observation Kalman filter with observation z_g .

Variance updates in Kalman filter [26] can be written in the following form:

$$\begin{aligned}P'[k+2] &= AP[k+1]A^T + Q \\ &= A \left[(P'[k+1])^{-1} + (R_g)^{-1} \right]^{-1} A^T + Q.\end{aligned}\quad (38)$$

According to the Woodbury matrix identity, the equation can be rewritten as follows:

$$\begin{aligned}P'[k+2] &= AP'[k+1]A^T \\ &\quad - AP'[k+1] \left[P'[k+1] + R_g \right]^{-1} P'[k+1] A^T \\ &\quad + Q.\end{aligned}\quad (39)$$

When the system reaches the steady state, $P'[k+2]$ is equal to $P'[k+1]$, so the time index can be omitted as follows:

$$AP'A^T - P' - AP'[P' + R_g]^{-1} P'A^T + Q = 0 \quad (40)$$

and the equation is a standard Riccati equation

$$AXA^T - X - AX[X + R]^{-1} XA^T + Q = 0 \quad (41)$$

where

$$A = A, X = P', R = R_g, Q = Q. \quad (42)$$

As a typical nonlinear equation for optimal control problems, there is no closed-form solution for the matrix Riccati equation. Numerical methods are commonly used to obtain the solution. The solution of the Riccati equation in this article represents stable P' which is the prediction variance. To get the optimized estimation variance, the following transformation is needed:

$$P = A^{-1}(P' - Q)(A^T)^{-1}. \quad (43)$$

Therefore, the steady-state variance P of the cooperative tracking algorithm is obtained by (40) and (43).

REFERENCES

- [1] I. Yaqoob, L. U. Khan, S. M. A. Kazmi, M. Imran, N. Guizani, and C. S. Hong, "Autonomous driving cars in smart cities: Recent advances, requirements, and challenges," *IEEE Netw.*, vol. 34, no. 1, pp. 174–181, Jan./Feb. 2020.
- [2] K. Jo, J. Kim, D. Kim, C. Jang, and M. Sunwoo, "Development of autonomous car—Part I: Distributed system architecture and development process," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 7131–7140, Dec. 2014.
- [3] C. Basnayake, T. Williams, P. Alves, and G. Lachapelle, "Can GNSS drive V2X?" *GPS World*, vol. 21, no. 10, pp. 35–43, Oct. 2010.
- [4] G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, "Simultaneous localization and mapping: A survey of current trends in autonomous driving," *IEEE Trans. Intell. Veh.*, vol. 2, no. 3, pp. 194–220, Sep. 2017.
- [5] "Automotive vertical sector," 5G Infrastruct. Assoc., 5G-PPP, Flanders, Belgium, White Paper, Oct. 2015. [Online]. Available: <https://5g-ppp.eu/wp-content/uploads/2014/02/5G-PPP-White-Paper-on-Automotive-Vertical-Sectors.pdf>
- [6] S. Kuutti, S. Fallah, K. Katsaros, M. Dianati, F. McCullough, and A. Mouzakitis, "A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 829–846, Apr. 2018.
- [7] N. M. Elfatih et al., "Internet of vehicle's resource management in 5G networks using AI technologies: Current status and trends," *IET Commun.*, vol. 16, no. 5, pp. 400–420, Dec. 2022.
- [8] X. Cheng, D. Duan, L. Yang, and N. Zheng, "Societal intelligence for safer and smarter transportation," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 9109–9121, Jun. 2021.
- [9] X. Cheng, R. Zhang, and L. Yang, "Wireless toward the era of intelligent vehicles," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 188–202, Feb. 2019.
- [10] X. Cheng, C. Chen, W. Zhang, and Y. Yang, "5G-enabled cooperative intelligent vehicular (5GenCIV) framework: When Benz meets Marconi," *IEEE Intell. Syst.*, vol. 32, no. 3, pp. 53–59, May/Jun. 2017.
- [11] X. Cheng, Z. Huang, and S. Chen, "Vehicular communication channel measurement, modelling, and application for beyond 5G and 6G," *IET Commun.*, vol. 14, no. 19, pp. 3303–3311, Nov. 2020.
- [12] P. Yang, D. Duan, C. Chen, X. Cheng, and L. Yang, "Multi-sensor multi-vehicle (MSMV) localization and mobility tracking for autonomous driving," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, pp. 14355–14364, Dec. 2020.
- [13] W. Pi et al., "Malicious user detection for cooperative mobility tracking in autonomous driving," *IEEE Internet Things J.*, vol. 7, no. 6, pp. 4922–4936, Jun. 2020.
- [14] L. Kleeman, "Optimal estimation of position and heading for mobile robots using ultrasonic beacons and dead-reckoning," in *Proc. ICRA*, vol. 3, Nice, France, May 1992, pp. 2582–2587.
- [15] K. S. Chong and L. Kleeman, "Accurate odometry and error modelling for a mobile robot," in *Proc. ICRA*, vol. 4, Albuquerque, NM, USA, Apr. 1997, pp. 2783–2788.
- [16] H. Durrant-Whyte, "Where am I? A tutorial on mobile vehicle localization," *Ind. Robot.*, vol. 21, no. 2, pp. 11–16, 1994.
- [17] S.-X. Wu and M. Yang, "Landmark pair based localization for intelligent vehicles using laser radar," in *Proc. IEEE Intell. Veh. Symp.*, Istanbul, Turkey, Jun. 2007, pp. 209–214.
- [18] H.-S. Tan and J. Huang, "DGPS-based vehicle-to-vehicle cooperative collision warning: Engineering feasibility viewpoints," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 415–428, Dec. 2006.

- [19] T. Takasu and A. Yasuda, "Evaluation of RTK-GPS performance with low-cost single-frequency GPS receivers," in *Proc. IGSSS*, Tokyo, Japan, 2008, pp. 852–861.
- [20] A. Y. Hata and D. F. Wolf, "Feature detection for vehicle localization in urban environments using a multilayer LIDAR," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 2, pp. 420–429, Feb. 2016.
- [21] J. Castorena and S. Agarwal, "Ground-edge-based LIDAR localization without a reflectivity calibration for autonomous driving," *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 344–351, Jan. 2018.
- [22] C. Li, B. Dai, and T. Wu, "Vision-based precision vehicle localization in urban environments," in *Proc. CAC*, Changsha, China, Nov. 2013, pp. 599–604.
- [23] X. Du and K. K. Tan, "Vision-based approach towards lane line detection and vehicle localization," *Mach. Vis. Appl.*, vol. 27, no. 2, pp. 175–191, Nov. 2016.
- [24] O. Kanhere, S. Ju, Y. Xing, and T. S. Rappaport, "Map-assisted millimeter wave localization for accurate position location," in *Proc. GLOBECOM*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
- [25] T. Bailey and H. Durrant-Whyte, "Simultaneous localization and mapping (SLAM): Part II," *IEEE Robot. Autom. Mag.*, vol. 13, no. 3, pp. 108–117, Sep. 2006.
- [26] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic. Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
- [27] C.-L. Wang and D.-S. Wu, "Decentralized target positioning and tracking based on a weighted extended Kalman filter for wireless sensor networks," *Wireless Netw.*, vol. 19, no. 8, pp. 1915–1931, Apr. 2013.
- [28] L. Mihaylova, D. Angelova, S. Honary, D. R. Bull, C. N. Nanagarajah, and B. Ristic, "Mobility tracking in cellular networks using particle filtering," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3589–3599, Oct. 2007.
- [29] M. Shen, J. Sun, H. Peng, and D. Zhao, "Improving localization accuracy in connected vehicle networks using Rao–Blackwellized particle filters: Theory, simulations, and experiments," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 6, pp. 2255–2266, Jun. 2019.
- [30] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.
- [31] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proc. IEEE*, vol. 97, no. 2, pp. 427–450, Feb. 2009.
- [32] S. Wang, X. Jiang, and H. Wymeersch, "Cooperative localization in wireless sensor networks with AOA measurements," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6760–6773, Aug. 2022.
- [33] B. Li, N. Wu, Y. Wu, and Y. Li, "Convergence-guaranteed parametric Bayesian distributed cooperative localization," *IEEE Trans. Wireless Commun.*, vol. 21, no. 10, pp. 8179–8192, Oct. 2022.
- [34] X. Chu, Z. Lu, D. Gesbert, L. Wang, and X. Wen, "Vehicle localization via cooperative channel mapping," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 5719–5733, Jun. 2021.
- [35] W. Meng, X. Chu, Z. Lu, L. Wang, X. Wen, and M. Li, "V2V communication assisted cooperative localization for connected vehicles," in *Proc. WCNC*, Nanjing, China, Apr. 2021, pp. 1–6.
- [36] K. Lassoued, P. Bonnifait, and I. Fantoni, "Cooperative localization with reliable confidence domains between vehicles sharing GNSS pseudorange errors with no base station," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 1, pp. 22–34, Jan. 2017.
- [37] N. Alam, A. T. Balaei, and A. G. Dempster, "Relative positioning enhancement in VANETs: A tight integration approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 47–55, Mar. 2013.
- [38] J. Liu, B.-G. Cai, and J. Wang, "Cooperative localization of connected vehicles: Integrating GNSS with DSRC using a robust cubature Kalman filter," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 8, pp. 2111–2125, Aug. 2017.
- [39] J. Yao, A. T. Balaei, M. Hassan, N. Alam, and A. G. Dempster, "Improving cooperative positioning for vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 6, pp. 2810–2823, Jun. 2011.
- [40] S. Fang, H. Li, and M. Yang, "LiDAR SLAM based multivehicle cooperative localization using iterated split CIF," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 21137–21147, Nov. 2022.
- [41] R. Lu, R. Zhang, X. Cheng, and L. Yang, "UAV-assisted data dissemination with proactive caching and file sharing in V2X networks," in *Proc. GLOBECOM*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
- [42] L. Chen, H. Hu, and K. McDonald-Maier, "EKF based mobile robot localization," in *Proc. EST*, Sep. 2012, pp. 149–154.
- [43] A. Khattab, Y. A. Fahmy, and A. A. Wahab, "High accuracy GPS-free vehicle localization framework via an INS-assisted single RSU," *Int. J. Distrib. Sens. Netw.*, vol. 11, no. 5, May 2015, Art. no. 795036.
- [44] A. Botta, A. Pescapé, and G. Ventre, "Quality of service statistics over heterogeneous networks: Analysis and applications," *Eur. J. Oper. Res.*, vol. 191, no. 3, pp. 1075–1088, Dec. 2008.
- [45] M. D. Bernardo, P. Falcone, A. Salvi, and S. Santini, "Design, analysis, and experimental validation of a distributed protocol for platooning in the presence of time-varying heterogeneous delays," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 2, pp. 413–427, Mar. 2016.



Sijiang Li (Graduate Student Member, IEEE) received the B.S. degree in electronic information science and technology from Peking University, Beijing, China, in 2021, where he is currently pursuing the Ph.D. degree with the School of Electronics.

His current research interests are cooperative localization, perception, and intelligence of connected mobile agents.



Dongliang Duan (Senior Member, IEEE) received the B.S. degree in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2006, the M.S. degree in electrical engineering from the University of Florida, Gainesville, FL, USA, in 2009, and the Ph.D. degree in electrical engineering from the Colorado State University, Fort Collins, CO, USA, in 2012.

Since graduation, he joined the Department of Electrical and Computer Engineering, University of Wyoming, Laramie, WY, USA, and he is currently

an Associate Professor. His current research interests are signal processing and data analytics for power and intelligent transportation systems.

Dr. Duan is currently a Subject Editor for IET COMMUNICATIONS, and a Guest Editor for the special issue on the IoT for Power Grid on IEEE INTERNET OF THINGS JOURNAL.



Jianan Zhang (Member, IEEE) received the B.E. degree from Tsinghua University, Beijing, China, in 2012, and the M.S. and Ph.D. degrees from Massachusetts Institute of Technology, Cambridge, MA, USA, in 2014 and 2018, respectively.

He is currently an Assistant Professor with the School of Electronics, Peking University, Beijing, China. Prior to joining Peking University, he was a Senior Software Engineer with Google LLC, Menlo Park, CA, USA, from 2018 to 2023. His research interests are network theory and networked intelligence, with focus on resource allocations, cross-system optimizations, robust network designs, and multiagent systems.



Xiang Cheng (Fellow, IEEE) received the Ph.D. degree jointly from Heriot-Watt University, Edinburgh, U.K., and the University of Edinburgh, Edinburgh, in 2009.

He is currently a Boya Distinguished Professor with Peking University. His general research interests are in areas of channel modeling, wireless communications, and data analytics, subject on which he has published more than 280 journal and conference papers, nine books, and holds 23 patents.

Prof. Cheng is a recipient of the IEEE Asia-Pacific Outstanding Young Researcher Award in 2015, a Distinguished Lecturer of IEEE Vehicular Technology Society, and a Highly Cited Chinese Researcher since 2020. He was a co-recipient of the 2016 IEEE JSAC/IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Best Paper Award: Leonard G. Abraham Prize, and the 2021 IET Communications Best Paper Award: Premium Award. He has also received the Best Paper Awards at IEEE ITST'12, ICC'13, ITSC'14, ICC'16, ICNC'17, GLOBECOM'18, ICCS'18, and ICC'19. In 2021 and 2023, he was selected into two world scientist lists, including World's Top 2% Scientists released by Stanford University and Top Computer Science Scientists released by Guide2Research. He has served as the symposium lead chair, a co-chair, and a member of the Technical Program Committee for several international conferences. He led the establishment of three Chinese standards (including one industry standard and two group standards) and participated in the formulation of ten 3GPP international standards and two Chinese industry standards. He is currently a Subject Editor of *IET Communications* and an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE WIRELESS COMMUNICATIONS LETTERS, and the *Journal of Communications and Information Networks*, and is an IEEE Distinguished Lecturer.



Liuqing Yang (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Minnesota, Minneapolis, MN, USA, in 2004.

She has been a Faculty Member with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA; Colorado State University, Fort Collins, CO, USA; and the University of Minnesota, and is currently a Chair Professor with Hong Kong University of Science and Technology (GZ), Hong Kong. She is

also the Acting Head of the Intelligent Transportation Thrust and the Director of Guangzhou Municipal Key Laboratory of Seamless Connectivity and Connected Intelligence. Her general interests are in communications, sensing, and connected intelligence—subjects on which she has published more than 380 journal and conference papers, four book chapters, and six books.

Dr. Yang was the recipient of the ONR Young Investigator Program Award in 2007, and the NSF Faculty Early Career Development (CAREER) Award in 2009, the Best Paper Award at IEEE ICUBW'06, ICC'13, ITSC'14, GLOBECOM'14, ICC'16, WCSP'16, GLOBECOM'18, ICCS'18, and ICC'19. She is the Editor-in-Chief for *IET Communications*, an Executive Editorial Committee Member for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and a Senior Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING. She has also served as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE INTELLIGENT SYSTEMS, and *PHYCOM: Physical Communication*, and as a program chair, a track/symposium, or a TPC chair for many conferences.