

# MetaTrading: An Immersion-Aware Model Trading Framework for Vehicular Metaverse Services

Hongjia Wu<sup>ID</sup>, Hui Zeng<sup>ID</sup>, Zehui Xiong<sup>ID</sup>, Jiawen Kang<sup>ID</sup>, Zhiping Cai<sup>ID</sup>, Tse-Tin Chan<sup>ID</sup>, Member, IEEE, Dusit Niyato<sup>ID</sup>, Fellow, IEEE, and Zhu Han<sup>ID</sup>, Fellow, IEEE

**Abstract**—Timely updating of Internet of Things (IoT) data is crucial for achieving immersion in vehicular metaverse services. However, challenges such as latency caused by massive data transmissions, privacy risks associated with user data, and computational burdens on metaverse service providers (MSPs) hinder the continuous collection of high-quality data. To address these challenges, we propose an immersion-aware model trading framework that enables efficient and privacy-preserving data provisioning through federated learning (FL). Specifically, we first develop a novel multi-dimensional evaluation metric for the immersion of models (IoM). The metric considers i) the freshness and accuracy of the local model, and ii) the amount and potential value of raw training data. Building on the IoM, we design an incentive mechanism to encourage metaverse users (MUs) to participate in FL by providing local updates to MSPs under resource constraints. The trading interactions between MSPs and MUs are modeled as an equilibrium problem with equilibrium constraints (EPEC) to analyze and balance their costs and gains, where MSPs as leaders determine rewards, while MUs as followers optimize resource allocation. To ensure privacy and adapt to dynamic network conditions, we develop a distributed dynamic reward algorithm based on deep reinforcement learning, without acquiring any private information from MUs and other MSPs. Experimental results show that the proposed framework outperforms state-of-the-art benchmarks, achieving improvements in IoM of 38.3% and 37.2%, and reductions in training time to reach the target accuracy of 43.5% and

Received 15 April 2024; revised 6 August 2025; accepted 7 September 2025. Date of publication 15 September 2025; date of current version 11 December 2025. This work was supported in part by NSF under Grant ECCS-2302469 and Grant CMMI-2222810; in part by Toyota; in part by Amazon; in part by the Japan Science and Technology Agency (JST) Adopting Sustainable Partnerships for Innovative Research Ecosystem (ASPIRE) under Grant JPMJAP2326; in part by NSFC under Grant 62472434; in part by the Science and Technology Innovation Program of Hunan Province under Grant 2022RC3061; and in part by the Research Matching Grant Scheme from the Research Grants Council of Hong Kong. (*Corresponding author: Tse-Tin Chan*)

Hongjia Wu and Tse-Tin Chan are with the Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong, China (e-mail: whongjia@edu.hk; tsetinchan@edu.hk).

Hui Zeng and Zhiping Cai are with the College of Computer Science and Technology, National University of Defense Technology, Qinhuangdao 066104, China (e-mail: zenghui116@nudt.edu.cn; zpc@nudt.edu.cn).

Zehui Xiong is with the Pillar of Information Systems Technology and Design, Singapore University of Technology and Design, Singapore 487372 (e-mail: zhuixiong@sutd.edu.sg).

Jiawen Kang is with the School of Automation, Guangdong University of Technology, Guangzhou 510520, China (e-mail: kavinkang@gdut.edu.cn).

Dusit Niyato is with the College of Computing and Data Science, Nanyang Technological University, Singapore 639798 (e-mail: dniyato@ntu.edu.sg).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea (e-mail: hanzhu22@gmail.com).

This article has supplementary downloadable material available at <https://doi.org/10.1109/TSC.2025.3609730>, provided by the authors.

Digital Object Identifier 10.1109/TSC.2025.3609730

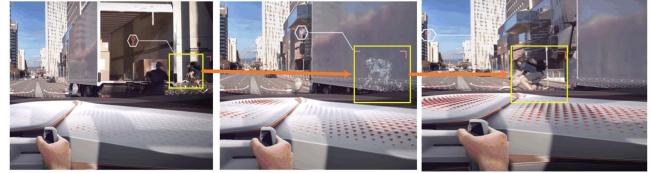


Fig. 1. An example of AR services in the vehicular metaverse<sup>1</sup>.

49.8%, on average, for the MNIST and GTSRB datasets, respectively. These findings validate the effectiveness of our approach in incentivizing MUs to contribute high-value local models to MSPs, providing a flexible and adaptive scheme for data provisioning in vehicular metaverse services.

**Index Terms**—Equilibrium problem with equilibrium constraints, immersion-aware, incentive mechanism, resource allocation, vehicular metaverse.

## I. INTRODUCTION

THE metaverse, envisioned as a major evolutionary step for the Internet, aims to create a fully immersive, self-sustaining virtual environment for activities such as playing, working, and socializing [1]. This vision is propelled by advances in fifth- and sixth-generation (5G/6G) communication technologies, which offer low latency and high data throughput. These technologies play a critical role in seamlessly integrating the Internet of Things (IoT) data into metaverse services, thus bringing the once-fictional concept of immersive experiences closer<sup>1</sup> to reality.

Metaverse services are beginning to reveal their vast potential across a broad spectrum of industries, from gaming and autonomous driving to education and marketing. Notably, the application of vehicles within the metaverse has attracted significant interest, particularly for the enhanced traffic safety enabled by state-of-the-art augmented reality (AR) technologies. Market report [2] projected that the global automotive metaverse market would grow from \$1.9 billion in 2022 to \$6.5 billion by 2030. Automakers such as BMW are actively investing in AR technologies. As shown in Fig. 1, augmented information improves driving safety by displaying potential hazards obscured by the vehicle ahead. Moreover, Nissan's [3] upcoming technology utilizes a 3D AR interface that merges the real world with the

<sup>1</sup>Image source: <https://www.jasoren.com/ar-in-automotive/>

virtual world to provide the driver with augmented information about the surrounding area. In addition, Neuron mobility<sup>2</sup> has announced the launch of an innovative AR parking assistance system to improve passengers' parking routines and trip-end experiences.

The ability to capture information from the "real" world, particularly the ability to collect and process massive amounts of data from IoT devices, is the key to determining the success of immersive services (e.g., AR) in the vehicular metaverse. Meanwhile, the data must be processed and presented in a meaningful and responsive manner with adequate privacy protection. Technically, a high-quality AR experience relies on accurately detecting and classifying real-world objects (e.g., cars and pedestrians) under complex conditions [4]. To achieve this goal, sufficient valid data need to be collected and deeply processed to detect and classify objects accurately. Therefore, it is essential to focus on effectively collecting, processing, and protecting the data that supports a safer and more enjoyable driving experience.

**Motivations:** The widely used data collection method, as adopted by Nissan [5] and studied in [6], involves gathering massive amounts of data through vehicle sensors, cameras, and roadside devices, and all data being processed centrally. While this approach is effective for various applications, when it comes to the situation with multiple metaverse users (MUs) and metaverse service providers (MSPs) associated with different companies, the centralized data collection approach is not applicable and may lead to the following issues. First, AR services in the vehicular metaverse need to be highly immersive so that MUs feel fully immersed in the rendered environment, such as visualized driving. However, data synchronization is hindered by the latency from massive real-time data updates under unstable and resource-limited communication conditions. Note that the value of real-time data diminishes over time [7]. Also, delays can severely impact the MU's experience and cause dizziness [8]. Second, the data to be transmitted may be sensitive and private, such as location, movement, and biometrics, which can create a better immersive experience but may inevitably increase the privacy risk of MUs.

Federated learning (FL) has been adopted in prior work [4], [9], [10], [11] to enable collaborative model training without sharing raw data (e.g., sensor/imaging data from vehicles). In this approach, local updates are uploaded by individual MUs to MSPs for AR services, rather than transmitting large volumes of data centrally, thus significantly reducing the communication burden. However, MUs are typically self-interested and are reluctant to share local models with MSPs due to the additional computation, communication, and energy overhead. To overcome this issue, incentive mechanisms using strategies such as contract theory [12], Stackelberg game [13], [14] and multi-winner sealed-bid auction [15], have been proposed to encourage MUs to contribute local models. However, existing studies fail to explicitly assess the model value from multiple dimensions, making it difficult for both MSPs and MUs to

<sup>2</sup><https://immersive-technology.com/augmentedreality/neuron-introduces-new-ar-parking-assistant/>

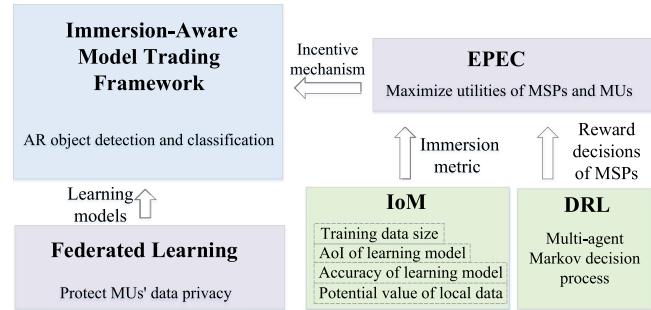


Fig. 2. The outline of an immersion-aware framework for FL-assisted vehicular metaverse.

quantify the benefits of the models for MSPs. Furthermore, most of these solutions are designed for a single MSP and do not consider the joint optimization of MUs' limited computational and communication resources. As a result, an efficient and privacy-preserving framework for data synchronization is needed to improve the immersive AR experience in the vehicular metaverse.

To address the above research gaps, we propose an immersion-aware model trading framework for a multi-MSP, multi-MU vehicular metaverse. This framework is designed to incentivize MUs to become active contributors by providing local models tailored to the specific needs of MSPs. However, MUs have different sampling costs and limited computational and communication resources, while MSPs differ in their model preferences and compete with each other. These inherent asymmetries and competitive dynamics make it challenging to establish an efficient model trading ecosystem. This leads to two key research questions that underpin our work:

- How can MUs be effectively incentivized to contribute high-value local models that benefit vehicular metaverse services for diverse MSPs?
- How can the dynamic competitive interactions among MSPs be modeled to achieve a trade-off in gains between MUs and MSPs, ensuring sustainable and stable trading of local models in the vehicular metaverse?

**Proposed Framework and Contributions:** As depicted in Fig. 2, our model trading framework consists of four components: FL mechanism, metric design, game modeling, and dynamic algorithms. Specifically, this work first designs a new metric called the *immersion of the learning model* (IoM) to evaluate the value of local models contributed by MUs to MSPs. This metric jointly considers the freshness and accuracy of the local model, as well as the amount and potential value of raw training data. Building on this metric, we propose an immersion-aware incentive mechanism that aligns the interests of both MUs and MSPs. Then, we model the dynamic competitive trading interactions as an equilibrium problem with equilibrium constraints (EPEC), which is a hierarchical optimization problem with equilibria at two levels [16]. Moreover, given the dynamic networks and the privacy concerns of MSPs, we formulate the reward decisions of MSPs as a multi-agent Markov decision process (MAMDP) and develop a multi-agent deep reinforcement learning (DRL)-based dynamic reward (MDDR) approach

to obtain the reward decisions in a fully distributed manner. In summary, the key contributions of this work are as follows.

- *Incentive mechanism design for immersion-aware model trading:* From the perspectives of both MUs and MSPs, we propose an incentive mechanism that encourages MUs to contribute high-value local models tailored to the specific demands of MSPs. To our knowledge, this is the first study focusing on incentive mechanism design for efficient and privacy-preserving data synchronization in multi-MU and multi-MSP vehicular metaverse environments.
- *Novel design of multi-dimensional metric:* To quantify immersion enhancement provided by MUs for AR services, we design an immersion metric of the local model integrating four critical dimensions: the freshness and accuracy of the local model, as well as the amount and potential value of raw training data. Freshness is captured through age of information (AoI), while potential value is evaluated by the difference between model predictions and true labels. This metric enables fine-grained evaluation of model value under resource constraints and provides a basis for incentive and decision-making strategies in model trading.
- *Theoretical analysis and algorithm designs:* Given multiple resource-constrained MUs and competing MSPs, we model their interactions as an equilibrium problem with equilibrium constraints (EPEC). We theoretically prove the existence and uniqueness of the equilibrium, and further develop a fully distributed MDDR approach that adapts to dynamic environments and operates without accessing any private information of MUs or MSPs.
- *Performance evaluation:* Extensive numerical simulations are conducted based on AR-related vehicle datasets to validate the efficacy and efficiency of MDDR and the proposed immersion-aware model trading framework. Numerical results show that our proposed mechanism improves the IoM by 38.3% and 37.2%, and reduces the training time to reach the target accuracy by 43.5% and 49.8%, on average, for the MNIST and GTSRB datasets, respectively, compared with state-of-the-art benchmarks.

The rest of this paper is organized as follows. Section II discusses the related work. In Section III, we present the system overview and design the immersion metric of the local model. Section IV gives the game formulation, and Section V analyzes the existence of the equilibria at two levels. In Section VI, we give the detailed design of MDDR. Section VII shows numerical experiments to evaluate the framework performance, and finally, Section VIII concludes the paper.

## II. RELATED WORK

In this section, we discuss the work related to our study in terms of edge-enabled vehicular metaverse services, FL for AR, and incentive mechanisms for data synchronization in the metaverse.

### A. Edge-Enabled Vehicular Metaverse Services

Vehicular metaverse services require ultra-low latency and high reliability to enable immersive experiences, such as AR

navigation and real-time digital twins. To meet these demands, researchers have extensively explored edge-based strategies for task offloading and resource management. For instance, Feng et al. [17] proposed an innovative resource allocation framework for AR-empowered vehicular edge metaverse, which significantly improves data utility and reduces vehicle energy consumption through multi-dimensional optimization. Similarly, Khan et al. [18] introduced a cooperative framework integrating task offloading, sensing, learning, and communication to reduce both transmission energy consumption and latency in resource-constrained environments. To cope with dynamic network conditions, Tong et al. [19] proposed an attribute-aware auction mechanism for resource allocation, and applied a GPT-based deep reinforcement learning algorithm to adjust the auction clocks. Regarding security and privacy, Kang et al. [20] presented a cross-metaverse empowered dual pseudonym management framework to mitigate privacy leakage risks during the dynamic communications among vehicular edge metaverses. Moreover, a novel four-layer security framework was designed to protect vehicular metaverse systems from data poisoning attacks [21].

In addition to theoretical models and optimization frameworks for integrating autonomous vehicles into the metaverse, several studies have proposed practical decision-making approaches to evaluate and select among alternative implementation strategies. For example, Deveci et al. [22] proposed a framework that combines a fuzzy logarithmic weighting method with a similarity-based ranking technique to assess four traffic safety solutions enabled by the metaverse. Their results show that public transportation is the most practical and scalable option. Moreover, they applied a decision-making method based on Q-rung orthopair fuzzy sets to rank personal mobility alternatives, offering practical guidance for system planning in [23]. Gokasar et al. [6] introduced a two-stage fuzzy decision-making framework to evaluate integration options for self-powered sensors in autonomous vehicles. The study concludes that using these sensors for real-time traffic management in the metaverse is the most effective solution.

These works, from diverse perspectives, have significantly advanced the security, reliability, and performance of vehicular metaverse services, providing valuable insights for future research. Notably, distinct from existing studies, our work focuses on designing an incentive mechanism that transforms MUs into active contributors. These contributors efficiently provide high-value local models to MSPs, thereby enhancing the overall service experience.

### B. Federated Learning for AR in the Metaverse

Privacy preservation is critical in vehicular metaverse services, particularly when training models using sensitive user data. Federated learning (FL) has emerged as a promising technique to enable collaborative model training without sharing raw data. For example, Chen et al. [9] proposed a framework that integrates mobile edge computing and FL to solve the computational efficiency and low-latency object classification problem for AR services. However, they did not address how

communication and computational resources should be allocated within the FL framework.

To overcome this limitation, Zhou et al. [10] developed an FL-assisted MAR system via non-orthogonal multiple access for the metaverse. Furthermore, a resource allocation algorithm was designed to balance accuracy and energy consumption. To further improve learning efficiency, an adaptive and resource-efficient FL algorithm tailored for AR applications was proposed in [4]. This approach mitigates the effects of non-IID data distributions, reduces resource usage, and enhances the quality of experience. Additionally, Hazarika et al. [11] explored the integration of quantum computing with FL to provide a cost-efficient and adaptive solution for the dynamic nature of vehicular environments. Nadimi et al. [24] proposed a multi-modal, multi-task federated foundation model (FedFMs) architecture. It combines the privacy-preserving benefits of FL with the representation capabilities of foundation models to address sensor diversity, personalized interactions, and device heterogeneity in extended reality systems.

Incorporating FL into AR services brings the benefits of reduced communication latency and privacy protection. However, the existing studies fail to consider the integration of FL into AR for the multi-MSP and multi-MU metaverse scenario, while overlooking the selfishness of MUs, i.e., whether they are willing to participate in FL learning.

### C. Incentive Mechanism for Data Synchronization in the Metaverse.

Immersive metaverse services depend on high-quality real-world sensing data. However, the process of collecting such data is often costly and may raise privacy concerns. To this end, various incentive mechanisms have been proposed to encourage MUs to contribute data and participate in model training. For instance, Zhang et al. [25] introduced a vehicle-assisted data sensing framework that incentivizes vehicles to upload sensing data, enhancing the driving experience. Xu et al. [26] developed a UAV swarm system using digital twins and semantic communication to improve data synchronization and reduce latency. Lin et al. [27] designed an incentive-based congestion control scheme for digital twin edge networks, which addresses the stochastic service demand and the long-term provider profit by combining Lyapunov optimization and contract theory. Zhao et al. [28] proposed a physical reality incentive mechanism based on a two-stage Stackelberg game to motivate MUs to collect physical reality sensing data for the metaverse. Although these studies successfully encourage raw data contributions, they also pose a risk to user privacy due to potential data leakage.

To address privacy concerns, Kang et al. [12] developed a cross-chain FL framework with an AoI-based contract theory model under prospect theory (PT) to incentivize data sharing. They utilized a hierarchical cross-chain architecture with a main chain and multiple subchains to perform decentralized, privacy-preserving, and secure data training in both virtual and physical spaces. Baccour et al. [13] designed a dual game-theoretic framework for federated meta-learning (FML), incorporating a reputation system and Stackelberg game to enhance privacy,

TABLE I  
INCENTIVE MECHANISMS FOR DATA SYNCHRONIZATION

Ref.	MSPs MUs	Theory Method	Main Metric	Data Privacy
[12]	×	Contract theory under PT	AoI	✓
[13]	×	Dual game-theoretic	Reputation scores	✓
[14]	×	Stackelberg game	Model quality and latency	✓
[15]	✓	IISG and PMS-AM	Model quality and freshness	✓
[25]	✓	Coalition formation, Stackelberg game	Volume of sensing data	✗
[26]	✓	Deep learning based auction	Quality, reliability, latency	✗
[27]	✓	Contract theory	Computing delay	✗
[28]	✓	Stackelberg game	Energy consumption	✗
[29]	✓	Multi-objective optimization	Peak signal-to-noise ratio	✓
Our work	✓	EPEC	IoM	✓

reduce energy consumption, and improve model performance. Li et al. [14] designed a new metric named “Satisfaction” to balance training latency and model quality. They integrated this metric into utility functions to incentivize node participation in FL while optimizing resource allocation. However, these works are designed for a single MSP scenario, limiting scalability and applicability in multi-server scenarios commonly encountered in large-scale metaverse environments.

To fill this gap, Zhang et al. [15] employed an imperfect information Stackelberg game (IISG) to optimize the strategies of MUs and MSPs, and introduced a privacy-preserving multi-winner sealed-bid auction mechanism (PMS-AM) to enhance the quality of non-fungible token (NFT) with FL assistance. Esmaeil et al. [29] proposed a metaverse FL and semantic communication (MFSC) framework that matches IoT device capabilities to training tasks, optimizing digital twin quality through adaptive resource allocation. Nonetheless, existing incentive criteria are often simplistic and fail to evaluate the values of contributed models comprehensively.

Unlike these prior works, our study focuses on vehicular metaverse services in multi-MSP and multi-MU scenarios. We propose a novel metric, IoM, which evaluates model value from multiple dimensions. Based on this metric, we design an incentive mechanism that better motivates MUs to contribute high-value local models, thereby improving the quality and user experience of metaverse services. A detailed comparison of these approaches is provided in Table I, highlighting the key contributions of this study.

### III. SYSTEM MODEL

We first present the<sup>3</sup> system overview of the immersion-aware model trading framework in Section III-A. Next, we introduce the FL mechanism adopted by our framework in

<sup>3</sup>Image source: <https://www.vanarama.com/blog/cars/4-ways-augmented-reality-will-revolutionise-the-automotive-industry>.

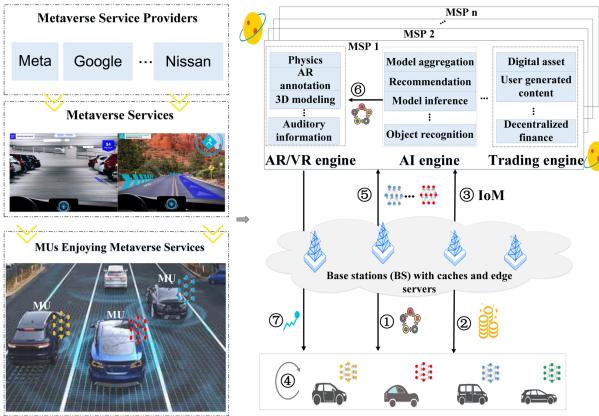


Fig. 3. Workflow of the immersion-aware model trading framework<sup>3</sup>.

Section III-B and design the immersion metric of the local model in Section III-C.

#### A. System Overview

Fig. 3 illustrates the workflow of our immersion-aware model trading framework, designed to enhance the immersive and interactive AR experiences of MUs. MUs can contribute local models to the AR services of MSPs for better immersive experiences. For example, services such as object detection can be improved by training on street and pedestrian images captured by vehicles' cameras. In this way, AR services can be applied more effectively to visible driving, with more accurate and timely hazard warnings from windscreens. The framework mainly consists of MSPs owned by different companies (e.g., Meta, Nissan, and Google), an infrastructure layer, and an interaction layer with the MUs that enjoy vehicular metaverse services. MSPs are powered by various technologies supporting vehicular metaverse services, such as AI engines, AR/VR engines, and metaverse trading engines. The infrastructure layer with base stations (equipped with caches and edge servers) provides the basis of 5G/6G communication services for the interactions among MUs and MSPs. MUs play essential roles as both consumers and contributors to vehicular metaverse services.

The framework adopts a multi-engine architecture, and this paper focuses on two components: the trading engine and the AI engine of the MSP. The trading engine is responsible for incentivizing MUs to participate in mutually beneficial collaborations with MSPs, providing them with a trading platform to make their decisions (i.e., computational and communication resource allocation of MUs and reward decisions of MSPs). Then, each MU deciding to contribute utilizes the FL mechanisms integrated into AI engines to provide local models for MSPs under the strategic guidance of the trading results. The specific workflow of the framework involves the following two phases.

**Phase I (Incentive process):** Initially, MSPs' AI engines broadcast their global models to all MUs in the vehicular metaverse<sup>4</sup> (① in Fig. 3). Then, the trading engines of MSPs

<sup>4</sup>The FL global models owned by different MSPs we consider could be object detection or classification models for intelligent terrain mapping, intuitive road

TABLE II  
DEFINITIONS OF NOTATION

Notation	Definition
$m, M$	Index of an MU, Number of MUs
$n, N$	Index of an MSP, Number of MSPs
$\Delta_{mn}$	The average AoI of MU $m$ 's model provided for MSP $n$
$V_{mn}$	The IoM contributed by MU $m$ to MSP $n$
$f_{mn}$	MU $m$ 's computational resource used for MSP $n$
$B_{mn}$	MU $m$ 's communication resource used for MSP $n$
$X_{mn}$	MU $m$ 's training data set for MSP $n$
$\omega_{mn}$	The potential value of MU $m$ 's local data for MSP $n$
$\theta_m$	MU $m$ 's local accuracy threshold

determine their digital currency prices per IoM, i.e., rewards to MUs, and broadcast them to MUs (②). Next, each MU determines its allocation of computational and communication resources for contributing local models to various MSPs based on the IoM, costs, and rewards given by MSPs (③). Based on the responses from MUs, the trading engines of MSPs adjust the digital currency prices to maximize their own utilities. Steps ② and ③ iterate until an agreement is reached between MUs and MSPs.

**Phase II (FL process):** Each MU, guided by decisions obtained through tradings, allocates computational resources to local training based on different global models from MSPs and generates local models for different MSPs (④). Then, MUs upload their local updates to the corresponding AI engines of MSPs by consuming communication resources (⑤). Following this, AI engines aggregate local updates from all MUs to update their global models and transmit them to their AR/VR engines for utilization (⑥). Finally, the AR/VR engines provide enhanced AR services to MUs (⑦). Steps ④-⑦ are repeated until the guidance time  $T$  ends. To ease the presentation, we summarize some important notations in Table II.

#### B. FL Framework Mechanism Adopted

We consider a set of  $\mathcal{M} = \{1, \dots, M\}$  MUs equipped with local computational capabilities and a set of  $\mathcal{N} = \{1, \dots, N\}$  MSPs with FL synchronous tasks<sup>5</sup>. Each MSP initiates an FL synchronous task with a virtual deadline  $\tau_n$  ( $n \in \mathcal{N}$ ). The MUs can participate in FL to generate local models and contribute to enhancing vehicular metaverse services. As shown in Fig. 4, MU  $m$  provides local models to different MSPs simultaneously for a given time period  $T$ , meaning multiple FL tasks can be processed in parallel. Moreover, the data collected at the  $(r - 1)$ -th round are used for local training at the  $r$ -th round.

Without loss of generality, we consider that the trading decisions among MUs and MSPs are made at the beginning of period  $T$ .

Let  $X_{mn} = \{(y_i, z_i)\}_{i=1}^{|X_{mn}|}$  be the set of input-output pairs sampled from MU  $m$  for task  $n$ , where  $y_i$  is the input with

safety, visible driving, etc. Note that it is feasible to train on multiple tasks with the same local data. For instance, for the same image, intuitive road safety focuses on pedestrians and vehicles, while intelligent terrain mapping focuses on landmarks and routes.

<sup>5</sup>That is, the update time of the global model is limited by the slowest MU, i.e., the AI engine of the MSP must wait to receive local updates from all MUs before performing model aggregation. The advantages of synchronous FL are high model accuracy and fast convergence [30].

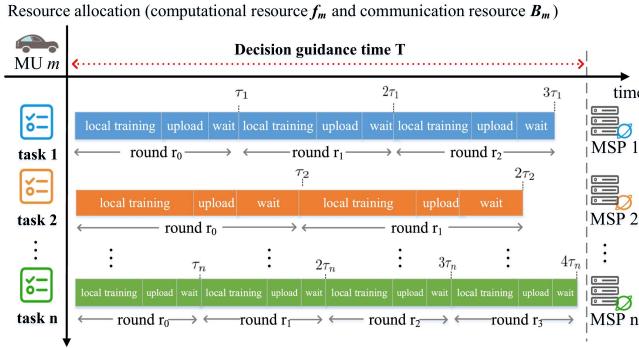


Fig. 4. Illustration of the FL mechanism with  $n$  tasks for MU  $m$ .

$d$  features and  $z_i$  is its corresponding ground-truth label. Here,  $|X_{mn}|$  denotes the size of the set  $X_{mn}$ . The data, such as images of streets and pedestrians, can be generated by high-definition cameras inside and outside the vehicles. The steps involved in each iteration  $r \in \{1, 2, \dots, R\}$  are as follows:

- *Publishing Tasks:* The AI engine of MSP  $n$  broadcasts its global model  $j_n^{(r)}$  (e.g., object detection models used in AR services for visible driving and intuitive road safety) to all MUs in the  $r$ -th round.
- *Local Training:* Each MU, for example, MU  $m$ , trains  $j_n^{(r)}$  with the most recently collected data  $X_{mn}^{(r)}$  by stochastic gradient descent (SGD) within certain local rounds. The number of local training rounds depends on its local accuracy threshold.
- *Uploading and Aggregation:* MUs transmit their local updates  $j_{mn}^{(r)}$  to the AI engine of MSP  $n$ . Upon receiving all the local updates, the MSP  $n$ 's AI engine aggregates them through the following weighted average:

$$j_n^{(r+1)} = j_n^{(r)} + \sum_{m=1}^M \frac{|X_{mn}^{(r)}|}{|X_n^{(r)}|} (j_{mn}^{(r)} - j_n^{(r)}), \quad (1)$$

and then obtains the new global model  $j_n^{(r+1)}$  for the next iteration.

- *Model Utilization:* After global aggregation, MSP  $n$ 's AR engine obtains an updated global model that can be used for corresponding 3D modeling and other functions [9]. The global rounds are iterated until a specific requirement is met, such as reaching a certain level of accuracy or a deadline.

### C. Immersion Metric of Local Model

MSPs determine the rewards for MUs based on the immersion of the local model (IoM) metric. Correspondingly, MUs need to decide the IoMs of local models provided to MSPs by allocating their computational and communication resources to maximize the benefits. This process highlights the key role of “IoM” in the interactions between the two parties, serving as both an evaluation criterion and a basis for decision-making. Unlike traditional FL and incentive mechanisms, vehicular metaverse scenarios

with immersion requirements necessitate consideration of both the contribution and freshness of the local model,<sup>6</sup> which will collectively affect the immersive experience of AR services, i.e., whether virtual objects can be placed in the physical world accurately and promptly.

Driven by the above considerations, we design the IoM based on four key dimensions: the freshness and accuracy of the local model, as well as the amount and potential value of raw training data. The freshness is captured through *age of information* (AoI)  $\Delta_{mn}$  [31], while the other three dimensions are jointly reflected in *contribution prediction*  $I_{mn}$ . To measure the value of a local model that MU  $m$  brought to the vehicular metaverse service of MSP  $n$ , we define the IoM as

$$V_{mn} = I_{mn}(\tau_n - \bar{\Delta}_{mn}), \quad (2)$$

where the components are detailed as follows. In the rest of this paper, we refer to the trading engine of MSP as MSP.

1) *Contribution Prediction I:* The contribution prediction  $I_{mn}$  from MU  $m$  to MSP  $n$  is determined by the accuracy  $\theta_m$  of the local model, the total amount of training data  $\lfloor \frac{T}{\tau_n} \rfloor |X_{mn}|$ , and the potential value  $\omega_{mn}$  of local data. Prior work [32] characterized model contributions using a logarithmic function of the amount of local training data. However, evaluating the contribution solely by the training data size is one-sided, as there may be a large amount of redundant data or the validity of the local model cannot be guaranteed. It is more practical to comprehensively evaluate the amount of training data, model accuracy, and the potential value of local data; thus, we denote  $I_{mn}$  as

$$I_{mn} = \frac{\omega_{mn} \epsilon \ln(1 + \eta \lfloor \frac{T}{\tau_n} \rfloor |X_{mn}|)}{\theta_m}, \quad (3)$$

where  $\lfloor \frac{T}{\tau_n} \rfloor$  is the number of iterations that the task of MSP  $n$  can be performed within  $T$ .  $\epsilon$  and  $\eta$  are system parameters from experiments as obtained in [33].  $\theta_m \in (0, 1)$  characterizes the accuracy of local training as decided by the MU. Here,  $\theta \rightarrow 0$  means that a high-accuracy local model is available, while  $\theta \rightarrow 1$  implies that the accuracy of the local model is low. The local data potential value  $\omega_{mn}$  describes the difference between model predictions and true labels, defined as

$$\omega_{mn} = \frac{1}{|X_{mn}|} \sum_{i=1}^{|X_{mn}|} (\hat{z}_i - z_i)^2, \quad (4)$$

where  $\hat{z}_i$  denotes the inferred label from the current global model and  $z_i$  is its ground-truth label. If the value of  $\omega_{mn}$  is large, it indicates that the current global model is not working well on local data or there are unseen samples. Therefore, updating the model by training on these data may lead to better performance of the global model. If the value of  $\omega_{mn}$  is small, the current data has little potential to improve the global model further since the knowledge of the data has already been learned or the data size is small. Note that to shorten the decision time, we use the

<sup>6</sup>For example, an MSP may receive fresher local updates but with lower contributions, or less fresh local updates but with higher contributions.

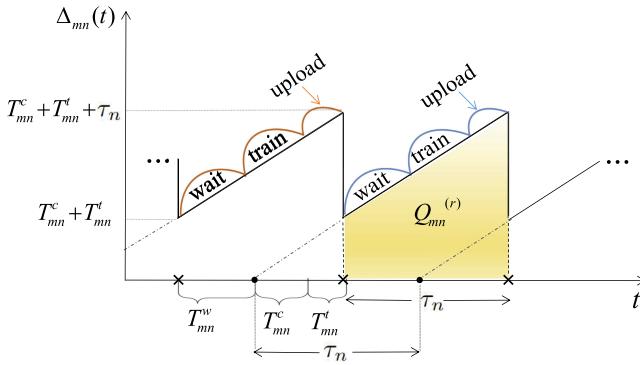


Fig. 5. The AoI evolution of the local updates sent from MU  $m$  to MSP  $n$ .

potential value of the initial local data to be a proxy for the average potential value of all data within  $T$ . This is a reasonable approach when  $T$  is not too long.

2) *Age of Information*  $\Delta$ : The whole FL process involves model training, uploading, and waiting, which together constitute a complete cycle of model updating. From the perspective of MSP  $n$ , the age evolution of the model contributed by MU  $m$  is shown in Fig. 5. Here, the filled circles denote the time instances when MU  $m$  starts training the local model. The intersections show the time instances when MSP  $n$  receives the corresponding local updates, at which point the AoI drops to the lowest.

We denote that at time  $t$ , the generation time of the last update received by MSP  $n$  from MU  $m$  is  $s(t)$ . Then, the instantaneous AoI of MU  $m$ 's local model measured at MSP  $n$ ,  $\Delta_{mn}(t)$ , is

$$\Delta_{mn}(t) = t - s(t), \quad (5)$$

which shows the elapsed time since the generation of the latest local updates. A smaller instantaneous AoI means that the latest local updates received are fresher at that time.

To obtain a comprehensive understanding of the overall freshness of model updates over a certain period, we use *average AoI* to evaluate the freshness of the models. With the aid of Fig. 5, the time-averaged AoI of the local model from MU  $m$  to MSP  $n$  can be expressed as

$$\bar{\Delta}_{mn} = \lim_{Z \rightarrow \infty} \frac{1}{Z} \int_0^Z \Delta_{mn}(t) dt = \lim_{Z \rightarrow \infty} \sum_{r=1}^Z \frac{Q_{mn}^{(r)}}{Z} = \frac{1}{2}\tau_n + T_{mn}^c + T_{mn}^t, \quad (6)$$

where  $Q_{mn}^{(r)}$  is the  $r$ -th trapezium under the curve. A lower average AoI indicates that the local updates are generally fresher over a long period.

Since MU  $m$  continuously collects data during the FL process, the amount of training data used in each round is  $|X_{mn}| = x_m \tau_n$ , where  $x_m$  denotes the data (number of floats) collected per unit time. The cumulative time  $T_{mn}^c$  for local training is determined by the amount of training data, computational resource  $f_{mn}$  allocated to MSP  $n$ , and its local accuracy threshold  $\theta_m$ ,

which is equivalent to

$$T_{mn}^c = \log(1/\theta_m) \frac{x_m \tau_n}{f_{mn}}. \quad (7)$$

A smaller value of  $\theta$  indicates higher accuracy, but leads to a higher MU cost, i.e., the number of local iterations, which is upper bounded by  $\log(1/\theta_m)$  [34].

Furthermore, the upload time  $T_{mn}^t$  depends on the size of local updates  $b_{mn}$  (in bits) and the resource  $B_{mn}$  (in Hz) allocated for communication between MU  $m$  and MSP  $n$  (i.e., uplink bandwidth), defined as

$$T_{mn}^t = \frac{b_{mn}}{B_{mn} \log_2(1 + \varsigma_{mn})}, \quad (8)$$

where  $\varsigma_{mn} = \frac{p_m^t g_{mn}}{\sigma^2}$  denotes the signal-to-noise ratio (SNR) [35] for the communication between MU  $m$  and MSP  $n$ .  $g_{mn}$  and  $p_m^t$  are the corresponding channel gain and the transmission power of MU  $m$ , respectively.  $\sigma^2$  is the power of additive white Gaussian noise (AWGN).

Finally, by substituting (7) and (8) into (6),  $\bar{\Delta}_{mn}$  can be expressed as

$$\bar{\Delta}_{mn} = \frac{1}{2}\tau_n + \frac{x_m \tau_n \log(1/\theta_m)}{f_{mn}} + \frac{b_{mn}}{B_{mn} \log_2(1 + \varsigma_{mn})}. \quad (9)$$

#### IV. GAME FORMULATION

We first define the utility functions of MU  $m$  and MSP  $n$  in Section IV-A and Section IV-B, respectively. Then, we formulate the interactions among MUs and MSPs as an equilibrium problem with equilibrium constraints (EPEC) in Section IV-C, where the equilibrium criterion exists at both the level of MUs and MSPs due to the conflicting interests between them.

##### A. Utility of MU

We define the utility of MU  $m$  composed of the rewards from MSPs and the cost incurred for contribution, which is also affected by resource constraints and immersion requirements for consuming vehicular metaverse services. The cost incurred due to local training and model uploading can be expressed as

$$C_{mn} = c_m^f \log(1/\theta_m) f_{mn} + c_m^B B_{mn}, \quad (10)$$

where  $c_m^f$  and  $c_m^B$  denote the cost factors of computational and communication resources, respectively. Let  $\mathbf{f}_m \triangleq (f_{m1}, f_{m2}, \dots, f_{mN})$  and  $\mathbf{B}_m \triangleq (B_{m1}, B_{m2}, \dots, B_{mN})$  be the computational and communication allocation profiles of MU  $m$ , respectively. Then, the utility maximization problem for MU  $m$  within  $T$  can be formulated as

*Problem 1:*

$$\begin{aligned} \max \Phi_m(\mathbf{f}_m, \mathbf{B}_m) &= \sum_{n=1}^N (p_{mn} V_{mn} - C_{mn}) \\ \text{s.t. } C1: & \sum_{n=1}^N f_{mn} < f_m^{max}, \quad \sum_{n=1}^N B_{mn} < B_m^{max}, \\ C2: & \log(1/\theta_m) \frac{x_m \tau_n}{f_{mn}} + \frac{b_{mn}}{B_{mn} \log_2(1 + \varsigma_{mn})} \leq \frac{1}{2} \tau_n, \\ C3: & \frac{S_m}{f_m^{max} - \sum_{n=1}^N f_{mn}} < T^{req}, \end{aligned} \quad (11)$$

where  $p_{mn}$  is the reward offered by MSP  $n$  to MU  $m$  for their contribution.  $S_m$  indicates the minimum computational resource required to enjoy other basic services. Constraints C1 and C3 indicate that the MU's total computational and communication resources are limited and cannot be fully used to contribute local models, given the MU's demand for basic services. Constraint C2 ensures that  $V_{mn}$  is non-negative, guaranteeing that the value derived from the task meets the MSP's minimum requirements. Specifically, the total time for local training and uploading in each round cannot exceed the time constraint  $\frac{1}{2} \tau_n$ . The goal of each MU is to choose the optimal allocation of computational  $\mathbf{f}_m$  and communication  $\mathbf{B}_m$  resources to maximize its utility.

### B. Utility of MSP

Each MSP has a gain function  $\psi$  associated with the IoM to quantify the cumulative benefit of all local models contributed by MUs. In this paper, we adopt  $\psi_n = \mu_n \ln(1 + \sum_{m=1}^M V_{mn})$  as a monotonically increasing, differentiable, strictly concave IoM function, which is a simplification of the widely adopted function [36]. Let  $\mathbf{p}_n \triangleq (p_{1n}, \dots, p_{Mn})$  and  $\mathbf{p}_{-n}$  represent the reward profile of MSP  $n$  and all other MSPs except MSP  $n$ , respectively. Then, the optimization problem for the MSP is defined as

*Problem 2:*

$$\begin{aligned} \max \Psi_n(\mathbf{p}_n, \mathbf{p}_{-n}) &= \psi_n - \sum_{m=1}^M p_{mn} V_{mn}, \\ \text{s.t. } p_{mn} &> 0, n \in \mathcal{N}, m \in \mathcal{M}, \end{aligned} \quad (12)$$

where  $\mu_n$  refers to the profit conversion coefficient [37] from IoM, which is adjusted according to the AR services offered by the different MSPs. Note that the other MSPs' decisions  $\mathbf{p}_{-n}$  are captured by the MUs' responses. Moreover, when MSP  $n$  determines its reward  $\mathbf{p}_n$  for different MUs, the MSP needs to consider the rewards offered by other MSPs (i.e.,  $\mathbf{p}_{-n}$ ) as well as the strategies of all MUs (i.e.,  $(\mathbf{f}_m, \mathbf{B}_m)$ ,  $m \in \mathcal{M}$ ). This thereby leads to reward competition among MSPs. Therefore, the optimization among MSPs can be considered a non-cooperative game, termed a multi-MSP rewarding game  $\Omega$ , as follows.

*Definition 1:* A multi-MSP rewarding game  $\Omega$  is a tuple  $\Omega = \{\mathcal{N}, \mathbf{p}, \Psi\}$  defined by

- *Players:* The set of MSPs;
- *Strategies:* The reward decisions  $\mathbf{p}_n$  of any MSP  $n$ ;

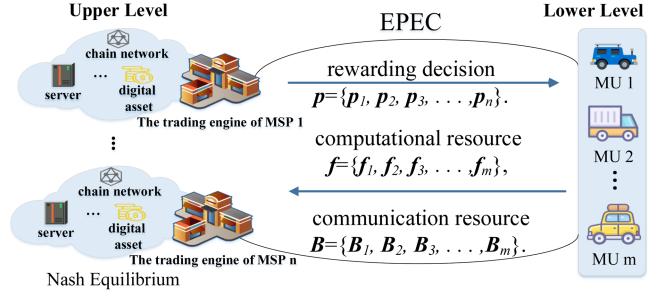


Fig. 6. The hierarchical structure of equilibrium problem with equilibrium constraints (EPEC).

- *Utilities:* The vector  $\Psi = \{\Psi_1, \Psi_2, \dots, \Psi_n\}$  contains the utility functions of all the MSPs defined in (12).

### C. Multi-MSP Multi-MU Game as EPEC

As shown in Fig. 6, MUs and MSPs negotiate allocation strategies of computational and communication resources and reward decisions to maximize their benefits. In the upper level, MSPs determine the rewards they are willing to offer by considering their cost, the responses of MUs, and the decisions of other MSPs. In the lower level, each MU  $m$  gives the optimal allocation response for computational and communication resources by considering its resource constraints, cost, and rewards from different MSPs.

The objective of EPEC is to find the equilibria at two levels, i.e., the point at which the MSPs' (leaders') utilities are maximized given that the MUs (followers) will choose their best responses. For the proposed EPEC, the equilibria at two levels are defined as follows.

*Definition 2:* Let  $(f_{mn}^*, B_{mn}^*)$  and  $p_{mn}^*$  denote the optimal computational and communication resource allocation of MU  $m \in \mathcal{M}$  and the optimal reward decision of MSP  $n \in \mathcal{N}$ , respectively. Then, the points  $(f_{mn}^*, B_{mn}^*)$  and  $p_{mn}^*$  are the equilibria at two levels if the following conditions hold:

$$\begin{aligned} \Phi_m((\mathbf{f}_m^*, \mathbf{B}_m^*), \mathbf{p}_n^*) &\geq \Phi_m((\mathbf{f}_m, \mathbf{B}_m), \mathbf{p}_n^*), \forall m \in \mathcal{M}, \\ \Psi_n(f_m^*(\mathbf{p}_n^*, \mathbf{p}_{-n}^*), B_m^*(\mathbf{p}_n^*, \mathbf{p}_{-n}^*), \mathbf{p}_n^*, \mathbf{p}_{-n}^*) \\ &\geq \Psi_n(f_m^*(\mathbf{p}_n^*, \mathbf{p}_{-n}^*), B_m^*(\mathbf{p}_n^*, \mathbf{p}_{-n}^*), \mathbf{p}_n, \mathbf{p}_{-n}^*), \forall n \in \mathcal{N}, \end{aligned} \quad (13)$$

where  $\mathbf{p}_{-n}^*$  denotes the optimal reward vector for all MSPs except  $n$ .

In summary, the MSPs' optimization problems are formulated as the following EPEC problems:

$$\begin{aligned} \max_{\mathbf{p}_n} \Psi_n &= \psi_n - \sum_{m=1}^M \left[ p_{mn} I_{mn} \left( \frac{1}{2} \tau_n - \frac{x_m \tau_n \log(1/\theta_m)}{f_{mn}^*} \right. \right. \\ &\quad \left. \left. - \frac{b_{mn}}{B_{mn}^* \log_2(1 + \varsigma_{mn})} \right) \right], \\ \text{s.t. } & \begin{cases} p_{mn} > 0, n \in \mathcal{N}, m \in \mathcal{M}, \\ (\mathbf{f}_m^*, \mathbf{B}_m^*) = \arg \max \Phi_m(\mathbf{f}_m, \mathbf{B}_m), \end{cases} \\ & \quad \text{s.t. } C1, C2, C3. \end{aligned} \quad (14)$$

To investigate the above EPEC, we address the lower level (Problem 1) and the upper level (Problem 2) by using the backward induction methods in the following section.

## V. EPEC ANALYSIS AND SOLUTIONS

In this section, we utilize backward induction to analyze the EPEC formulated in (14). Specifically, we prove the existence and uniqueness of the equilibrium solutions at two hierarchical levels: (1) the optimal allocation of computational and communication resources for MUs (lower level), and (2) the optimal reward decisions for MSPs (upper level).

### A. Lower Level: Optimal Resource Allocation for MUs

In the lower level of EPEC, for any reward decisions  $\mathbf{p}$  given by MSPs, MU  $m$  aims to solve **Problem 1** in (11) to determine its optimal computational and communication resource allocations, i.e.,  $\mathbf{f}_m^*$  and  $\mathbf{B}_m^*$ , to maximize its utility. Below, we analyze and derive the unique optimal allocation.

*Theorem 1:* Problem 1 is strictly concave and has a unique globally optimal solution in the lower level. That is, for each MU  $m$ , there exists a unique resource allocation tuple  $(\mathbf{f}_m^*, \mathbf{B}_m^*)$  that maximizes its utility.

*Proof:* We first examine the Hessian matrix of MU  $m$ 's utility  $\Phi_m(\mathbf{f}_m, \mathbf{B}_m)$  with respect to  $(f_{mn}, B_{mn})$ . Let this Hessian matrix be denoted by  $\mathbf{H}_m$ , which can be block-diagonalized as

$$\mathbf{H}_m = \begin{bmatrix} \mathbf{H}_m^f & 0 \\ 0 & \mathbf{H}_m^B \end{bmatrix}. \quad (15)$$

The block matrix  $\mathbf{H}_m^f$  can be computed as the second-order partial derivative of  $\Phi_m(\mathbf{f}_m, \mathbf{B}_m)$  with respect to  $f_{mn}$ , i.e.,

$$\begin{aligned} \mathbf{H}_m^f &= \left[ \frac{\partial^2 \Phi_m(\mathbf{f}_m, \mathbf{B}_m)}{\partial f_{mn} \partial f_{mn'}} \right]_{n, n' \in \mathcal{N}} \\ &= -\text{diag}\left(h_{m1}^f, h_{m2}^f, \dots, h_{mN}^f\right) < \mathbf{0}, \end{aligned} \quad (16)$$

where  $h_{mn}^f = p_{mn} I_{mn} \frac{2x_m \tau_n}{f_{mn}^3} \log(1/\theta_m)$ . Similarly, the block matrix  $\mathbf{H}_m^B$  can be calculated by

$$\begin{aligned} \mathbf{H}_m^B &= \left[ \frac{\partial^2 \Phi_m(\mathbf{f}_m, \mathbf{B}_m)}{\partial B_{mn} \partial B_{mn'}} \right]_{n, n' \in \mathcal{N}} \\ &= -\text{diag}\left(h_{m1}^B, h_{m2}^B, \dots, h_{mN}^B\right) < \mathbf{0}, \end{aligned} \quad (17)$$

where  $h_{mn}^B = \frac{2b_{mn} p_{mn} I_{mn}}{B_{mn}^3 \log_2(1+\varsigma_{mn})}$ .

Since  $\mathbf{H}_m^f$  and  $\mathbf{H}_m^B$  are diagonal matrices with strictly negative diagonal elements, they are negative definite. Consequently, the block-diagonal Hessian  $\mathbf{H}_m$  is negative definite, implying  $\Phi_m(\mathbf{f}_m, \mathbf{B}_m)$  is strictly concave and continuous. Moreover, the objective function  $\Phi_m$  tends to  $-\infty$  as any variable approaches the boundary of the feasible set (e.g.,  $f_{mn} \rightarrow 0^+$  and  $B_{mn} \rightarrow 0^+$ , or  $f_{mn} \rightarrow \infty$  and  $B_{mn} \rightarrow \infty$ ). Hence, the maximum is necessarily attained in the interior. Combined with strict concavity over the convex feasible region defined by constraints C1-C3, this ensures that Problem 1 admits a unique global optimum  $(\mathbf{f}_m^*, \mathbf{B}_m^*)$  (see Section IV in [38]).  $\square$

Consequently, MU  $m$  has a unique best-response resource allocation for each MSP by solving Problem 1. From the Karush–Kuhn–Tucker (KKT) conditions, we derive the following proposition.

*Proposition 1:* Given the digital currency price offered by the MSPs and the Lagrange multipliers induced by the constraints, the optimal computational resource  $f_{mn}$  allocated for MSP  $n$  by MU  $m$  satisfies:

$$f_{mn}^* = \begin{cases} \sqrt{\frac{p_{mn} I_{mn} x_m \tau_n}{c_m^f}}, & p_{mn} \geq \frac{F_{mn}^2}{I_{mn}}, \\ F_{mn} \sqrt{\frac{x_m \tau_n}{c_m^f}}, & 0 < p_{mn} < \frac{F_{mn}^2}{I_{mn}}, \\ 0, & \text{otherwise,} \end{cases} \quad (18)$$

and the optimal communication bandwidth  $B_{mn}$  is:

$$B_{mn}^* = \begin{cases} \sqrt{\frac{p_{mn} I_{mn} b_{mn}}{c_m^B \log_2(1+\varsigma_{mn})}}, & p_{mn} \geq \frac{F_{mn}^2}{I_{mn}}, \\ F_{mn} \sqrt{\frac{b_{mn}}{c_m^B \log_2(1+\varsigma_{mn})}}, & 0 < p_{mn} < \frac{F_{mn}^2}{I_{mn}}, \\ 0, & \text{otherwise,} \end{cases} \quad (19)$$

where  $F_{mn} = 2 \log(1/\theta_m) \sqrt{\frac{x_m c_m^f}{\tau_n}} + \frac{2\sqrt{b_{mn} c_m^B}}{\tau_n \sqrt{\log_2(1+\varsigma_{mn})}}$ . For the detailed proof of Proposition 1, please refer to **Appendix A**, available in the supplemental materials online via IEEE Xplore and in [39].

According to Proposition 1, the optimal strategy for each MU  $m$  is influenced by five main factors:  $p_{mn}$ ,  $I_{mn}$ ,  $c_m^f$ ,  $c_m^B$  and  $\tau_n$ . Specifically, MU  $m$  focuses on the rewards paid by MSPs, and the MU tends to put more computational and communication resources into the contribution when  $p$  increases. Moreover, MUs are more willing to invest more resources when the potential contribution prediction  $I_{mn}$  is higher, resulting in greater benefits. In addition, the virtual deadline  $\tau_n$  set by the MSP  $n$  determines the minimum criteria for resource allocation.

### B. Upper Level: Optimal Reward Equilibrium Among MSPs

In the upper level of the EPEC, each MSP  $n$  competes with other MSPs and determines its reward vector  $\mathbf{p}_n$ . Specifically, given the responses  $(\mathbf{f}, \mathbf{B})$  from all MUs, and other MSPs' decisions  $\mathbf{p}_{-n}$ , the optimal reward decision  $\mathbf{p}_n$  of MSP  $n$  can be obtained by solving Problem 2, defined as follows.

*Proposition 2:* Given other MSPs' reward vectors  $\mathbf{p}_{-n}$ , the optimal strategy of MSP  $n$  is

$$\mathbf{p}_n^* = \arg \max_{p_{mn} > 0} \Psi_n(\mathbf{f}_n^*(\mathbf{p}_n, \mathbf{p}_{-n}), \mathbf{B}_n^*(\mathbf{p}_n, \mathbf{p}_{-n}), \mathbf{p}_{-n}). \quad (20)$$

Then, we analyze the existence and uniqueness of the optimal reward decision equilibrium in Theorem 2.

*Theorem 2:* There exists a unique Nash equilibrium in the multi-MSP rewarding game  $\Omega$ , ensuring a unique optimal reward decision profile  $\{\mathbf{p}_n^*\}_{n \in \mathcal{N}}$ .

*Proof:* We define the Hessian matrix of  $\Psi_n$  with respect to its reward vector  $\mathbf{p}_n$  as  $(\Lambda_n + \mathbf{H}_n)$ . The matrix  $\Lambda_n = \text{diag}(\frac{\partial^2 \Psi_n}{\partial p_{1n}^2}, \dots, \frac{\partial^2 \Psi_n}{\partial p_{Mn}^2})$  and the second-order partial derivative

matrix  $\mathbf{H}_n$  is expressed by

$$\mathbf{H}_n = \begin{bmatrix} 0 & \frac{\partial^2 \Psi_n}{\partial p_{1n} \partial p_{2n}} & \dots & \frac{\partial^2 \Psi_n}{\partial p_{1n} \partial p_{Mn}} \\ \frac{\partial^2 \Psi_n}{\partial p_{2n} \partial p_{1n}} & 0 & \dots & \frac{\partial^2 \Psi_n}{\partial p_{2n} \partial p_{Mn}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \Psi_n}{\partial p_{Mn} \partial p_{1n}} & \frac{\partial^2 \Psi_n}{\partial p_{Mn} \partial p_{2n}} & \dots & 0 \end{bmatrix}, \quad (21)$$

where

$$\frac{\partial^2 \Psi_n}{\partial p_{mn}^2} = \mu_n \frac{V''_{mn}(1 + \sum_{m=1}^M V_{mn}) - (V'_{mn})^2}{(1 + \sum_{m=1}^M V_{mn})^2} - 2V'_{mn} - p_{mn}V''_{mn} < 0, \forall m \in \mathcal{M}, \quad (22)$$

$$\frac{\partial^2 \Psi_n}{\partial p_{mn} \partial p_{m'n}} = -\mu_n \frac{V'_{mn} V'_{m'n}}{(1 + \sum_{m=1}^M V_{mn})^2} < 0, \forall m \in \mathcal{M}, m \neq m'. \quad (23)$$

We randomly choose a vector  $\mathbf{h} \in \mathbb{R}^{M \times 1}$  with elements not all 0. Then, we have  $\mathbf{h}^T(\Lambda_n + \mathbf{H}_n)\mathbf{h} = \sum_{m=1}^M (h^m)^2 \left( \frac{\mu_n V''_{mn}}{1 + \sum_{m=1}^M V_{mn}} - 2V'_{mn} - p_{mn}V''_{mn} \right) - \frac{\mu_n}{(1 + \sum_{m=1}^M V_{mn})^2} (\sum_{m=1}^M h^m V'_{mn})^2 < 0$ , indicating that the utility function  $\Psi_n$  is strictly concave. According to [40], there exists a unique Nash equilibrium in the multi-MSP rewarding game  $\Omega$ . The proof of  $\frac{\partial^2 \Psi_n}{\partial p_{mn}^2} < 0$ ,  $\frac{\partial^2 \Psi_n}{\partial p_{mn} \partial p_{m'n}} < 0$ , and  $\mathbf{h}^T(\Lambda_n + \mathbf{H}_n)\mathbf{h} < 0$  can be found in **Appendix B**, available in the supplemental materials online via IEEE Xplore and in [39].  $\square$

Utilizing Theorems 1 and 2, we prove the existence and uniqueness of the hierarchical equilibrium in the proposed EPEC model. Concretely, each MU has a unique best-response resource allocation strategy in reaction to any given MSP rewards. At the same time, each MSP determines unique optimal pricing to maximize its own utility, given the rewards of other MSPs.

Although the backward induction approach theoretically ensures an equilibrium solution to (14), solving it in real-world vehicular metaverse environments can be challenging due to dynamic network conditions and limited system knowledge (e.g., private cost parameters). In the following section, we address these practical issues by introducing a fully distributed reinforcement learning approach, enabling MSPs to adaptively and privately optimize their reward decisions in real time.

## VI. DYNAMIC REWARD STRATEGY FOR MSPS

According to the EPEC proven above, it is possible to obtain an optimal solution. However, the optimal approach faces the following practical challenges: (i) In a time-varying network, the optimal approach is time-consuming due to the complexity of (12), which is a non-linear problem with a complicated structure. Additionally, the reward decisions of MSPs are tightly coupled. (ii) MSPs must have prior knowledge of private information about MUs (e.g., cost factors) to determine their rewards, which raises privacy concerns for MUs. To address these challenges, we further formulate the multi-MSP rewarding game as a multi-agent Markov decision process (MAMDP) [36], adapting to dynamic channel conditions. In this framework, each MSP is

---

### Algorithm 1: Entire Process of the Framework.

---

```

1 Input: Action threshold  $[p_{min}, p_{max}]$  and number of episodes  $\mathcal{T}$ ;
2 Initialization: local observation  $s_n$ , actor network  $\alpha_n$ , critic network  $\beta_n$ , and episode buffer  $\mathcal{D}_n$  for each MSP agent  $n \in \mathcal{N}$ ;
3 # Trading process using MDDR:
4 for  $episode = 1, 2, \dots, \mathcal{T}$  do
5   Concurrently for each MSP agent  $n \in \mathcal{N}$ :
6     for  $epoch k = 1, \dots, |\mathcal{D}_n|$  do
7       Observe space  $s_n(k)$ ;
8       Choose price  $p_n(k) \in [p_{min}, p_{max}]$  by sampling from its current policy  $\pi_{\alpha_n}(s_n(k))$ ;
9       Broadcast reward decisions to MUs;
10      Interact with the environment and receive responses from the MUs;
11      Calculate utility according to (12);
12      Store transition  $e_n(k) = [s_n(k), p_n(k), \mathcal{U}_n(k), s_n(k+1)]$ ;
13   end
14   Update actor network  $\alpha_n$  and critic network  $\beta_n$  for each MSP agent  $n$ ;
15   Clear the episode buffer  $\mathcal{D}_n$  for each MSP agent  $n$ ;
16 end
17 Return: Reward decisions  $\mathbf{p}$  of MSPs and resource allocation  $[f, \mathbf{B}]$  of MUs;
18 # FL training guided by trading results within time T:
19 foreach MSP agent  $n \in \mathcal{N}$  do
20   foreach MU  $m \in \mathcal{M}$  do
21     Based on the trading results, utilize  $f_{mn}$  and  $B_{mn}$  to perform the local FL task for MSP  $n$ , and then upload the local updates to MSP  $n$ .
22   end
23   Aggregate received local updates and issue an updated global model;
24   Supply the updated model to the AR engine;
25 end

```

---

modeled as an individual agent that makes intelligent reward decisions.

#### A. MAMDP for Multi-MSP Rewarding Game $\Omega$

We train the reward model based on the state-of-the-art policy gradient method proximal policy optimization (PPO) for the reasons described in [41]. The  $MAMDP = \langle \mathcal{S}_n, \mathcal{A}_n, \mathcal{P}_n, \mathcal{U}_n \rangle$  for multi-MSP rewarding game is composed of state space  $\mathcal{S}_n \triangleq \{s_n\}$ , action space  $\mathcal{A}_n \triangleq \{p_n\}$ , state transition probability  $\mathcal{P}_n \triangleq \{P_n\}$ , and utility  $\mathcal{U}_n \triangleq \{\Psi_n\}$ . For the MSP agent, the local observation contains the channel information and the responses of MUs, all of which are captured in the IoM of (2). Thus, we set the local observation of MSP  $n$  at the  $k$ -th stage game defined as  $s_n(k) = [V_{1n}(k), V_{2n}(k), \dots, V_{mn}(k)]$ , and the state of the environment is  $\mathcal{S}(k) = [s_1(k), \dots, s_n(k)]$ . At the

$k$ -th stage game, MSP  $n$  observes a state  $s_n(k)$  and determines an action  $p_n(k)$  within  $[p_{\min}, p_{\max}]$ . When an action  $p_n(k)$  is applied to state  $s_n(k)$ , the agent  $n$  receives a utility  $\mathcal{U}_n(k)$  from the environment. Taking into account the competition among agents, we define the reward  $\mathcal{U}_n$  for each MSP by (12). The state transition probability  $P(s_n(k+1) | s_n(k), p_n(k))$  leads to the new state  $s_n(k+1)$  after executing an action  $p_n(k)$  at the state  $s_n(k)$ .

### B. The Multi-Agent DRL-Based Dynamic Reward (MDDR)

In our work, each MSP agent operates the DRL-based dynamic reward algorithm in a fully distributed manner. Algorithm 1 shows the pseudocode for model trading using MDDR under a dynamic environment (Lines 4-17) and the guidance for the FL process (Lines 19-25).

**Trading process:** At the start of the game, each MSP  $n$  initializes its actor  $\alpha_n$  and critic network  $\beta_n$ , and episode buffer  $\mathcal{D}_n$  for each MSP agent  $n$  (Line 2). Each agent  $n$  feeds the observed information into the policy network  $\alpha_n$  to derive its reward policy (Lines 7-8). Subsequently, each MSP agent  $n$  broadcasts its reward decisions, interacts with the environment, receives feedback from the MUs, calculates its utility, and stores the experience in the episode buffer  $\mathcal{D}_n$  (Lines 9-12). The experience is collected through this interactive process (Lines 7-12) until the buffer  $\mathcal{D}_n$  is full. Once the buffer is full, the critic network  $\beta_n$  evaluates the actor network  $\alpha_n$  based on the collected experience and updates it accordingly (Line 14). Finally, the episode buffer is cleared in preparation for the next episode (Line 15). The game terminates once the maximum number of training episodes is reached, after which the trading results are finalized.

**FL training guided by trading results:** After obtaining the trading results (MSPs' reward decisions  $p$  and resource allocation  $(f, B)$ ), MUs and MSPs enter the FL phase. MUs follow the resource allocation scheme for local training and model uploading (Line 21). Then, the MSP is responsible for aggregating and updating the global model (Line 23). Finally, the updated model is provided for use by the AR engine (Line 24). This process iterates multiple times until the total FL time  $T$  is reached. During real-time interactions, the MSP's AR service immersion and quality can be enhanced by utilizing the updated model.

## VII. PERFORMANCE EVALUATION

First, we verify the near-optimal performance of MDDR. Then, to evaluate the effectiveness of the proposed framework, we compare its performance with benchmark schemes in object detection and classification for AR services through FL.

### A. Simulation Settings

We conduct experiments on the MNIST dataset with ResNet-18 and the GTSRB dataset with Faster R-CNN. The GTSRB was recorded during daytime driving on different types of roads in Germany for vehicle AR scenario testing [42]. We set the number of MUs as  $M = 5$  and the number of MSPs as  $N = 3$ .

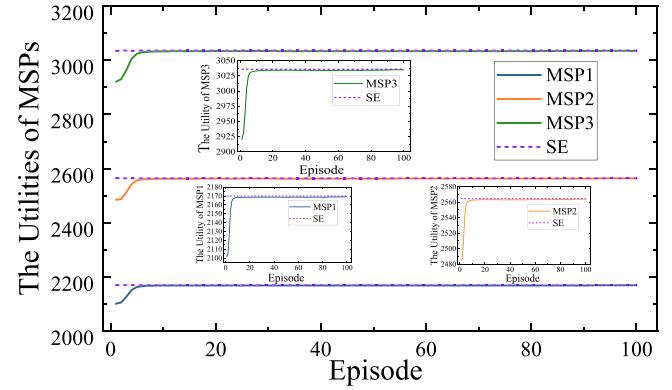


Fig. 7. The utilities of MSPs.

The configurations of the maximum computational and communication resource for each MU are randomly generated from the range [3,5] GHz and [1, 4] MHz [43]. For local training of MUs, we adopt the SGD optimizer with momentum = 0.9 and learning rate = 0.001. For model aggregation, MSPs adopt the Federated Averaging (FedAvg) algorithm, applying equal weights to all participants' updates. The trading guidance time  $T$  for MNIST and GTSRB are 30 s and 1,200s, respectively. All baseline methods are configured with the same settings for fair comparison.

### B. Benchmark Schemes

To our knowledge, there are no comparable solutions to the model trading problem in the vehicular metaverse. Therefore, under the uniform rewards of MSPs, we develop and extend four benchmark schemes for multi-MSP and multi-MU scenarios while meeting FL deadlines and resource constraints of devices.

- **$x_{based}$ :** Inspired by the incentive mechanism of the Stackelberg game for federated learning [32], the MU allocates its computational and communication resources based on the amount of data involved in local training.
- **$w_{based}$ :** The MU adjusts the resource allocation according to the inference loss [44], i.e., the potential value of local data in our work.
- **$w_x_{based}$ :** Combining the  $x_{based}$  and  $w_{based}$  designs, the MU allocates resources considering both the data size and the potential value of data.
- **$fixed$ :** Taking into account the differences in performance of the various devices that exist in reality, we randomly allocate a fixed set of computational and communication resources to reflect the heterogeneity of devices.

### C. Performance Evaluation of MDDR

First, we consider a relatively static communication environment (fixed channel conditions) to show the effectiveness of MDDR compared to the optimal solution and the general multi-agent depth approximation policy optimization (MAPPO) approach. We trained our proposed MDDR and MAPPO until convergence, and the results are shown in Fig. 7 and Table III. We find that both MDDR and MAPPO achieve good results,

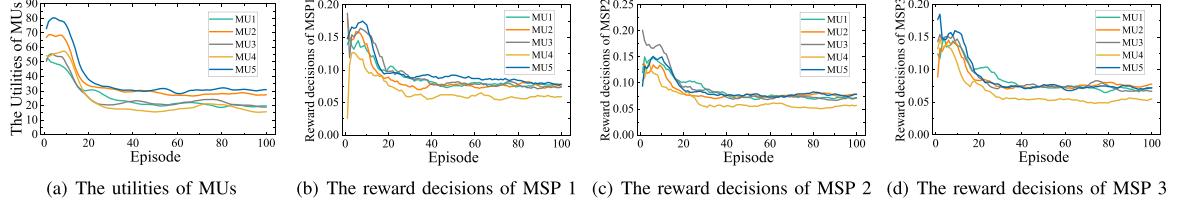


Fig. 8. The convergence processes of MDDR under dynamic networks.

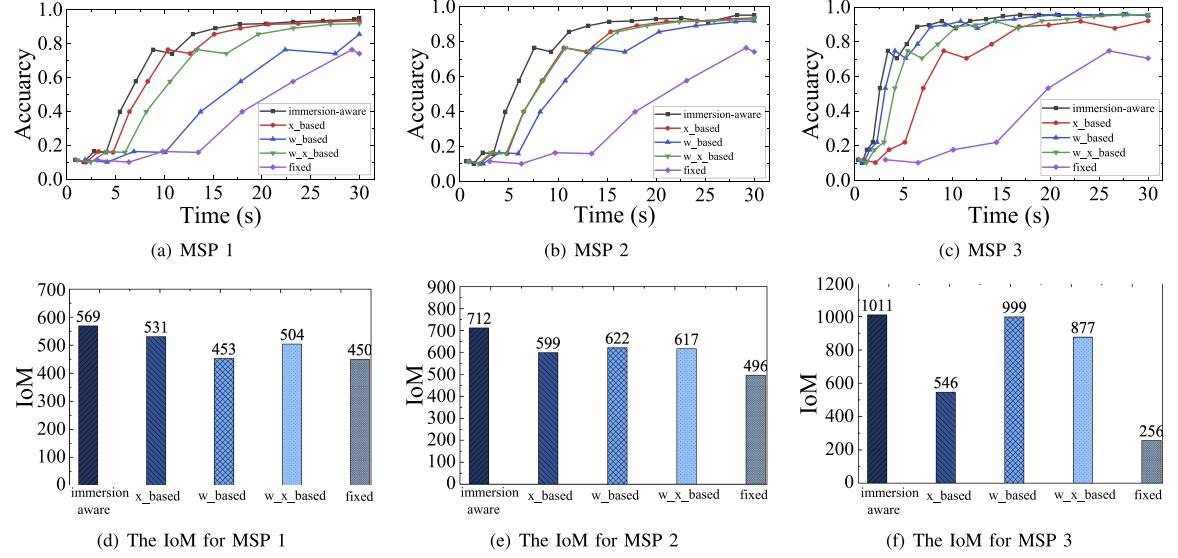


Fig. 9. Evaluation results for MNIST.

TABLE III  
THE UTILITIES OF THREE REWARD APPROACHES

	MSP 1	MSP 2	MSP 3	Total
Optimal	2169.99	2564.56	3035.96	7770.51
MDDR	2169.11	2563.95	3034.94	7768.00
MAPPO	2169.23	2563.96	3035.26	7768.45

but MAPPO slightly outperforms MDDR. This is attributed to MAPPO's centrally controlled yet distributed execution architecture [45] that can utilize more information for decision making. However, this approach is not suitable for non-cooperation due to the need to share parameters among MSPs.

For the fully distributed MDDR, it is more suitable for non-cooperative scenarios. The reason is that in MDDR, each agent has a self-contained actor-critic network, and information is not required to be shared. This suggests that the approach can achieve considerable utility while preserving privacy and saving communication costs significantly. Then, we deploy the MDDR in dynamic communication environments where channel conditions change over time. The trading behaviors of each MU and MSP are illustrated in Fig. 8(a), (b), (c) and (d), reflecting that the MDDR can converge rapidly and show good stability in dynamic network environments.

#### D. Benchmark Comparison

Fig. 9(a), (b), and (c) show the performance improvement in classification accuracy over time, where the *immersion-aware* outperforms the other schemes. This is because *immersion-aware* utilizes IoM to incentivize MUs to allocate computational and communication resources for providing local models to MSPs. The IoM metric not only evaluates the freshness and accuracy of the local model but also considers both the amount and potential value of raw data used for training. As a result, resource-constrained MUs are motivated to contribute more resources, providing valuable local models to MSPs more rapidly.

In contrast, baseline schemes like *x\_based*, *w\_based*, and *w\_x\_based* struggle to offer adequate incentives for MUs to contribute their available resources due to uniform rewards. For example, if an MU has already reached its maximum allocatable resources at a given reward  $p$ , increasing  $p$  further will not yield additional contributions. In this situation, raising the reward only increases the MSP's cost without eliciting additional contributions, so its utility actually declines. Consequently, the uniform reward  $p$  remains unchanged, preventing other MUs with more resources from contributing higher-value local models.

In addition, Fig. 9(a), (b), and (c) reveal that MSP 3 performs better than MSP 1 and MSP 2. For example, at the 10th second, the accuracy of MSP 3 is about 88%, which is 14% and 3% higher than those of MSP 1 and MSP 2, respectively. This is

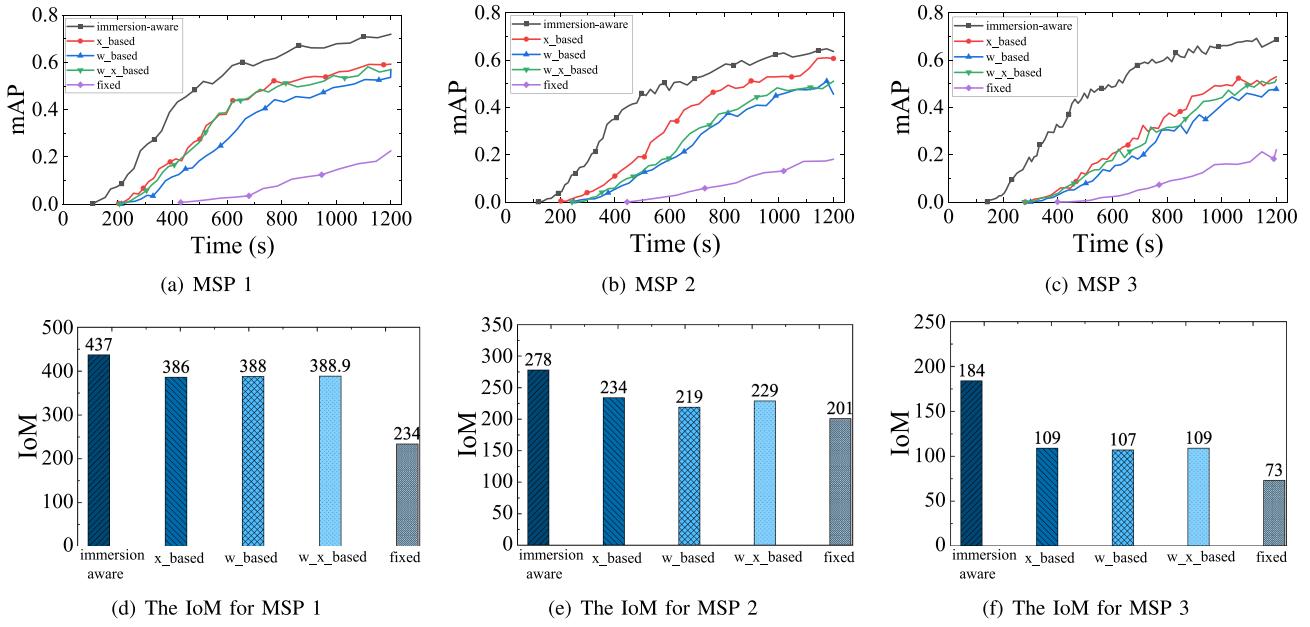


Fig. 10. Evaluation results for GTSRB.

TABLE IV  
TIME COMPARISON OF FIVE SCHEMES FOR ACHIEVING THE SAME PERFORMANCE

Schemes	MNIST			GTSRB		
	Training Time to reach 90% Accuracy (s)			Training Time to reach 50% mAP (s)		
	MSP 1	MSP 2	MSP 3	MSP 1	MSP 2	MSP 3
immersion-aware	15.23	13.09	7.59	513.16	581.09	623.89
x_based	17.79	18.03	19.80	812.53	898.59	1061.50
w_based	30+	24.04	10.28	1064.83	1175.58	1200+
w_x_based	23.20	19.08	12.34	901.61	1190	1194
fixed	30+	30+	30+	1200+	1200+	1200+

Note: 30+ or 1200+ indicates that the scheme required more than 30 seconds or 1200 seconds, respectively, to achieve the target accuracy, exceeding the guideline time for the trading results.

because MSP 3 is associated with a higher potential value  $\omega$ , resulting in more substantial rewards being offered. As a result, the MUs will allocate more computational and communication resources to provide local models for MSP 3 than for MSP 1 and MSP 2. For the three different MSPs, some fluctuations are normal for model training (e.g., at time 10 s). Notably, the *fixed* scheme performs the worst due to the limitations of synchronous FL, which is bottlenecked by the least capable MU.

Meanwhile, it is clear from Fig. 9(d), (e) and (f) that the effectiveness of the five schemes is influenced by IoM. With the higher IoM, the performance is improved faster, resulting in a better immersive experience. Although the IoM of *w\_based* in Fig. 9(e) is higher than *x\_based* and *w\_x\_based*, its performance improvement rate is comparatively slower. This is because the IoM represents the aggregate value across all MUs, while the performance improvement rate is hampered by the worst MU. Furthermore, from Fig. 9(b) and (c), it can be observed that both *w\_x\_based* and *w\_based* perform better when compared to Fig. 9(a). The main reason is the higher potential value  $\omega$  associated with MSP 2 and MSP 3.

For object detection on GTSRB, Fig. 10(a), (b) and (c) show the mean average precision (mAP) and Table IV gives the mAP@50 (The value of mAP when the intersection over

union (IOU) threshold is larger than 0.5) results. Compared to MNIST, the performance enhancement for object detection is slower due to the complexity of the detection task. Specifically, it takes at least 513.16 seconds to achieve a 50% mAP using the *immersion-aware* method. Our research introduces the flexibility to adjust decision guidance time, making it adaptable to various task types. Moreover, Fig. 10(d), (e), (f) and Table IV demonstrate that MSP 1 with greater IoM takes less time to achieve the same mAP, which indicates that IoM works well to capture immersion. For instance, MSP 1 has the highest IoM, and the time that it takes to reach 50% mAP is 513.16 s, which is 67.93 s and 110.73 s less than those of MSP 2 and MSP 3, respectively.

The comparison results validate the advantages of the proposed trading framework, which outperforms other schemes by improving IoM by 38.3% and 37.2% and reducing the training time to reach the target accuracy by 43.5% and 49.8%, on average, for the MNIST and GTSRB datasets, respectively.

## VIII. CONCLUSION

To facilitate the development of vehicular metaverse services, we propose an immersion-aware model trading framework,

incorporating FL to incentivize MUs to contribute local models for AR services such as object detection and classification. Specifically, we construct an EPEC problem with MSPs as leaders and MUs as followers to achieve an equilibrium among their interests. A new metric called “IoM” is designed to comprehensively evaluate the enhancement brought by the local models of MUs for AR services. Furthermore, considering the competitive relationship among MSPs and the dynamic network environment, we develop a fully distributed MDDR approach to obtain the reward decisions of MSPs. Extensive simulations on AR-related vehicle and MNIST datasets demonstrate that the proposed framework enables more efficient and immersive AR services in the vehicular metaverse.

## REFERENCES

- [1] Y. Wang et al., “A survey on metaverse: Fundamentals, security, and privacy,” *IEEE Commun. Surv. Tuts.*, vol. 25, no. 1, pp. 319–352, First Quarter, 2023.
- [2] MarketsandMarkets, “Metaverse market for automotive by products (software, hardware), technology (virtual reality (VR), augmented reality (AR), mixed reality (MR), non-fungible token (NFT), blockchain), function, application & region - global forecast to 2030,” Aug. 2022. [Online]. Available: <https://www.marketsandmarkets.com/Market-Reports/automotive-metaverse-market-10437470.html>
- [3] D. S. Sarwatt, Y. Lin, J. Ding, Y. Sun, and H. Ning, “Metaverse for intelligent transportation systems (ITS): A comprehensive review of technologies, applications, implications, challenges and future directions,” *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 7, pp. 6290–6308, Jul. 2024.
- [4] D. Qiao, L. Qian, S. Guo, J. Zhao, and P. Zhou, “AMFL: Resource-efficient adaptive metaverse-based federated learning for the human-centric augmented reality applications,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 5, pp. 7888–7902, May 2025.
- [5] Nissan Motor Corporation, “Invisible-to-visible (I2V),” Accessed: Apr. 20, 2025. [Online]. Available: <https://www.nissan-global.com/EN/INNOVATION/TECHNOLOGY/ARCHIVE/I2V/>
- [6] I. Gokasar, D. Pamucar, M. Deveci, B. B. Gupta, L. Martinez, and O. Castillo, “Metaverse integration alternatives of connected autonomous vehicles with self-powered sensors using fuzzy decision making model,” *Inf. Sci.*, vol. 642, Sep. 2023, Art. no. 119192.
- [7] F. Chiariotti, “Age of information analysis for a shared edge computing server,” *IEEE Trans. Commun.*, vol. 72, no. 12, pp. 7826–7841, Dec. 2024.
- [8] L. Feng et al., “Resource allocation for metaverse experience optimization: A multi-objective multi-agent evolutionary reinforcement learning approach,” *IEEE Trans. Mobile Comput.*, vol. 24, no. 4, pp. 3473–3488, Apr. 2025.
- [9] D. Chen et al., “Federated learning based mobile edge computing for augmented reality applications,” in *Proc. Int. Conf. Comput. Netw. Commun.*, 2020, pp. 767–773.
- [10] X. Zhou, Y. Li, and J. Zhao, “Resource allocation of federated learning-assisted mobile augmented reality systems in the metaverse via uplink and downlink NOMA,” *IEEE Trans. Wireless Commun.*, early access, Apr 08, 2025, doi: [10.1109/TWC.2025.3554361](https://doi.org/10.1109/TWC.2025.3554361).
- [11] B. Hazarika, K. Singh, O. A. Dobre, C.-P. Li, and T. Q. Duong, “Quantum-enhanced federated learning for metaverse-powered vehicular networks,” *IEEE Trans. Commun.*, vol. 73, no. 6, pp. 4168–4183, Jun. 2025, doi: [10.1109/TCOMM.2024.3502667](https://doi.org/10.1109/TCOMM.2024.3502667).
- [12] J. Kang et al., “Blockchain-empowered federated learning for healthcare metaverses: User-centric incentive mechanism with optimal data freshness,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 1, pp. 348–362, Feb. 2024.
- [13] E. Baccour, A. Erbad, A. Mohamed, M. Hamdi, and M. Guizani, “A blockchain-based reliable federated meta-learning for metaverse: A dual game framework,” *IEEE Internet Things J.*, vol. 11, no. 12, pp. 22697–22715, Apr. 2024.
- [14] X. Li et al., “Meta-computing enhanced federated learning in IIoT: Satisfaction-aware incentive scheme via DRL-based stackelberg game,” 2025, *arXiv:2502.06909*.
- [15] Q. Zhang, Z. Xiong, J. Zhu, S. Gao, and W. Yang, “A privacy-preserving auction mechanism for learning model as an NFT in blockchain-driven metaverse,” *ACM Trans. Multimedia Comput., Commun. Appl.*, vol. 20, no. 7, Mar. 2024, Art. no. 190.
- [16] N. Raveendran, H. Zhang, L. Song, L.-C. Wang, C. S. Hong, and Z. Han, “Pricing and resource allocation optimization for IoT fog computing and NFV: An EPEC and matching based perspective,” *IEEE Trans. Mobile Comput.*, vol. 21, no. 4, pp. 1349–1361, Apr. 2022.
- [17] J. Feng and J. Zhao, “Resource allocation for augmented reality empowered vehicular edge metaverse,” *IEEE Trans. Commun.*, vol. 73, no. 3, pp. 1987–2001, Mar. 2025.
- [18] L. U. Khan, M. Alghfeli, M. Guizani, N. Saeed, and S. Muhamadat, “TCS: A joint task offloading, communication, and sensing framework for vehicular metaverse,” *IEEE Internet Things J.*, vol. 12, no. 16, pp. 32994–33010, Aug. 2025, doi: [10.1109/JIOT.2025.3573837](https://doi.org/10.1109/JIOT.2025.3573837).
- [19] Y. Tong et al., “Multi-attribute auction-based resource allocation for twins migration in vehicular metaverses: A GPT-based DRL approach,” *IEEE Trans. Cogn. Commun. Netw.*, vol. 11, no. 1, pp. 638–654, Feb. 2025.
- [20] J. Kang et al., “Blockchain-based pseudonym management for vehicle twin migrations in vehicular edge metaverse,” *IEEE Internet Things J.*, vol. 11, no. 21, pp. 34254–34269, Nov. 2024.
- [21] I. Lotfi, M. Qaraqe, A. Ghayeb, and D. Niyato, “VMGuard: Reputation-based incentive mechanism for poisoning attack detection in vehicular metaverse,” *IEEE Trans. Veh. Technol.*, vol. 74, no. 7, pp. 10255–10267, Jul. 2025, doi: [10.1109/TVT.2025.3543650](https://doi.org/10.1109/TVT.2025.3543650).
- [22] M. Deveci, D. Pamucar, I. Gokasar, M. Koppen, B. B. Gupta, and T. Daim, “Evaluation of metaverse traffic safety implementations using fuzzy einstein based logarithmic methodology of additive weights and TOPSIS method,” *Technol. Forecasting Social Change*, vol. 194, Sep. 2023, Art. no. 122681.
- [23] M. Deveci, D. Pamucar, I. Gokasar, M. Koppen, and B. B. Gupta, “Personal mobility in metaverse with autonomous vehicles using Q-rung orthopair fuzzy sets based OPA-RAFSI model,” *IEEE Trans. Intell. Transp. Syst.*, vol. 24, no. 12, pp. 15642–15651, Dec. 2023.
- [24] F. Nadimi, P. Abdisarabshali, K. Borazjani, J. Chakareski, and S. Hosseinalipour, “Multi-modal multi-task federated foundation models for next-generation extended reality systems: Towards privacy-preserving distributed intelligence in AR/VR/MR,” Jun. 2025, *arXiv:2506.05683*.
- [25] Z. Zhang, F. Zeng, and F. Tang, “Vehicle-assisted data sensing in vehicle edge metaverse: A game theory approach,” *IEEE Trans. Veh. Technol.*, vol. 74, no. 4, pp. 5664–5675, Apr. 2025.
- [26] J. Xu et al., “Semantic-aware UAV swarm coordination in the metaverse: A reputation-based incentive mechanism,” *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 13821–13833, Dec. 2024.
- [27] X. Lin, J. Wu, J. Li, W. Yang, and M. Guizani, “Stochastic digital-twin service demand with edge response: An incentive-based congestion control approach,” *IEEE Trans. Mobile Comput.*, vol. 22, no. 4, pp. 2402–2416, Apr. 2023.
- [28] N. Zhao et al., “Two-stage Stackelberg game based physical reality sensing incentive mechanism in metaverse,” *Internet Things*, vol. 30, Mar. 2025, Art. no. 101507.
- [29] E. Almosherea, E. Baccour, A. Erbad, and M. Hamdi, “MFSC: A federated learning and semantic communication framework for efficient digital twin creation in metaverse,” in *Proc. IEEE Int. Conf. Intell. Metaverse Technol. Appl.*, 2024, pp. 01–08.
- [30] L. Gao, H. Fu, L. Li, Y. Chen, M. Xu, and C.-Z. Xu, “FedDC: Federated learning with non-IID data via local drift decoupling and correction,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 10112–10121.
- [31] K. Wang, Z. Ding, D. K. C. So, and Z. Ding, “Age-of-information minimization in federated learning based networks with non-IID dataset,” *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8939–8953, Aug. 2024.
- [32] Y. Zhan, P. Li, Z. Qu, D. Zeng, and S. Guo, “A learning-based incentive mechanism for federated learning,” *IEEE Internet Things J.*, vol. 7, no. 7, pp. 6360–6368, Jul. 2020.
- [33] S. Jiang and J. Wu, “A reward response game in the federated learning system,” in *Proc. IEEE 18th Int. Conf. Mobile Ad Hoc Smart Syst.*, 2021, pp. 127–135.
- [34] X. Lin, J. Wu, J. Li, X. Zheng, and G. Li, “Friend-as-learner: Socially-driven trustworthy and efficient wireless federated edge learning,” *IEEE Trans. Mobile Comput.*, vol. 22, no. 1, pp. 269–283, Jan. 2023.
- [35] F. Li, P. Rao, W. Sun, Y. Su, and X. Chen, “A low signal-to-noise ratio infrared small-target detection network,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 8643–8658, 2025.

- [36] Y. Zhan, C. H. Liu, Y. Zhao, J. Zhang, and J. Tang, "Free market of multi-leader multi-follower mobile crowdsensing: An incentive mechanism design by deep reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 19, no. 10, pp. 2316–2329, Oct. 2020.
- [37] W. Y. B. Lim et al., "When information freshness meets service latency in federated learning: A task-aware incentive scheme for smart industries," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 457–466, Jan. 2022.
- [38] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004. [Online]. Available: <https://web.stanford.edu/boyd/cvxbook/>
- [39] H. Wu et al., "Metatrading: An immersion-aware model trading framework for vehicular metaverse services," Apr. 2025, *arXiv:2410.19665*.
- [40] T. Kim, C. K. Kim, S.-S. Lee, and S. Lee, "Incentive-aware partitioning and offloading scheme for inference services in edge computing," *IEEE Trans. Serv. Comput.*, vol. 17, no. 4, pp. 1580–1592, Jul./Aug. 2024.
- [41] J. Ji, K. Zhu, and L. Cai, "Trajectory and communication design for cache-enabled UAVs in cellular networks: A deep reinforcement learning approach," *IEEE Trans. Mobile Comput.*, vol. 22, no. 10, pp. 6190–6204, Oct. 2023.
- [42] D. Toshniwal, S. Loya, A. Khot, and Y. Marda, "Optimized detection and classification on GTRSB: Advancing traffic sign recognition with convolutional neural networks," Mar. 2024, *arXiv:2403.08283*.
- [43] S. Jošilo and G. Dán, "Wireless and computing resource allocation for selfish computation offloading in edge computing," in *Proc. IEEE Conf. Comput. Commun.*, 2019, pp. 2467–2475.
- [44] H. Zeng, T. Zhou, Y. Guo, Z. Cai, and F. Liu, "FedCav: Contribution-aware model aggregation on distributed heterogeneous data in federated learning," in *Proc. 50th Int. Conf. Parallel Process.*, 2021, pp. 1–10.
- [45] Q. Han, X. Wang, and W. Shen, "Communication-dependent computing resource management for concurrent task orchestration in IoT systems," *IEEE Trans. Mobile Comput.*, vol. 23, no. 12, pp. 14297–14312, Dec. 2024.



**Hongjia Wu** received the PhD degree in computer science from the National University of Defense Technology, China, in 2023. She was a visiting student with the Singapore University of Technology and Design. She is currently a postdoctoral fellow with the Education University of Hong Kong in China. Her research interests mainly focus on computational offloading, edge intelligence, network economics and game theory, and Internet of Things.



**Hui Zeng** received the BS degree from Sun Yat-sen University, in 2020, and the MS degree from the National University of Defense Technology, in 2022. He is currently working toward the PhD degree with the National University of Defense Technology. His research interests include data privacy, federated learning, and distributed systems.



**Zehui Xiong** received the PhD degree from Nanyang Technological University and was a visiting scholar with Princeton University and University of Waterloo. He is currently a full professor with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, United Kingdom. Prior to that, he was with Singapore University of Technology and Design, and Alibaba-NTU Singapore Joint Research Institute. Recognized as a Clarivate Highly Cited Researcher, he has published more than 250 peer-reviewed research papers in leading journals, with numerous Best Paper Awards from international flagship conferences. His research interests include 5G/B5G/6G and cellular systems, autonomous vehicles and intelligent transportation, cooperative communication and green communication, land transportation, vehicular communications and connected vehicles, wireless networks.



**Jiawen Kang** received the MS and the PhD degrees from the Guangdong University of Technology, China, in 2015 and 2018. He is currently a full professor with the Guangdong University of Technology. He has been a postdoc with Nanyang Technological University from 2018 to 2021, Singapore. His research interests include mainly focus on blockchain, Metaverse, and AIGC in wireless communications, and networking.

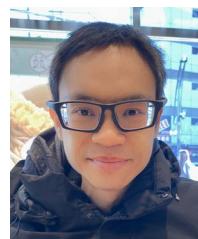


**Zhiping Cai** received the BEng, MSc, and PhD degrees in computer science and technology from the National University of Defense Technology (NUDT), China, in 1996, 2002, and 2005, respectively. He is a full professor with the College of Computer Science and Technology, NUDT. His current research interests include artificial intelligence, network security, and Big Data. He is a distinguished member of the CCF.



**Tse-Tin Chan** (Member, IEEE) received the BEng and PhD degrees in information engineering from the Chinese University of Hong Kong (CUHK), Hong Kong, China, in 2014 and 2020, respectively. From 2020 to 2022, he was an assistant professor with the Department of Computer Science, The Hang Seng University of Hong Kong (HSUHK), Hong Kong. He is currently an assistant professor with the Department of Mathematics and Information Technology, The Education University of Hong Kong (EdUHK), Hong Kong. His research interests include wireless

communications and networking, the Internet of Things (IoT), age of information (AoI), and artificial intelligence (AI)-native wireless communications.



**Dusit Niyato** (Fellow, IEEE) received the BEng degree from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, in 1999, and the PhD degree in electrical and computer engineering from the University of Manitoba, Canada, in 2008. He is a professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests include the areas of sustainability, edge intelligence, decentralized machine learning, and incentive mechanism design.



**Zhu Han** (Fellow, IEEE) received the BS degree in electronic engineering from Tsinghua University, in 1997, and the MS and PhD degrees in electrical and computer engineering from the University of Maryland, College Park, in 1999 and 2003, respectively. From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, Maryland. From 2003 to 2006, he was a research associate with the University of Maryland. From 2006 to 2008, he was an assistant professor with Boise State University, Idaho. Currently, he is a John and Rebecca Moores professor

with the Electrical and Computer Engineering Department as well as in the Computer Science Department, University of Houston, Texas. His main research targets on the novel game-theory related concepts critical to enabling efficient and distributive use of wireless networks with limited resources. His other research interests include wireless resource allocation and management, wireless communications and networking, quantum computing, data science, smart grid, security and privacy. He received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, IEEE Leonard G. Abraham Prize in the field of Communications Systems (best paper award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He was an IEEE Communications Society Distinguished Lecturer from 2015–2018, AAAS fellow since 2019, and ACM distinguished Member since 2019. He is a 1% highly cited Researcher since 2017 according to Web of Science. He is also the winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: "for contributions to game theory and distributed management of autonomous communication networks."