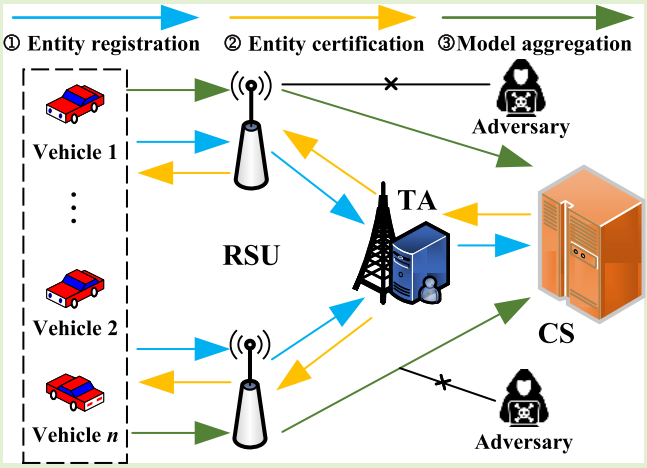
Federated Learning-Based Collaborative Authentication Protocol for Shared Data in Social IoV

Pengcheng Zhao, *Graduate Student Member, IEEE*, Yuanhao Huang, Jianping Gao, Ling Xing, *Member, IEEE*, Honghai Wu, and Huahong Ma

***Abstract*—In the Social Internet of Vehicles (SIoV), federated learning is able to significantly protect the private data of the vehicle's client, while reducing the transmission load between entities. Nevertheless, data can still be stolen by an adversary who analyzes the parameters uploaded by the client to steal it. In this paper, to effectively prevent data leakage and reduce the propagation delay of data, we design a federated learning collaborative authentication protocol for shared data. The parameters of the vehicle client model are encrypted by the protocol in the federated learning. The vehicle and other entities of the protocol realize efficient anonymous mutual authentication and key agreement. The security of the proposed protocol is proved in the stochastic predictive machine model. The simulation results on the SUMO and OMNeT++ platforms show that the authentication delay is the lowest compared to other protocols and the packet loss rate is reduced to 4.68%. Moreover, the overfitting of the globally aggregated model is effectively resolved.**



***Index Terms*—Social Internet of Vehicles (SIoV), data security, federated learning, anonymous mutual authentication.**

I. Introduction

THE Internet of Vehicles (IoV) is formed by the combination and evolution of the Internet of Things (IoT) and Intelligent Transportation System (ITS) [1], [2]. Vehicles share data using dedicated short-range communication (DSRC), dynamic long-term evolutionary in-vehicle communication (LTE-V) and C-V2X technologies, which enable open connectivity for Vehicle-to-Vehicle (V2V), Vehicle-to-RSU (V2R), and Vehicle-to-People (V2P) communication methods [3], [4]. These connectivity approaches enhance vehicle engagement as social users, while the network's cross-modal and cross-platform social capabilities have led to the birth of social vehicle networking theory and application scenarios [5]. Therefore, the Social Internet of Vehicles (SIoV) can be seen as a complex intelligent transportation system comprising several elements [6], including people [7], vehicles and other entities [8].

In the vehicle sharing data structure illustrated in Fig. 1, the vehicle-installed on-board mobile data center (MDC) is used to store data during vehicle sensing, broadcasting, and sharing [9]–[11]. The vehicle client uses multi-sensor fusion technology (such as LIDAR, GPS/BeiDou Navigation Satellite System, and other intelligent devices) to collect a series of environmental data, such as vehicle driving status, real-time speed, precise location, and pictures. These data are uploaded to the RSU for broadcasting via the integrated On-Board Unit (OBU), then shared among the vehicles in close proximity within the surrounding area [12]–[14]. The characteristics of SIoV, such as the high-speed mobility of vehicle nodes, the dynamic variability of the network topology, and the open network autonomous connectivity, have more significant

Manuscript received January 29, 2022; accepted February 17, 2022. Date of publication February 22, 2022; date of current version March 31, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62071170, Grant 62171180, and Grant 62072158; in part by the Program for Innovative Research Team in University of Henan Province under Grant 21IRTSTHN015; in part by the Key Science and Research Program in University of Henan Province under Grant 21A510001; and in part by the Henan Province Science Fund for Distinguished Young Scholars under Grant 222300420006. The associate editor coordinating the review of this article and approving it for publication was Dr. Varun Bajaj. *(Corresponding author: Jianping Gao.)*

Pengcheng Zhao, Yuanhao Huang, Ling Xing, Honghai Wu, and Huahong Ma are with the College of Information Engineering, Henan University of Science and Technology, Luoyang 471003, China (e-mail: zpchaust@163.com; hdyrhk@126.com; xingling_my@haust.edu.cn; honghai2018@haust.edu.cn; mhh@haust.edu.cn).

Jianping Gao is with the College of Vehicle and Traffic Engineering, Henan University of Science and Technology, Luoyang 471003, China (e-mail: gaojp@haust.edu.cn).

Digital Object Identifier 10.1109/JSEN.2022.3153338

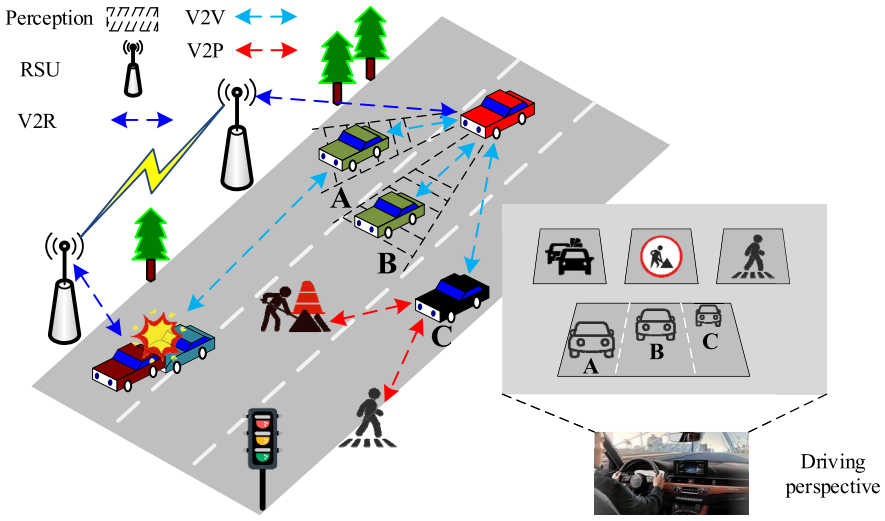


Fig. 1. Social IoV model and vehicle data source.

and urgent requirements for the safe sharing of data and transmission efficiency.

Federated learning, as a type of distributed machine learning, is widely used in the SIoV. This is an efficient method for data sharing and protection in distributed application scenarios. The authors upload the vehicle client data to a central server for aggregation, which largely solves the data privacy issue [15]. This approach of federated learning reduces vehicle data transmission, broadcasting, and sharing costs through training work distribution [16].

However, we know that V2V and V2R communication relies on public wireless networks. At the same time, federated learning makes the uploaded parameters vulnerable to different types of attackers. The attacker can infer the training model by analyzing the parameters uploaded by the client, and private data can still be compromised. Thus, the anonymous mutual authentication and key agreement protocol protects the security vulnerability of the uploaded parameters and the forward security of federated learning. Likewise, federated learning provides a low communication overhead for secure protocols. The main contributions of our work can be summarized as follows:

- We propose a vehicle social community, which is constructed based on the context of the vehicle broadcast messages and common interests. We use this social community to verify the performance indicators of the method in this paper.
- We design a federated learning collaborative authentication protocol for shared data. This protocol protects the safety of uploading parameters of vehicle client models and reduces communication costs in the federated learning. At the same time, anonymous mutual authentication and key agreement are considered by us in the protocol.
- We proved the security of the proposed protocol, such as resistance to replay attacks, sybil attacks, and so on. Our experiment evaluated the security, computational cost and authentication delay of the protocol. Meanwhile, the overfitting of federated learning is verified by us.

The rest of this paper is organized as follows: In Section II, we describe the related work that is the research status. Section III, we describe the background knowledge of federated learning in SIoV, followed by the threat model and security goals. In Section IV, we provide details for the federated learning model and the process of protocol. In Section V, the security analysis of the proposed protocol is explained in detail by us. We present the simulation details and results in Section VI followed by a conclusion in Section VII.

II. Related Work

In this section, we mainly reviewed research of the collaborative data security solution of federated learning in the SIoV. Further, we reviewed the advantages and disadvantages of different vehicle certification protocols in detail.

We find that federated learning in SIoV can be fused with different typical data protection algorithms to create collaborative data security solutions, which has the characteristics of generalization, inferential, and formalize [17]. In [18], [19], the authors use deep learning with reinforcement learning to optimize the association mechanism of federated learning. While this mechanism meets the confidentiality and integrity requirements for data shared among vehicles, it also increases the training model burden and algorithmic computational overhead of clients and servers. In [20]–[22], the authors integrate federated learning with Blockchain, which enhances the data security and aggregation efficiency of telematics; however, the communication overhead does not meet the new and higher demand. Therefore, the author proposed an edge computing-federated learning network architecture, which has high communication efficiency but leads to an imbalance of interoperability and weakened generalization capability [23]–[25].

In view of the above research content, we find that federated learning can be combined with the vehicle collaborative authentication protocol, which can realize the protection of private data and the vehicle user authentication mechanism. That is able to provide a conditional data security mechanism

for open public wireless vehicle networking communication technology [26]. Researchers have previously proposed identity encryption schemes based on bilinear pairs, although the computational overhead of the pairing operation is relatively large. We determined that this scheme is unable to meet the current delay requirements for vehicle broadcast traffic-related applications [27]. In [28], the authors proposed a binary authentication tree-based approach, which implements bulk verification of message signatures. However, we find that this scheme relies heavily on semi-trusted roadside units (RSUs). To overcome the semi-trusted RSU problem, in [29] the authors employ the fast computational efficiency of the hash function to design a lightweight vehicle authentication scheme. The goal is to achieve mutual authentication of onboard OBU, RSU, and trusted institutions. Trusted institutions are added to the protocol, while they use public key encryption with symmetric encryption to improve the security and computational efficiency of the vehicle protocol [30]. However, we find that both this scheme and that outlined in the paper [29] are not resistant to replay and tampering attacks. In [31], the authors proposed a secure authentication protocol based on a key distribution center to address denial of service attacks in vehicular networking. In [32], the authors proposed a PPAS authentication mechanism with some vehicle location anonymity. The authors proposed a vehicle cloud privacy-preserving authentication model [33]. This model is able to include the vehicle cloud in the authentication process, which indirectly helps the vehicle to obtain the correct data for sharing decisions.

III. Background Knowledge

In this section, we present background knowledge on the application of federated learning in SIoV [34]. Furthermore, we present the adversary objectives and adversary models faced by the protocol. Finally, we define the security goals based on the security requirements of vehicle data.

A. Federated Learning in SIoV

Federated learning is essentially a distributed machine learning technique. In the context of SIoV architecture, we find that this technique allows data to be transmitted, broadcast, and shared between vehicles in a way that both protects private data and enables model training to make driving decisions without the need for raw data from vehicle users [35]. Federated learning is composed primarily of n-clients and a central server, where local clients with the same data structure collaborate to learn ML models with the aid of the central server [36], [37]. The vehicle client is responsible for local data training to obtain local model parameters. The central server is a secure aggregation of the uploaded parameters for each client model, and is able to iteratively update the global model without the need to learn the local data.

Assume that there are $\{C_1, C_2, \ldots, C_i\}$ vehicles (participants), each of which has client data $\{D_1, D_2, \ldots, D_i\}$ where $i \in \{1, 2, \ldots, n\}$. At the Base Stations (BSs) side of the SIoV, this port collects the local training model parameters of the n clients. The model training parameters for a particular vehicle C_i are represented as a vector w_i. We observe that the central server performs a secure aggregation of the parameters received by the n clients, as follows:

$$w = \sum_{i=1}^{n} p_i w_i \tag{1}$$

Each vehicle client database D_i is trained to derive the upload parameters w_i for the i-th client, where w is the global model parameters after aggregation on the central server. Equation (2) represents the model and optimal model parameters following global aggregation; these are aggregated with the purpose of minimizing a certain loss function.

$$w^* = argmin_w \sum_{i=1}^{n} p_i F_i(w, \mathrm{D}_i) \tag{2}$$

In the above $\left|D'\right| = \sum_{i=1}^{N} |D_i|$, $p_i = \frac{|D_i|}{|D'|} \geq 0$ and $\sum_{i=1}^{N} p_i = 1$, w^* denotes the global model parameter after iterative updating. We define $F_i(\cdot)$ as representing the local loss function computed for the D_i client. In the context of federated learning for the SIoV application scenario. We know that the vehicle OBU is only responsible for uploading the model parameters trained by the local client. These parameters are sent to a central server for global aggregation [38].

B. Threat Model Analysis

SIoV communication is carried out over open public wireless channels [39], [40]. Therefore, there are many potential attacks that may seriously affect the proper functioning of such communications on the road. Utilizing the adversary capability hypothesis, we assume that probabilistic polynomial time (PPT) adversary $\mathcal{A}$ has a strong antagonistic capability. $\mathcal{A}$ is able to perform different types of tasks due to the fact that all parameters pass through an insecure channel (except for the registration process). In the specific field of the SIoV, the following network attacks are assumed to be capabilities of attacker $\mathcal{A}$:

(1) Tampering attacks [41]: Adversary $\mathcal{A}$, by means of interception, modification, deletion, replacement, etc., will alter the transmitted and broadcast data during or after transmission, while it interferes with the normal communication between the vehicle and the entity.

(2) Replay attack [42]: $\mathcal{A}$ constantly repeats or replays valid data and messages previously sent by the vehicle, and these actions disrupt normal communication.

(3) Sybil attack [43]: $\mathcal{A}$ simulates communication between entities, initiating attacks using false identities or false RSUs, and these actions cause errors in traffic emergency data commands.

(4) Man-in-the-middle attack (Wormhole Attack) [44]: $\mathcal{A}$ is able to change the communication channel between the two parties. This behavior causes both parties to believe that they are communicating directly with each other to complete data sharing.

C. Safety Objectives

In the SIoV federated learning architecture, we can observe that the vehicle client datasets are stored in the local MDC.

However, intermediate parameters need to be shared with the central server, which may result in the leakage of private client data, such as through tampering attacks, Sybil attacks, etc. Moreover, by analyzing the global parameters, we can determine that it is possible for privacy leakage to occur in the broadcast (downlink channel) phase. Therefore, data security in SIoV needs to achieve the following security objectives:

(1) Our primary security goal is the authenticity and integrity of the data source. We know that the data is sent by OBU or broadcast by RSU. $\mathcal{A}$ is unable to tamper with and intercept the vehicle client upload parameter w_i through attack capability during transmission, while the receiver is able to verify the vehicle's authenticity. The integrity of the information source is achieved through signatures. The proof here can refer to the V, B-(1).

(2) One of our goals is vehicle anonymity. When the sender provides the vehicle location and trajectory data, the adversary cannot identify the vehicle user ID through active and/or passive attacks. The identity information of the vehicle should be anonymous to the receiver of the parameter, and the adversary should be unable to trace the vehicle path trajectory. We can refer to security analysis to prove anonymity during in the paper V.

(3) The goal of the protocol is to have forward security. This protocol should protect communication previously carried out by the vehicle from the threat of passwords and keys. An attacker is not able to establish a pre-session to obtain the vehicle client upload parameter w_i, and also cannot reason to determine information about the global aggregation model. The proof of forward security can refer to the paper V, B-(5).

IV. Proposed Lightweight Protocol in an SIoV

In this section, we propose a lightweight collaborative authentication mechanism for SIoV based on federated learning [45], [46]. First, we define the concept of social community for SIoV based on the knowledge of social network evolution. We then introduce the system model of the protocol and the system initialization, registration, authentication, and aggregation processes.

A. Vehicle Social Community

In the SIoV, vehicles are considered to be entities, and these entities generate social awareness that constitute social relationships. Vehicles are social relationships that use shared environmental data to build trust with other vehicles within range. The vehicle social community structure is built based on broadcast context messages and common interests in the vehicle client network. The amount of data broadcasted and shared by different community structures is different, and these structures require different communication resources. The proposed protocol needs to verify the vehicle communication overhead and security in different community structures.

In the below, we classify the social community structure of SIoV into three types:

- All-Static Community: We define a community of vehicles parked in the parking lot with other vehicles and infrastructure (charging stations, parking information, etc.). This community is in social form.
- All-Dynamic Community: We define a fully dynamic community as one in which vehicles drive on roads such as highways and mountain roads. Vehicular OBU with RSU and TA works to form a network topology community structure with highly dynamic multi-hops.
- Mixed-State Community: We define a network link formed by vehicles traveling between complex urban roads and fixed urban units (office buildings, gas stations, etc.). Such an approach constitutes a hybrid state community structure.

We conduct a comparison of the shared and broadcast data volume of different vehicle social communities: All-static community < All-dynamic community < Mixed-state community. The structure of vehicle social communities is illustrated in Fig. 2.

B. System Model

Fig. 2 illustrates federated learning based on a Shared Collaborative authentication protocol for Perceptual data in SIoV. Under the federated learning mechanism, the system model is the communication system between the vehicle and the central server (content service provider; CS). The system model comprises four types of participants: Trusted Agency (TA), Content Service Provider (CS), Roadside Unit (RSU), and Vehicle Unit (OBU).

(1) Trusted Agency (TA): TA is a vehicle registration and authentication agency. This is a third-party agency trusted jointly by the user C_i, content provider CS, and roadside unit RSU. It is assumed that, ideally, this authority is resistant to all types of known Internet attacks and has sufficient computing and storage resources.

(2) Content Service Provider (CS): Based on the federated learning framework, CS is a server that provides SIoV applications and value-added services. Instructions such as automatic vehicle driving decisions and road traffic environment data are issued by CS. At the same time, CS is a social service that provides a range of weather, news and entertainment advice for vehicles, drivers and passengers. We know that the server updates the optimal model by securely aggregating the parameters uploaded by the vehicle clients in the n communities to complete the optimal model. This is done to achieve minimization of the loss function.

(3) Roadside Unit (RSU): This entity belongs to a light roadside base station, which is the key decision command sender and vehicle client broadcast upload parameter relay for SIoV; it can request authentication and obtain CS aggregation model parameters from the TA. It adopts DSRC technology, and they obtain relevant information of vehicle users through the broadcasting mechanism. Both the RSU and the vehicle user need to request registration from the TA, and the TA broadcasts the parameters for the RSU and vehicle authentication. The RSU verifies the identity of the vehicle as it passes within the range of the RSU. To sum up, this is a broadcast storage mechanism of RSU.

(4) On-board unit (OBU): The OBU unit in vehicle user C_i is a communication module with a certain degree of computing power. This is a device for communication and data sharing

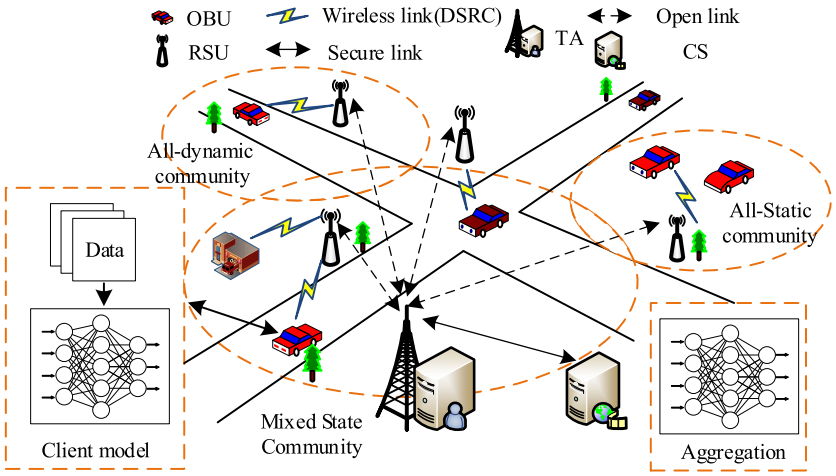


Fig. 2. System model.

TABLE I
PROPERTIES SYSTEM NOTATION AND DESCRIPTION

Notation	Description
C_i	User vehicles
ID_i/PW_i	Vehicle C_i username/password
ID_{cs} / s_j	CS's ID and private key
ID_{RSU}/s_k	RSU's ID and private key
$\{D_1, D_2, \ldots, D_i\}$	Vehicle client data
w_i	The i client training parameter vector
$F_i(\cdot)/h(\cdot)$	Vehicle client loss / hash function
$\oplus/\|$	exclusive-or/ string concatenation
$T_i/\triangle$ T	Timestamp / Lifecycle Threshold

between the vehicle and RSU or other vehicles. The vehicle unit OBU uses DSRC or C-V2X to establish a communication link with the RSU. This link is capable of uploading vehicle client training parameters.

C. Phases Involved in the Proposed Protocol

The protocol comprises different phases, such as system initialization, registration, authentication, and aggregation. The registration phase is a registration process that contains three entities: vehicle users, content service providers, and roadside units. The symbols and descriptions involved in the protocol are defined in Table I.

1) System Initialization Phase: The user vehicle C_i needs to be initialized for the system model. The RSU and CS of the community in which the vehicle is located register the vehicle information in advance with the TA. The vehicle obtains a shared key K. The user vehicle is registered in the TA and initializes the system parameters. We define a finite field F_P of order large prime p; moreover, TA is an elliptic curve E_q over the domain of F_P: $y^2 = x^3 + ax + bmodp$, and $a, b \in F_P, 4a^3 + 27b^2 \neq 0\ modp$. We choose a group G on E_q whose order is q. P is the element generated by the algorithm. We then randomly choose a private key $s \in Z_q^*$ of the system and compute the system public key $PK = sP$. TA selects the appropriate hash function $h(\cdot): \{0,1\}^* \rightarrow Z_p^*$, and It broadcasts the system parameters $H = \{PK, P, h(\cdot)\}$.

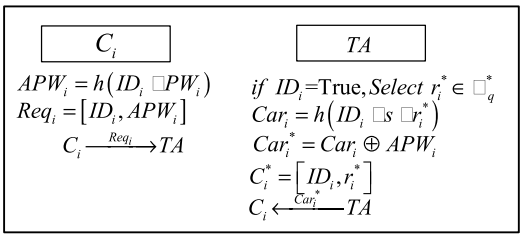


Fig. 3. Vehicle registration.

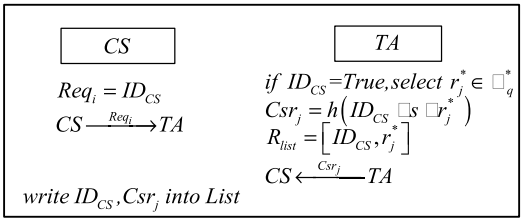


Fig. 4. CS registration.

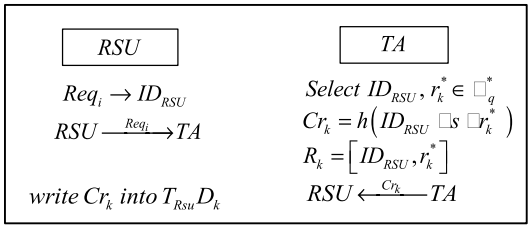


Fig. 5. RSU registration.

2) Protocol Entity Registration Phase:

- Vehicle user registration: Each community vehicle user C_i is registered to TA. The i represents the vehicle user serial number, where $i = 1, 2, \ldots, N$ and N represents the number of vehicle users. C_i selects the username ID_i and registration password PW_i to compute the pseudo password $APW_i = h(ID_i \| PW_i)$, and then we send vehicle ID_i and pseudo-password to form an information tuple Req_i to TA to request registration through a

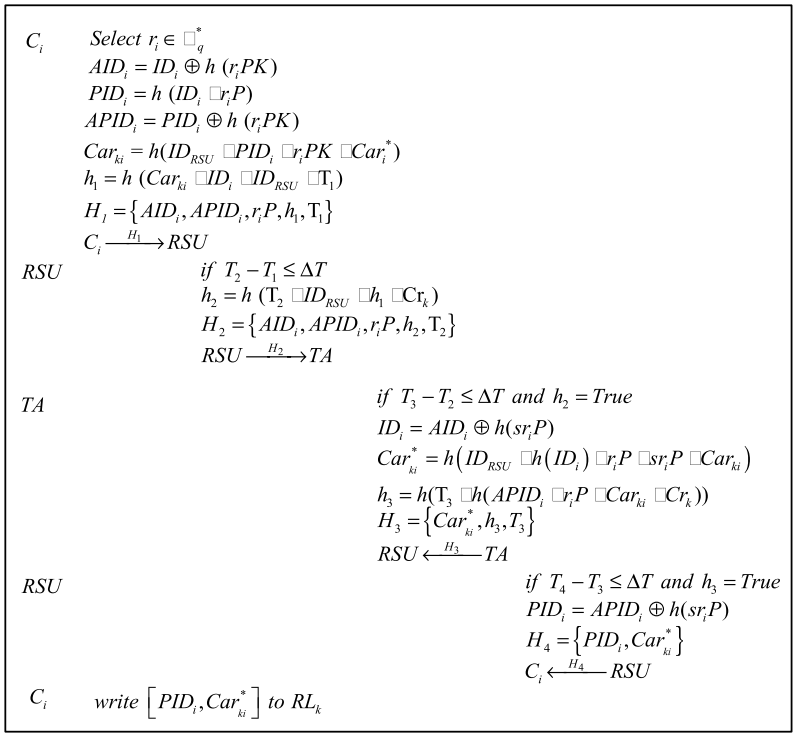


Fig. 6. Vehicle authenticate process.

secure channel. The TA receives the relevant message and checks the validity of the username. The protocol first checks whether the revocation list (RL) contains the vehicle ID: if it contains, it means that the registration has been completed, then the registration is rejected; if it is not, the ID registration is valid. The protocol is to select a random number $r_i^* \in Z_q^*$ compute the long-term authentication credentials Car_i, Car_i^*. The authentication credentials are sent to the vehicle user C_i through a secure channel. C_i is computed by $Car_i^* = Car_i \oplus APW_i$. The tuple ID_i and r_i^* are written to the vehicle OBU in the registration list C_i^*, and vehicle registration is complete. Otherwise, the vehicle discards Car_i, Car_i^*.

- Content Service registration: The content service provider CS sends a registration request to TA after the system model initialization is completed. CS selects the appropriate ID_{cs} to send to TA and then requests registration. TA checks the revocation list RL after receiving the request. When the validity of the identity ID_{cs} is true, we select the random number $r_j^* \in Z_q^*$ to generate the service provider credentials Csr_j, then we send the tuple to CS via a secure channel. The ID_{cs} in the tuple, Csr_j, are written to the server storage list. After CS receives the tuple, it selects the private key s_j. The protocol computes the public key $PK_j = s_j P$, which is published in the responding vehicle's social community.
- RSU registration: The roadside unit RSU sends a registration request to the TA after system initialization. TA receives the request and it selects username ID_{RSU} and $\mathrm{r}_k^* \in Z_q^*$ for this unit. The above is used to generate the registration credentials Cr_k and we send the tuple to the RSU via a secure channel. We write ID_{RSU}, r_k^* in the tuple to the RSU registration list R_k. RSU receives the tuple and selects the private key $sk_k = s_k \in Z_p^*$ to write to the security module $T_{\mathrm{Rsu}}\mathrm{D}_k$, which computes the public key $PK_k = s_k P$ for publication.

3) Certification Stage: We know that vehicle user C_i and content service provider CS are bi-directionally authenticated and initialized with system keys via RSU. Vehicle C_i dynamically drives into different vehicle social communities that are connected (interconnected) with RSUs. The purpose of the vehicle is to generate temporary registration credentials Car_{ki}^* for C_i after the vehicle passes TA authentication. The authentication process of the vehicle user is outlined in Fig. 6.

- After the vehicle C_i is authenticated by login, a random number $r_i \in Z_q^*$ is selected, and the pseudonym for PID_i is computed. There is $PID_i = \mathrm{h}(\mathrm{ID}_i \| r_i P)$, pseudo-identified vehicle user name ID_i and pseudonym PID_i are $AID_i = ID_i \oplus h(r_i PK)$ and $\mathrm{A}PID_i = PID_i \oplus h(r_i PK)$. The protocol computes $h_1 = h(ID_i||ID_{RSU}||Car_{ki}||T_1)$, where $Car_{ki} = h(ID_{RSU}||PID_i||r_i PK \| Car_i^*)$, and T_1 is the timestamp. The composition tuple $\mathrm{H}_1 = \{AID_i, \mathrm{A}PID_i, r_i P, h_1, T_1\}$ sends the timestamp to RSU for authentication through a secure channel.
- After the RSU receives H_1, it verifies the freshness of timestamp T_1. Once the RSU verifies that h_1 is valid, it computes $h_2 = h(T_2||ID_{RSU}||h_1||Cr_k)$, where T_2 is the timestamp, and we sends the tuple $\mathrm{H}_2 = \{AID_i, \mathrm{A}PID_i, r_i P, h_2, T_2\}$ to TA request authentication. After TA verifies and h_2 are valid. The ID_i is resolved from AID_i. The temporary credential

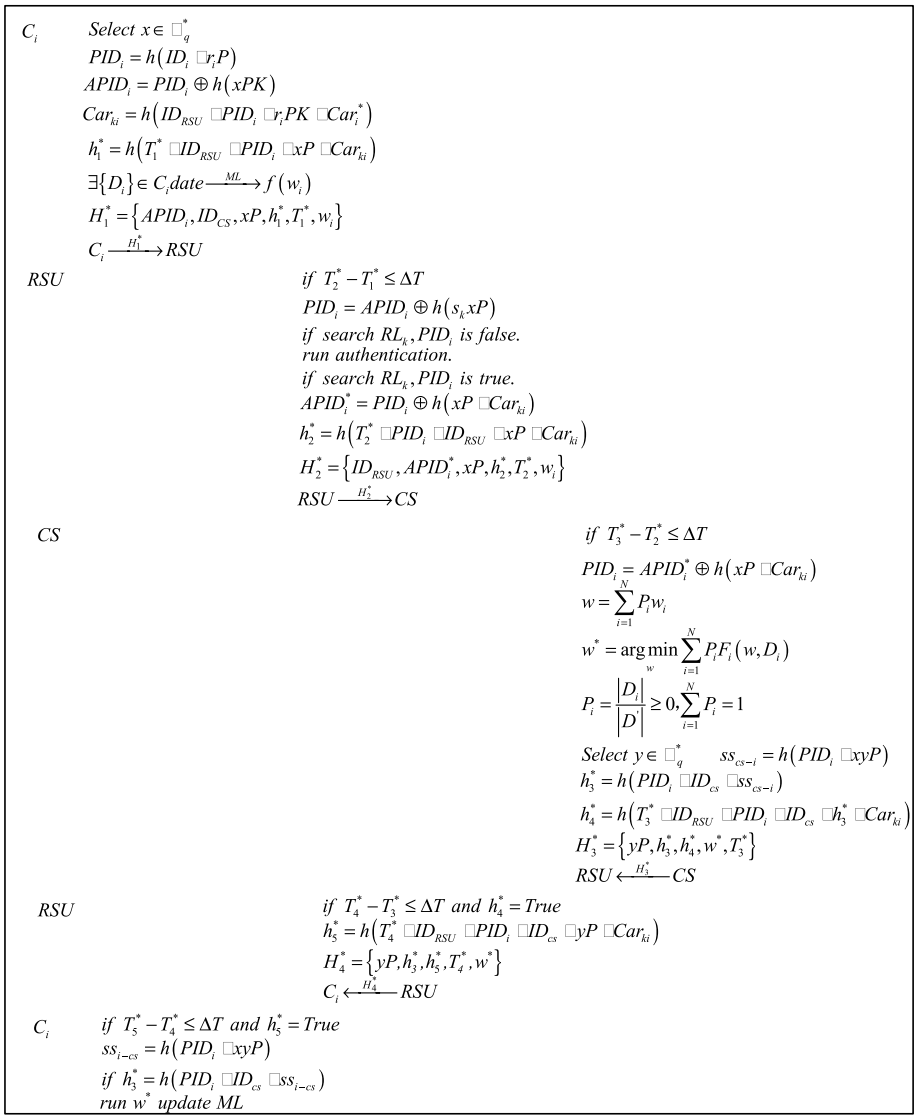


Fig. 7. Vehicle aggregate process.

Car_{ki}^* is computed after querying the user registration table and it return true. We can compute $h_3 = h(T_3 \| h(APID_i \| r_i P \| Car_{ki} \| Cr_k))$ and where $Cr_k = h(ID_{RSU} \| s \| r_k^*)$. The tuple $\mathrm{H}_3 = \{Car_{ki}^*, h_3, T_3\}$ is returned to RSU.

- After RSU verifies that timestamps T_3 and h_3 are valid, PID_i is resolved from $APID_i$. RSU computes $\mathrm{H}_4 = \{Car_{ki}^*, PID_i\}$ and sends it to the corresponding vehicle. The vehicle writes the temporary credential log Car_{ki}^* to the vehicle credential list RL_k. This is done to facilitate fast and accurate authentication when the vehicle uploads the tuple data in the next round.

4) Aggregation Stage: In the framework of a federated learning-based collaborative authentication protocol, the training model of C_i is obtain from the data $\{D_i\}$. The parameters of the model are w_i. When a vehicle accesses the CS to participate in global model aggregation, the RSU is authenticated based on the Car_{ki} in the security module $T_{\mathrm{Rsu}}\mathrm{D}_k$. TA's involvement is not required in this process. We know that such an approach ensures secure data sharing while also being able to reduce the communication and computational costs between entities. The aggregation process is outlined in Fig. 7.

- Vehicle C_i chooses a random number $x \in Z_q^*$, which is used to compute the pseudonym ID_i and pseudo-identifier PID_i. The vehicle computes the temporary voucher Car_{ki} and generates h_1^*. Meanwhile, C_i computes the training model parameters in the local vehicle client dataset $\{D_i\}$. The locally trained ML parameters are denoted as w_i (denoting the $i-th$ vehicle client training parameter vector). The vehicle client computes the tuple $H_1^* = \{APID_i, ID_{cs}, xp, h_1^*, T_1^*, w_i\}$, after which H_1^* is sent from the vehicle unit OBU to the RSU to request authentication. w_i is protected in the protocol, it can ensure the integrity of the data, the detailed proof is in the A and B-(1) of the V.
- RSU verifies the freshness of timestamp T_1^*, then checks the mapping relationship between temporary credentials Car_{ki} and PID_i. If PID_i is valid, the next step is to

compute and resolve PID_i from APID_i and search the list

- RL_k stored in the vehicle. PID_i computes $h_2^* = h\left(T_2^* || PID_i || ID_{RSU} || xP || Car_{ki}\right)$. RSU sends the tuple H_2^* to CS to request global aggregation; if it is false, the vehicle authentication process is re-executed.
- CS verifies that the timestamp T_2^* is valid and PID_i is computed in APID$_i^*$ and verifies h_2^*. If it is valid, the vehicle client uploads parameter w_i in tuple H_2^*, which is involved in the global security aggregation $w = \sum_{i=1}^{n} p_i w_i$, and the optimal model parameter w^* is iteratively trained. Based on ECC, CS selects a random number $y \in Z_q^*$ for using in computing the session key ss_{cs-i}. They then compute h_3^*, h_4^*. The session key is involved in H_3^* encryption w^*, which is sent to RSU through CS. w^* is protected to avoid Sybil attacks and Man-in-the-middle attacks. The proof is in V, B-(3), (4).
- The roadside unit RSU verifies that the timestamp T_3^* is fresh and h_4^* is true. It computes h_5^*, which is broadcasts the tuple H_4^* to the vehicles.
- Vehicle C_i receives the H_4^* authentication timestamp T_4^* and $h_4^* = True$. C_i computes the session key ss_{cs-i} to authenticate h_3^*. The vehicle decrypts the received tuple from the global model w^*, which updates the vehicle client model and executes relevant decisions and instructions.

V. Security Analysis

This section explains the security analysis of the proposed protocol. We have completed the stochastic prediction model for formal analysis. We perform security proofs against Adversary targets and threat models in Section III-B.

A. Proof of Security: A Formal Analysis

Based on the adversary goal and model, we propose that the security proof of the protocol be defined as a game between challenger $\mathcal{C}$ and adversary $\mathcal{A}$. For the purposes of this exercise, adversary $\mathcal{A}$ is able to adaptively execute four types of prophecy machine queries in this game.

- Send query: The query attack was initiated by Adversary $\mathcal{A}$ to challenger $\mathcal{C}$. $\mathcal{C}$ chooses a random number $r_i^* \in Z_P^*$. The tuple H_i^* is added to the Hash table species and r_i^* is returned to send to $\mathcal{A}$.
- Execute queries: Adversary $\mathcal{A}$ launches queries that are executed by individual entities, including the process of sharing data among entities.
- Reveal query: h_i^* is to return the session key, while the vehicle OBU$_i$ and CS are to return ss_{cs-i}.
- Test asks: If adversary $\mathcal{A}$ wins the Game, return session key s_k; otherwise, return session key string.

Based on the stochastic predictive machine model, the protocol is satisfyingly secure.

Definition Game: The protocol in this paper is denoted as F. We assume that $succ_i$ denotes the security of $\mathcal{A}$ by asking to destroy F. If $Adv_F\mathcal{A}$ can be ignored, then F satisfies its security.

G_0: We define real vehicle social community application scenarios based on the Game. The probability advantage of the adversary winning the Game can be defined as follows:

$$Adv_F\mathcal{A} = |2Prsucc_i - 1| \tag{3}$$

G_1: $\mathcal{A}$ initiates the execution of the query to challenger $\mathcal{C}$, which simulates the adversary $\mathcal{A}$ launching a passive attack against OBU$_i$ with RSU. In the F protocol, C_i and CS independently compute PID_i and $x_i p$ hidden in the authentication tuple H$_i^*$. We can determine that $Pr_{succ_0} = Pr_{succ_1}$.

Under the stochastic predictive machine model, the protocol is able to resist active attacks by the adversary.

G_2: We simulate the Game as G_1 based on sending an interrogation, which simulates' active attack capability. Vehicle C_i forges OBU$_i$ to send H$_i^*$ under adversary's imitation attack to obtain an advantage in the Game, denoted as (4).

$$Pr_{succ_2} - Pr_{succ_1} \le \frac{q_n^2}{2^{\varepsilon+1}} + q_e + \frac{q_s^2}{2q} \tag{4}$$

The protocol is able to guarantee the anonymity of the vehicle and RSU, and is protected from vehicle track data theft by attackers.

G_3: Adversary $\mathcal{A}$ initiates a decryption query. When OBU$_i$ is leaked as RSU sends H$_i^*$, $\mathcal{A}$ guesses that the advantage of the correct vehicle ID$_i$ and PW_i to win the Game is as indicated in (5); when the RSU cache is leaked, $\mathcal{A}$ guesses that the advantage of ID$_i$ and temporary credential Car_{ki} to win the Game is as in (6).

$$P_{r_{succ_3}} - P_{r_{succ_2}} \le \frac{q_s}{2^{\alpha+\beta}} \tag{5}$$

$$P_{r_{succ_3}} - P_{r_{succ_2}} \le \frac{q_s}{2^{\varepsilon}} \tag{6}$$

G_4: Adversary $\mathcal{A}$ initiates detection queries to solve the Elliptic Curve Diffie–Hellman under the stochastic predictive machine security model, and $\mathcal{A}$ wins the game by the advantage of (7).

$$Pr_{succ_4} - Pr_{succ_3} \le q_h \cdot Adv_F\mathcal{A} \tag{7}$$

In summary, considering all attack capabilities of $\mathcal{A}$ under the security model, its maximum advantage of winning the Game is expressed in (8).

$$Adv_F\mathcal{A} \le \max\left(\frac{q_n^2}{2^{\epsilon}} + q_e + \frac{q_s^2}{2q}, \frac{q_s}{2^{\epsilon}}\right) + \frac{q_s}{2^{\alpha+\beta}} + q_n \cdot Adv_F\mathcal{A} \tag{8}$$

$T_n(x), n \in [1, q-1]$, where q_n, q_e, q_s denote the number of requests sent, executed, and tested queries. Moreover, ϵ, α, β represent Hash functions, ID and PW_i length respectively.

The protocol is able to guarantee the integrity of messages in the random prediction model.

In the proposed protocol, the verification tuple H_i^* at the receiving end is used to determine the integrity of the upload parameter w_i and the RSU broadcast w^* during data sharing. If successful, it indicates that their integrity is protected. Our proposed protocol is secure and the $\mathcal{A}$ advantage is negligible.

B. Formal Proof of Threat Model Attacks

(1) Message integrity and validation: We determine that, during the transmission and broadcast of vehicle messages, the upload parameter w_i is trained by the vehicle client in the federated learning mechanism, while w^* is optimized by the CS central server global aggregation. These elements are embedded in H_1^* sent by OBU and H_3^* broadcast by CS to obtain the effective integrity guarantee of H_i^*. RSU is verified by vehicle ID through pseudonym PID_i and temporary registration credentials Car_{ki}.

(2) Replay attack resistance: Obviously, in the protocol, this attack simulation does not work; the timestamp T_i is attached to the corresponding H_i. Time synchronization is maintained for all vehicles in the SIoV community structure. If adversary $\mathcal{A}$ steals $H_1^* = (ID_{CS}, APID_i, xP, w_i, h_1^*, T_1^*)$ and initiates a replay attack, the time period $T_{i+1} - T_i \leq T$ is not a transmission delay recognized by both vehicles; therefore, the receiving vehicle rejects the message.

(3) Sybil attack resistance: Adversary $\mathcal{A}$ is unable to compute the random number r_i, meaning that they cannot obtain a valid PID_i, which indirectly indicates that the theft of temporary vehicle credentials cannot be achieved. The vehicle cannot impersonate OBU_i through a fake client. Moreover, $\mathcal{A}$ is unable to compute the temporary credential logs in order to obtain RSUs and thus cannot pass the authentication effectively; that is, they cannot impersonate a fake RSU and thus steal the upload parameter w_i from the vehicle client within the social community.

(4) Man-in-the-middle attack resistance: In the protocol, the vehicle ID is pseudo-identified as $APID_i$. The adversary cannot trace the real identity of the vehicle user C_i, meaning that the vehicle identity privacy is protected. Meanwhile, temporary credentials Car_{ki} are able to authenticate each other through lists, and these protocol processes can resist man-in-the-middle attacks.

(5) Forward security: Vehicle C_i and the central server perform computations independently. Each time they do not share the secret, they compute the session key $\text{ss} = h(\cdot)$. The protocol chooses random numbers $x, y \in Z_p^*$ with a certain freshness (this property is determined by the one-way hash property). The adversary $\mathcal{A}$ is unable to compute the generated C_i and CS session keys, providing forward security.

VI. Protocol Implementation and Analysis

The simulation experiments were conducted in three groups. The first group is the computational and communication costs used as performance metrics. Computational cost refers to the time required to execute the necessary protocol processes, while communication cost denotes the number of bytes transmitted over the communication channel. The second set of simulation experiments uses the OMNeT++ simulation tool, which analyzes the network performance of the protocol. The third set of simulations uses the Python language. We complete the vehicle client training and global aggregation fit simulation of the algorithm in a federated learning framework. The simulations are checked for protocol safety and convergence. The simulations use the C/C++ cryptographic library of MIRACL to measure the cryptographic operations related to the proposed protocol and other protocols. The simulation experiments are performed using an i7-11700F 8-core DDR6 and NVIDIA RTX3060ti in a Windows 11.

TABLE II
TMIE COST OF EACH OPERATION EXECUTED

Notation	Operation	Cost(ms)
T_h	Hash (SHA-256)	0.0002
T_{pm-ecc}	ECC multiplication	2.3510
T_{ske-d}	Symmetric	0.0076
T_{pke}	Public key encryption	3.1210
T_{pkd}	Public key decryption	0.6260
T_{bp}	Bilinear pairs	4.2110

TABLE III
COMPUTATIONAL COST

Name	Operation execution process	TC(ms)
SUAA[30]	$3T_h$+$4T_{ske-d}$+T_{pke}+T_{pkd}	6.89660
RFID[31]	$5T_{pm-ecc}$+$8T_h$+T_{pke}+T_{pkd}	15.5036
PPAS[32]	$8T_h$+$2T_{pm-ecc}$+$2T_{bp}$+T_{ske-d}	15.1332
VCC[33]	$2T_{bp}$+$4T_h$+$2T_{pm-ecc}$+T_{ske-d}	13.1324
Paper Protocol	$9T_h$+$5T_{pm-ecc}$	11.7568

A. Computation and Communication Cost

We analyze the performance of the protocols in this paper to ensure their effectiveness. This set of simulations is carried out by comparing four other authentication protocols published in prior works in terms of their computation and communication costs. According to the relevant findings in the paper [30], [31], the basic cryptographic operation times associated with the protocols of this paper are shown in Table II.

Computational costs are analyzed with reference to the user registration phase, authentication and aggregation processes. Vehicle client model training, $\oplus$ and $\|$ can be ignored. The protocol computes costs in two parts. For vehicle user C_i in the registration phase, they generate pseudonyms that require two Hash operations and one ECC multiplication operation, while pseudonym anonymization $\oplus$ is negligible. They further generate temporary user credentials with messages employing two Hash operations and one ECC. It is known that the cost of the registration phase calculation is $4T_h + 2T_{pm-ecc} = 4.7028$ms. The computed cost of the vehicle from the authentication-aggregation-broadcast phase is $5T_h + 3T_{pm-ecc} = 7.054$, while the total time required is determined to be 11.7568ms. The protocol cost performance comparison is presented in Table III.

The communication efficiency of the protocol is evaluated by the communication cost. The byte lengths of the data structures populated in this set of simulations are listed in Table IV. The vehicle social community sharing data is processed by an ANN with a high generalization capability. In the ANN training model, we assume that x is the input, the intermediate layer is y, and the output layer is z. There are ten model parameters included in the equation (9) – (10), and the number

TABLE IV
LENGTH OF PROTOCOL DATA

Notation	Operation	Length(bytes)
L_h	Hash	20
L_t	Timestamp	4
L_{pid}	Pseudonym	46
L_{id}	ID length	256
L_{ecc}	ECC point	128
L_s	Session key	256
L_w	Parameters	Para*4

TABLE V
COMMUNICATION COST

Name	Tuple Structure	CC(bytes)
SUAA[30]	$3L_h$+L_{id}+$18L_s$+$8L_t$	4956
RFID[31]	$8L_h$+$5L_t$+$5L_{ecc}$+$3L_{id}$+$2L_{pid}$	1680
PPAS[32]	$8L_h$+$2L_{ecc}$+$3L_{pid}$+L_{id}+$12L_s$	3882
VCC[33]	$4L_h$+$2L_{ecc}$+$4L_{pid}$+$4L_{id}$+$2L_t$	1552
Paper Protocol	$9L_h$+$4L_{ecc}$+$2L_{id}$+$5L_t$+L_w	1264

of $L_w = 40$ bytes is computed from Table IV.

$$y_1 + y_2 + y_3 = (w_1 + w_2 + w_3)\, x + (b_1 + b_2 + b_3) \quad (9)$$

$$z = y_1 w_4 + y_2 w_5 + y_3 w_6 + b_4 \quad (10)$$

Our protocol communication cost is shown below. User registration phase: two Hash, one vehicle ID, one ECC multiplication operation, $4L_h + L_{id} + L_{ecc} + L_t = 468$ bytes. The protocol is computed from the authentication to the aggregation phase and the RSU broadcast phase to the on-board OBU is: $5L_h + 3L_{ecc} + L_{id} + 4L_t + L_w = 796$ bytes. The total communication cost of the protocol is computed as 1264 bytes. The protocol pair is shown in Table V.

As shown in Table III and V, the computational cost of the protocol proposed in this paper is 1.7 times higher than that in [30]; however, the communication cost in [30] is 3.9 times higher than that of the protocol in this paper. Moreover, compared to the computational costs of the authentication protocol proposed in [31]–[33], that of our proposed method is reduced by 3.7464ms, 3.3764ms, 1.3756ms; furthermore, the communication efficiency is improved by 416bit, 2618bit and 288bit.

B. Network Simulation Analysis

To analyze our protocols, the simulation uses the IoV simulation tools OMNeT++, Sumo and Veins. OMNeT++ is an extensible, modular wireless self-organizing network simulator using a C++ simulation library. Sumo is an open-source highly portable road traffic simulation package, while Veins is the middleware. We use vehicle movement tracking files using IEEE 802.11p with the DSRC standard. DSRC is a set of standards defined for vehicle data communication that supports two types of physical devices: fixed-unit RSUs, which are typically installed on the sides of the road, and onboard units (OBU), which are installed in the vehicle. The meanings of these network simulation parameters are listed in Table VI. The simulated roads use the city traffic simulation package OSMWebWizard2 that comes with Sumo. Fig. 8 shows the

TABLE VI
NETWORK SIMULATION PARAMETER DESIGN

Parameters	Values
Simulation area	25000*25000(m^2)
Path loss model	Two-Ray
Path obstacle model	Obstacle shadowing
Maximum distance	2600m
Propagation power	50mW
Data transmission rate	6Mbps
Sensitivity	-89dBm
Noise volume	-98dBm
Network use channel	CCH
Analog time	600s

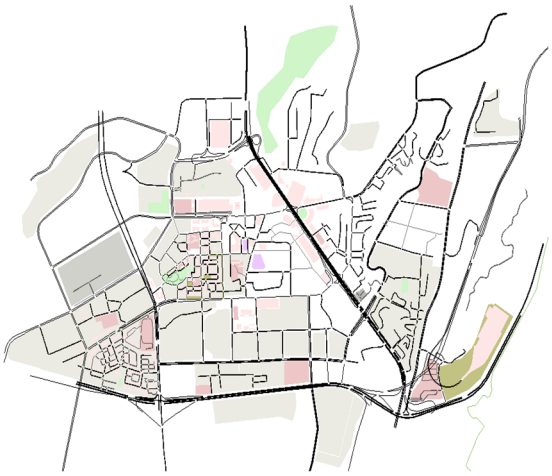


Fig. 8. SUMO simulation map (chongqing city area).

city simulation map. After we extract the map ranges, Sumo generates a map configuration (.cfg) file that is subsequently converted to a vehicle tracking file (.xml). This file is extracted as a vehicle movement tracking (.tcl) file to facilitate simulation of the protocol.

Vehicles are required to authenticate themselves each time they enter an area covered by a new RSU. The authentication delay refers to the sum of vehicle transmission delay and data verification delay, which is verified using (11).

$$AvgAuthDelay = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}\left(T_r^j - T_s^j\right)\right) \quad (11)$$

N in the above formula refers to the number of vehicle clients, n_i denotes the amount of data received C_i, and $T_r^j T_s^j$ indicates the time required to receive and send the vehicle shared data, respectively. The simulation is set up as a simulation scenario. The vehicle sends traffic environment data periodically and the simulation results are computed as shown in Fig. 9. In Fig. 9-(1), x represents the average speed of a vehicle dynamically driving at 5–30m/s (18km/h–108km/h). We observe that the increase in vehicle speed in the figure has a small effect on the vehicle authentication protocol delay and can therefore conclude that the protocol is robust. The x in Fig. 9-(2) shows the simulation results for different numbers of vehicles within different social communities. When the number of vehicles = 50, the vehicle shared protocol remains consistent with the VCC [33] protocol time for both all-static

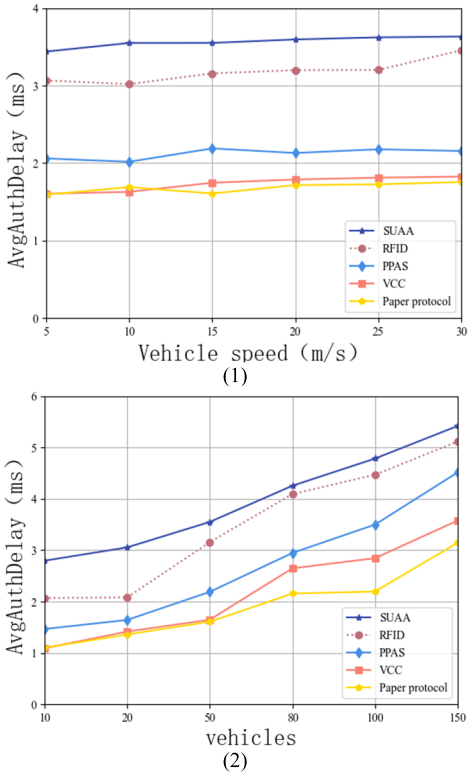


Fig. 9. Certification delay results: (1) Vehicle speed and certification delay, (2) Number of vehicles and certification delay.

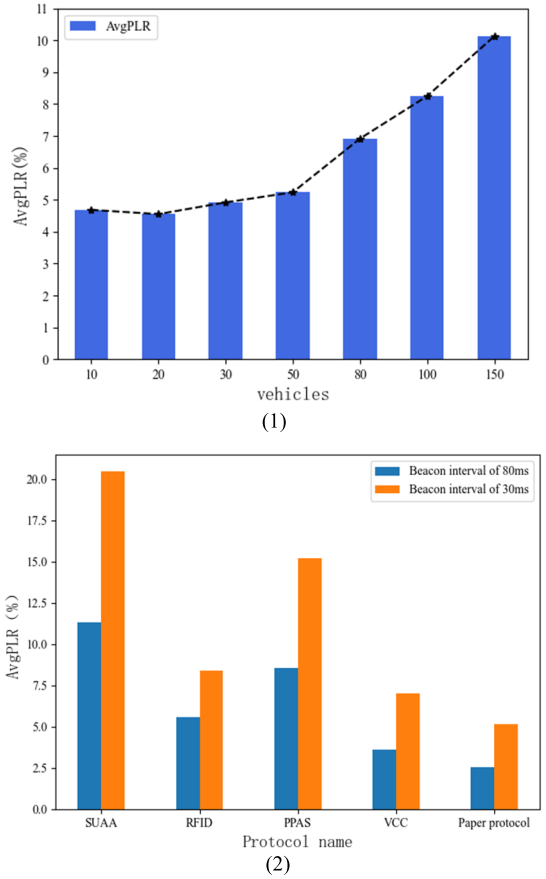


Fig. 10. Average vehicle loss rate: (1) Number of vehicles and AvgPLR, (2) Vehicle speed and AvgPLR.

and all-dynamic communities. The mixed-state community structure-wide shared collaborative protocol exhibits good performance. The packet loss rate of data transmitted and shared by vehicle clients in SIoV is the main indicator of protocol security. The packet loss rate is generated using the IEEE 802.11p protocol and (4) is used to compute the packet loss rate. N^i_{lost} within the formula indicates the number of packets lost by vehicle C_i, while $N^i_{received}$ indicates the amount of data received by the vehicle.

$$AvgPLR = \frac{1}{N}\sum_{i=1}^{N}\left\{\frac{N^i_{lost}}{N^i_{received}+N^i_{lost}}\right\} \tag{12}$$

The simulation is set up to simulate a SIoV application scenario. The vehicles drive within different social communities and send data periodically, at intervals of 30 ms and 80 ms respectively. We compute the effect of the number of vehicles on the protocol packet loss rate in different vehicle social communities. The simulation results are presented in Fig. 10. Fig. 10-(1) shows that the protocol is robust. We observe that the mixed-state social community vehicle packet loss rate exhibits an incremental increase within the permitted range when the number of vehicle users reaches 50, and that the trend is consistent with the authentication delay trend in Fig. 9-(2). We conclude that the size of the vehicle sharing data has some impact on the protocol. Fig. 10-(2) illustrates the low packet loss rate of the authentication protocol. The protocol yields satisfactory traffic data security during vehicle movement.

C. Federated Learning Collaborative Certification Protocol Convergence

This set of simulations is performed in a federated learning framework. We refer to the simulation as experimental set 1. We use ANN algorithms and MNIST data (Train/Test set images: train-images-idx3-ubyte.gz/Train/Test set labels: train-labels-idx1-ubyte.gz) to evaluate the convergence and goodness of fit of the authentication protocols in the paper. This simulation experimental set examines the convergence of the protocol assurance vehicle train w_i, which is learned by CS global aggregation followed by federated learning. The simulation experiments are chosen to represent different vehicle social community environments. We set the all-static, all-dynamic, and mixed-state community vehicle clients as clients = 10, 30, and 50 respectively, then performed cross-validation by varying the number of selected clients N. The experimental results for all groups are shown in Figs. 11 and 12. Fig. 11 shows that different loss functions and fits are applicable to different communities with respect to the vehicle training and detection sets. As the number of vehicles in the community increases, they gradually produce over-fitting. In our protocol, vehicle client upload parameters are securely aggregated at the central server content service provider. As shown in Fig. 12, moreover, the federated learning converges quickly for the global model in the all-static

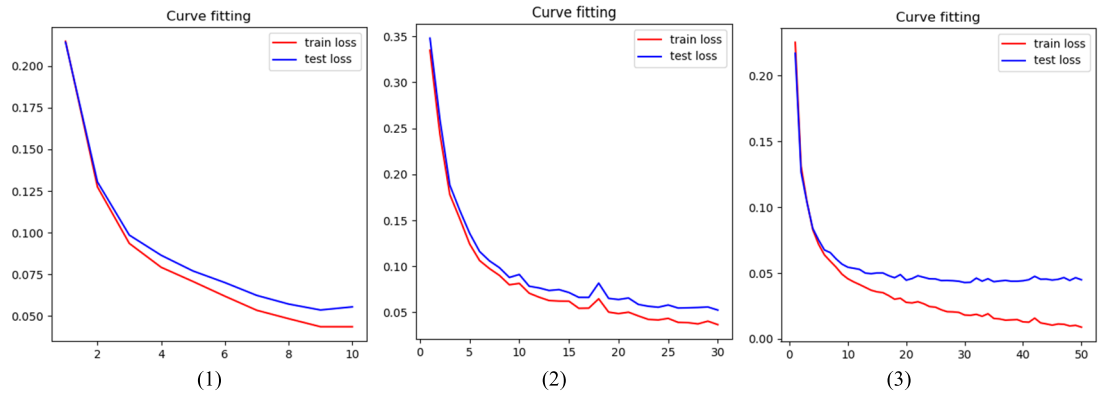


Fig. 11. Vehicle client training loss function: The vertical axis represents the value of the loss function of the training set/detection set. The horizontal axis represents the number of vehicle clients. (1) All-static community, (2) All-dynamic community, (3) Mixed-static community.

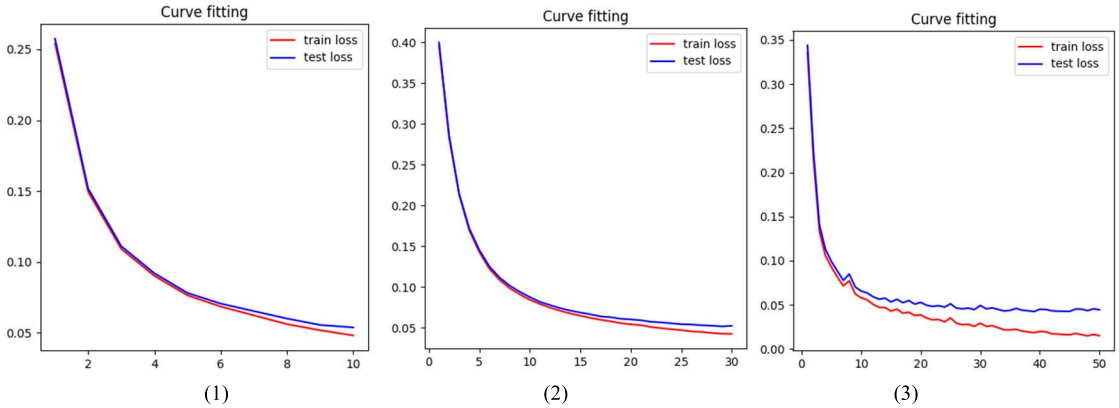


Fig. 12. Global aggregation loss function: The vertical axis represents the value of the loss function of the training set/detection set. The horizontal axis represents the number of vehicle clients. (1) All-static community, (2) All-dynamic community, (3) Mixed-static community.

Fig. 12(1) and the all-dynamic Fig. 12(2) social communities. Their training and detection set loss function fits tend to be smooth. Model Fig. 12(3) in the mixed-state community after global aggregation exhibits improved fit relative to Fig. 11(3). This simulation is a supplementary experiment. This experimental result indirectly proves that the protocol can guarantee the integrity and security of the parameters uploaded by the vehicle client. We can clearly find that the model is overfitting in the client training phase of the federated learning framework. After safe aggregation, it can be found that the model fit is enhanced. This result shows that the client weight parameters are safely shared and propagated to the content service provider to complete the aggregation. At the same time, the convergence of federated learning in different vehicle social communities meets the needs of practical applications.

VII. Conclusion

In this paper, the security issue of data sharing between vehicles and other entities in the SIoV has been effectively resolved by us. We design a federated learning collaborative authentication protocol for shared data. Meanwhile, the protocol reduces the number of vehicle certifications for each dynamic RSU, which improves the efficiency of vehicle communication. This scheme protects the safety of vehicle upload training model parameters in the federated learning architecture. In addition, anonymous mutual authentication and key agreement are considered by us in the protocol. Three groups of simulation results confirm the effectiveness of our proposed protocol in terms of efficiency and safety.

References

[1] X. Yong *et al.*, "Anonymous mutual authentication and key agreement protocol of multi-server architecture for the Internet of Vehicles," *J. Comput. Res. Develop.*, vol. 53, no. 10, pp. 2323–2333, 2016.

[2] Z. Haibo, H. Hongwu, L. Kaijian, and H. Xiaofan, "Provable and secure anonymous and traceable fast group authentication protocol in Internet of Vehicles," *J. Commun.*, vol. 42, no. 6, pp. 213–225, 2021.

[3] G. Wang, F. Xu, and C. Zhao, "QoS-enabled resource allocation algorithm in Internet of Vehicles with mobile edge computing," *IET Commun.*, vol. 14, no. 14, pp. 2326–2333, Aug. 2020.

[4] Z. Wenfang, L. Liting, W. Xiaomin, and W. Yu, "Cloud service-oriented secure and efficient certificateless aggregation signature car networking authentication key agreement protocol," *Acta Electronica Sinica*, vol. 48, no. 9, pp. 1814–1823, 2020.

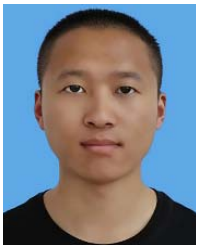
[5] I. Sarrigiannis, L. M. Contreras, K. Ramantas, A. Antonopoulos, and C. Verikoukis, "Fog-enabled scalable C-V2X architecture for distributed 5G and beyond applications," *IEEE Netw.*, vol. 34, no. 5, pp. 120–126, Sep. 2020.

[6] T. A. Butt, R. Iqbal, S. C. Shah, and T. Umar, "Social Internet of Vehicles: Architecture and enabling technologies," *Comput. Electr. Eng.*, vol. 69, pp. 68–84, Jul. 2018.

[7] Z. Ning *et al.*, "A cooperative quality-aware service access system for social Internet of Vehicles," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2506–2517, Aug. 2018.

[8] X. Wang *et al.*, "A city-wide real-time traffic management system: Enabling crowdsensing in social Internet of Vehicles," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 19–25, Sep. 2018.
[9] P. Xie, L. Xing, H. Wu, J. Seo, and I. You, "Cooperative jammer selection for secrecy improvement in cognitive Internet of Things," *Sensors*, vol. 18, no. 12, p. 4257, Dec. 2018.
[10] G. Luo *et al.*, "Software-defined cooperative data sharing in edge computing assisted 5G-VANET," *IEEE Trans. Mobile Comput.*, vol. 20, no. 3, pp. 1212–1229, Mar. 2021.
[11] Z. Zhou, C. Gao, C. Xu, Y. Zhang, S. Mumtaz, and J. Rodriguez, "Social big-data-based content dissemination in Internet of Vehicles," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 768–777, Feb. 2018.
[12] C. Chen, H. Xiang, T. Qiu, C. Wang, Y. Zhou, and V. Chang, "A rear-end collision prediction scheme based on deep learning in the Internet of Vehicles," *J. Parallel Distrib. Computing*, vol. 117, pp. 192–204, Jul. 2018.
[13] T. D. T. Nguyen, V. Nguyen, V.-N. Pham, L. N. T. Huynh, M. D. Hossain, and E.-N. Huh, "Modeling data redundancy and cost-aware task allocation in MEC-enabled Internet-of-Vehicles applications," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1687–1701, Feb. 2021.
[14] L. Zhu *et al.*, "PRIF: A privacy-preserving interest-based forwarding scheme for social Internet of Vehicles," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2457–2466, Aug. 2018.
[15] K. Deng, L. Xing, M. Zhang, H. Wu, and P. Xie, "A multiuser identification algorithm based on Internet of Things," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–11, May 2019.
[16] W. Zhou, L. Xing, J. Xia, L. Fan, and A. Nallanathan, "Dynamic computation offloading for MIMO mobile edge computing systems with energy harvesting," *IEEE Trans. Veh. Technol.*, vol. 70, no. 5, pp. 5172–5177, May 2021.
[17] J. Kang *et al.*, "Blockchain for secure and efficient data sharing in vehicular edge computing and networks," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 4660–4670, Jun. 2019.
[18] T. Wang *et al.*, "Privacy-enhanced data collection based on deep learning for Internet of Vehicles," *IEEE Trans. Ind. Informat.*, vol. 16, no. 10, pp. 6663–6672, Oct. 2020.
[19] X. Zhang, M. Peng, S. Yan, and Y. Sun, "Deep-reinforcement-learning-based mode selection and resource allocation for cellular V2X communications," *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6380–6391, Jul. 2020.
[20] Y. Lu, X. Huang, K. Zhang, S. Maharjan, and Y. Zhang, "Blockchain empowered asynchronous federated learning for secure data sharing in Internet of Vehicles," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4298–4311, Apr. 2020.
[21] S. R. Pokhrel and J. Choi, "Federated learning with blockchain for autonomous vehicles: Analysis and design challenges," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4734–4746, Aug. 2020.
[22] Y. Qi, M. S. Hossain, J. Nie, and X. Li, "Privacy-preserving blockchain-based federated learning for traffic flow prediction," *Future Gener. Comput. Syst.*, vol. 117, pp. 328–337, Apr. 2021.
[23] Y. Li, Z. Zhang, Z. Zhang, and Y.-C. Kao, "Secure federated learning with efficient communication in vehicle network," *J. Internet Technol.*, vol. 21, no. 7, pp. 2075–2084, 2020.
[24] J. Cao, K. Zhang, F. Wu, and S. Leng, "Learning cooperation schemes for mobile edge computing empowered Internet of Vehicles," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2020, pp. 1–6.
[25] S. R. Pokhrel and J. Choi, "Improving TCP performance over WiFi for Internet of Vehicles: A federated learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 6, pp. 6798–6802, Jun. 2020.
[26] D. He, S. Zeadally, B. Xu, and X. Huang, "An efficient identity-based conditional privacy-preserving authentication scheme for vehicular ad hoc networks," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2681–2691, Dec. 2015.
[27] D. Genkin, L. Pachmanov, I. Pipman, E. Tromer, and Y. Yarom, "ECDSA key extraction from mobile devices via nonintrusive physical side channels," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, Oct. 2016, pp. 1626–1638.
[28] Y. Jiang, M. Shi, X. Shen, and C. Lin, "BAT: A robust signature scheme for vehicular networks using binary authentication tree," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1974–1983, Apr. 2009.
[29] B. Ying and A. Nayak, "Anonymous and lightweight authentication for secure vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 10626–10636, Dec. 2017.
[30] N. M. R. Lwamo, L. Zhu, C. Xu, K. Sharif, X. Liu, and C. Zhang, "SUAA: A secure user authentication scheme with anonymity for the single & multi-server environments," *Inf. Sci.*, vol. 477, pp. 369–385, Mar. 2019.
[31] X. Jian, L. Wenjiang, G. Hongyang, and Z. Yingbo, "An RFID security authentication protocol that can resist DoS attacks in the Internet of Vehicles," *J. Beijing Univ. Posts Telecommun.*, vol. 42, no. 2, pp. 114–119, 2019.
[32] H. Zhu, T. Liu, G. Wei, and H. Li, "PPAS: Privacy protection authentication scheme for VANET," *Cluster Comput.*, vol. 16, no. 4, pp. 873–886, Dec. 2013.
[33] T. Limbasiya and D. Das, "Secure message confirmation scheme based on batch verification in vehicular cloud computing," *Phys. Commun.*, vol. 34, pp. 310–320, Jun. 2019.
[34] S. Yu, X. Chen, Z. Zhou, X. Gong, and D. Wu, "When deep reinforcement learning meets federated learning: Intelligent multitimescale resource management for multiaccess edge computing in 5G ultradense network," *IEEE Internet Things J.*, vol. 8, no. 4, pp. 2238–2251, Feb. 2021.
[35] C. Bing *et al.*, "Overview of federal learning security and privacy protection," *J. Nanjing Univ. Aeronaut. Astronaut.*, vol. 52, no. 5, pp. 675–684, 2020.
[36] Z. Yu, J. Hu, G. Min, Z. Zhao, W. Miao, and M. S. Hossain, "Mobility-aware proactive edge caching for connected vehicles using federated learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 8, pp. 5341–5351, Aug. 2021.
[37] F. O. Olowononi, D. B. Rawat, and C. Liu, "Federated learning with differential privacy for resilient vehicular cyber physical systems," in *Proc. IEEE 18th Annu. Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2021, pp. 1–5.
[38] K. Wei *et al.*, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3454–3469, 2020.
[39] S. Abbasi, A. M. Rahmani, A. Balador, and A. Sahafi, "Internet of Vehicles: Architecture, services, and applications," *Int. J. Commun. Syst.*, vol. 34, no. 10, p. e4793, Jul. 2021.
[40] B. Vaidya and H. T. Mouftah, "IoT applications and services for connected and autonomous electric vehicles," *Arabian J. Sci. Eng.*, vol. 45, no. 4, pp. 2559–2569, Apr. 2020.
[41] S. Yu, J. Lee, K. Park, A. K. Das, and Y. Park, "IoV-SMAP: Secure and efficient message authentication protocol for IoV in smart city environment," *IEEE Access*, vol. 8, pp. 167875–167886, 2020.
[42] C.-M. Chen, B. Xiang, Y. Liu, and K.-H. Wang, "A secure authentication protocol for Internet of Vehicles," *IEEE Access*, vol. 7, pp. 12047–12057, 2019.
[43] Y.-Y. Zhang, J. Shang, X. Chen, and K. Liang, "A self-learning detection method of Sybil attack based on LSTM for electric vehicles," *Energies*, vol. 13, no. 6, p. 1382, Mar. 2020.
[44] S. Agrawal, M. L. Das, and J. Lopez, "Detection of node capture attack in wireless sensor networks," *IEEE Syst. J.*, vol. 13, no. 1, pp. 238–247, Mar. 2019.
[45] A. M. Vegni and V. Loscrí, "A survey on vehicular social networks," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2397–2419, 4th Quart., 2015.
[46] A. M. Vegni, C. Souza, V. Loscri, E. Hernandez-Orallo, and P. Manzoni, "Data transmissions using hub nodes in vehicular social networks," *IEEE Trans. Mobile Comput.*, vol. 19, no. 7, pp. 1570–1585, Jul. 2020.

Pengcheng Zhao (Graduate Student Member, IEEE) received the B.S. degree in automation from the Luoyang Institute of Technology, China, in 2017, and the M.S. degree in software engineering from the Henan University of Science and Technology, China, in 2020, where he is currently pursuing the Ph.D. degree in control theory engineering. His research interests include data security and privacy, and cyberspace security.



Yuanhao Huang received the B.E. degree in computer science and technology from the Zhengzhou University of Science and Technology, China. He is currently pursuing the M.S. degree in electronic information with the Henan University of Science and Technology, China. His research interests include social network user identification and user alignment.

Jianping Gao was born in 1977. He received the B.S. degree in automotive engineering from the Luoyang Institute of Technology in 2000, and the M.S. degree in mechanical engineering and the Ph.D. degree in vehicle engineering from the Beijing Institute of Technology in 2003 and 2009, respectively. He completed his Ph.D. Joint Training Program at Michigan State University in 2008. His main research interests are new energy vehicles and intelligent networked vehicle technology.

Ling Xing (Member, IEEE) was born in 1978. She received the B.S. degree in electrical engineering from the Southwest University of Science and Technology, China, in 2002, the M.S. degree in communication and information systems from the University of Science and Technology of China in 2005, and the Ph.D. degree in communication engineering from the Beijing Institute of Technology in 2008. She was a Visiting Scholar at the Illinois Institute of Technology, USA, in 2008. Her main research interests are semantic multimedia, privacy preservation, and trusted computing.

Honghai Wu was born in 1979. He received the B.S. degree in industrial automation from Zhengzhou University in 2001, and the M.S. degree in signal and signal processing and the Ph.D. degree in computer science and technology from the Beijing University of Posts and Telecommunications in 2007 and 2015, respectively. He has participated in and chaired the National Natural Science Foundation of China. His research interests are in mobile multimedia computing and mobile edge computing.

Huahong Ma was born in 1979. She received the B.S. degree from Zhengzhou University, China, in 2001, the M.S. degree from Yunnan University, China, in 2005, and the Ph.D. degree from the Henan University of Science and Technology, Luoyang, China. Since 2005, she has been working at the College of Information Engineering, Henan University of Science and Technology. Now, she is an Associate Professor. Her main research interests are crowd sensing networks and the Internet of Things.