Socially-Inspired Semantic Communication Codec Updating for NTN-Enabled Intelligent Transportation Systems

Guhan Zheng, *Member, IEEE*, Qiang Ni, *Senior Member, IEEE*, Keivan Navaie, *Senior Member, IEEE*, and Charilaos Zarakovitis, *Member, IEEE*

***Abstract*—In navigating the challenges of real-time semantic communication (SC) codec updates in the 6G-era non-terrestrial network (NTN)-assisted vehicular networks (NTN-VNs), a crucial component of intelligent transportation systems (ITS), this article introduces a novel approach inspired by human society. Facing complexities like 3-dimensional updating, network dynamism, and updating costs, NTN-VNs are treated as social networks. The proposed NTN-VN federated learning (NTN-VN-FL) framework asynchronously addresses challenges such as uplink and downlink SC codec updates, device decentralization, and asynchronous updating. By viewing device behaviors during updating as social behaviors with economic costs, an NTN-VN social management system ensures the proper functioning of the social network in the context of NTN-VN-FL. An economical social behavior selection mechanism, based on the reverse auction game for NTN-VN-FL, minimizes training delay and device energy costs, considering social relationships. The article also presents a two-stage Stackelberg game with the Vickrey auction rule to maximize social welfare in the auction. Simulation results highlight the superiority of NTN-VN-FL over existing potential application algorithms, effectively addressing the unique challenges of SC codec updating in NTN-VN. The efficacy of the social management system and social behavior selection mechanism is demonstrated in achieving optimal outcomes.**

***Index Terms*—Semantic communication, non-terrestrial network assisted vehicular networks, federated learning, social management, game.**

I. Introduction

In the era of 6G, characterized by advancements in Low Earth Orbit (LEO) satellites and aerial platforms, vehicles can access these facilities, benefiting from affordable, high-throughput, and seamless connectivity services [1], [2]. This led to the transformation of 2-dimensional (2D) vehicular networks (VNs), which is a key component of intelligent transportation systems (ITS). The 2D-VNs, often referred to as traditional terrestrial VNs, rely on terrestrial-based infrastructure such as fixed roadside units to provide connectivity. These VNs, while effective in urban and densely populated areas, face limitations in regions where deploying infrastructure is challenging or economically infeasible. The incorporation of LEO satellites and aerial platforms into traditional terrestrial VNs, forming 3-dimensional (3D) non-terrestrial network (NTN) enabled VNs, i.e., NTN-VNs, is recognized as a crucial element in the context of 6G [3]. In other words, vehicles can choose the most suitable access point based on their specific conditions, ensuring seamless communication. For instance, in remote areas like deserts or disaster-stricken regions where terrestrial services are unavailable, vehicles have the option to connect through existing aerial platforms or satellites.

T-ITS-24-10-4025Nevertheless, in NTN-VNs, the extensive coverage provided by satellites and aerial platforms leads to a substantial number of served vehicles, resulting in a significant volume of content transmission within a given time frame. A large number of vehicles place immense pressure on limited spectrum resources. Moreover, considering the extensive propagation distance and intricate channel conditions, ensuring the reliability of transmitted information becomes a critical concern. To tackle these challenges, Semantic Communication (SC) stands out as a transformative and pivotal technical enabler [4].

SC enables the extraction of semantic information through goal-oriented semantic codecs employing machine learning (ML) techniques. Unlike conventional communication paradigms that transmit all symbols and bits, SC encoders selectively transmit semantic information. This approach markedly enhances spectrum efficiency in VNs and augments the reliability of transmission information through ML-based decoding [5]. Envisionably, the adoption of SC codecs in vehicles for communication with non-terrestrial devices holds the potential to significantly boost the spectrum efficiency of NTN-VNs, resulting in increased network throughput and heightened reliability of transmission information. Hence, the integration of SC in NTN-VNs is deemed imperative [6]. Yet, the introduction of SC presents unique technical challenges in NTN-VN.

A. Challenges

Federated learning (FL) has demonstrated its efficacy as a machine learning approach for updating SC codecs in conventional terrestrial VNs [9]. In this approach, vehicles locally train their SC codec models for new content, which is then uploaded to a coordinator for model aggregation. This allows the aggregation of an efficient new SC codec model for all

Received 10 October 2024; revised 09 January 2025 and 09 February 2025; accepted 24 February 2025. This work was supported in part by the Open Networks Ecosystem Competition Western Open Radio Access Network(O-RAN) Deployment (ONE WORD) Project. (Corresponding author: Qiang Ni.)

G.Zheng, Q. Ni, and K. Navaie are with the School of Computing and Communications, Lancaster University, LA1 4WA, UK (Email: g.zheng2, q.ni, k.navaie@lancaster.ac.uk).

C. Zarakovitis is with the National Centre for Scientific Research Demokritos, Greece. (Email: c.zarakovitis@iit.demokritos.gr)

vehicles in the VN. Qin et al. [9] explored the potential of FL for SC codec updating, while Xie and Qin [10] proposed a lite federated SC system. However, the latter involves the training data leaking to devices, posing privacy risks to participants. Similarly, [11] and [8] both introduced different FL-based frameworks for training SC codec models for user-transmitted content.

While relevant, most existing works focus on 2D networks, addressing vehicles and terrestrial edges, making them less directly applicable to NTN-VNs. Additionally, current distributed updating schemes primarily cater to vehicles transmitting content (uplink) and overlook vehicles receiving content (downlink). Based on these considerations, we identify challenges in updating SC codecs in general terrestrial VNs, including the economic imperative for consistent encoder/decoder models [8] and the absence of a comprehensive SC updating framework and algorithm accommodating both uplink and downlink scenarios.

In the context of updating SC codecs in sophisticated NTN-VNs, several additional unique challenges emerge alongside the aforementioned issues. Unlike 2D-VNs, the application of existing FL-based studies for SC codec updating in NTN-VNs is not straightforward due to the intricate multi-layer framework inherent to NTN-VNs. Designing a multi-layer updating framework poses a significant challenge. Moreover, due to the high mobility of vehicles, vehicles may choose various devices for SC, such as terrestrial roadside units, aerial base stations, and satellites. There may be multiple devices in each layer of the NTN (terrestrial facility layer, aerial platform layer, and LEO satellite layer) whose SC encoders/decoders for serving vehicles require updating, reflecting the decentralized nature of the NTN-VNs.

Coordinating multi-device cooperative SC codec updates becomes complex, involving decisions on coordinator selection and accounting for device heterogeneity. The variable duration of each aggregation epoch, with the potential for asynchronous aggregation, adds an additional layer of complexity. Furthermore, compared to VNs, LEO satellites, acting as edges/base stations, exhibit dynamic movement, while vehicles are relatively geographically stationary.

Building upon the above considerations, the main challenges associated with updating SC codecs in NTN-VNs involve tackling multi-layer NTN-VNs, implementing updates in decentralized NTN-VNs, navigating asynchronous updates while accounting for device heterogeneity, and adapting to the dynamic nature of NTN-VNs.

B. Related works

In [6], a PSFed method was first proposed for 3D networks updating the SC codec for task offloading. Nevertheless, this method was still for uplink codec updating and neglected the decentralization and synchronicity of codec updating in NTN-VNs.

Several decentralized techniques leveraging FL show promise for addressing SC codec updates in NTN-VN. For instance, Guha Roy et al. [12] introduced the BrainTorrent framework, randomly designating a node as the coordinator for each federated aggregation epoch. However, concerns like latency and energy efficiency were overlooked in this approach. Similarly, Che et al.'s CMFL [13], inspired by practical Byzantine fault tolerance (PBFT), enhances training robustness and employs an election mechanism to select the coordinator based on trust value. Despite these advantages, CMFL lacks consideration for latency and energy efficiency during coordinator selection. Notably, these frameworks neglect the temporal dynamics and asynchronous nature of nodes, resulting in inefficient training and suboptimal resource utilization [14].

To address the challenge of asynchronous training, Ghosh et al. [15] proposed grouping devices with similar training and transmission capabilities into clusters. This approach utilizes a synchronous update strategy within clusters and an asynchronous update strategy between clusters. However, determining cluster divisions poses challenges in a decentralized dynamic environment. In [16], a semi-asynchronous learning approach, which aggregates only a portion of participating devices in each round, is explored. Nevertheless, it necessitates the availability of aggregation weights from the previous epoch at the time of aggregation. In a decentralized setting, a new aggregation node may lack the weights from the previous epoch. Additionally, Bian et al. [17] suggested using an unmanned aerial vehicle (UAV) as a relay node, but this introduces additional training resource consumption and is not applicable in a decentralized environment.

Moreover, GossipFL [18] tackles both asynchronous and decentralized challenges by having devices exchange models with a single peer at each communication round. While this approach results in devices possessing perfectly convergent ML models that may not be identical, maintaining uniform SC codec models after training is economically desirable. Furthermore, GossipFL faces limitations in adapting to dynamic environments as it necessitates fixed participants. Additionally, none of the previously mentioned frameworks are directly applicable to the NTN-VN.

C. Motivation and contributions

Given the above, we contend that there is an urgent need for a framework and algorithm facilitating SC codec updating in decentralized, asynchronous, and dynamic environments, with a specific focus on addressing training latency and energy concerns. This need is particularly essential in the context of NTN-VNs. Consequently, the primary aim of this paper is to introduce a framework and algorithm tailored for updating SC codecs in NTN-VNs, effectively tackling these challenges.

Drawing inspiration from human socialization, our novel proposal introduces the NTN-VN Federated Learning (NTN-VN-FL) framework for asynchronous SC codec updating across NTN-VNs. Within this innovative approach, vehicles and non-terrestrial devices autonomously forge social relationships and optimize social behavior and training decisions to align with both economic and social benefits. To ensure the stability of the social network and associated behaviors, we present the NTN-VN Social Management System tailored for NTN-VN-FL. Additionally, we delve into the selection of social behavior, considering aspects such as social relationships, delay, and energy. To address this, we introduce an economic social behavior selection mechanism grounded in various game theories.

The main contributions of this paper are as follows:

- Introducing a novel and efficient NTN-VN-FL framework, our proposal extends beyond updating uplink SC codecs to encompass the crucial aspect of updating downlinks as well. The primary objective of this framework is twofold: enhancing asynchronous training accuracy and ensuring convergence in dynamic and decentralized NTN-VNs. To strike a balance between technical factors such as delay and energy, we uniquely integrate sociological and economic concepts into NTN-VN-FL. Devices achieve economical SC codec updating through distinct social behaviors—namely, "leading," "following," and "plagiarizing"—while dynamically cultivating social relationships. This adaptive framework not only addresses the current challenges of SC codec updating but also anticipates future VNs, making significant strides in resolving both aspects of SC codec updating and codec updating in NTN-VNs.
- Our subsequent proposition introduces a social management system for NTN-VN-FL, meticulously designed to ensure the stability of both the "social network" and "social behavior" during training. Central to this system is the definition of "social relations," serving as a metric to quantify the impact of each social behavior on others. Trustors gain a comprehensive assessment of a trustee's social relations, allowing them to act in their best interests and optimize performance and welfare based on this evaluation. Empowering trustors further, they can make informed decisions about including a trustee in their social networks and participating in SC codec updating activities. This system's adaptive nature facilitates the dynamic adjustment of social management parameters in response to evolving social relationships, ensuring ongoing effectiveness.
- Additionally, we introduce an innovative game-theoretical and economic social behavior selection mechanism for NTN-VN-FL, leveraging the proposed social NTN-VN and social management system. This mechanism aims to optimize social welfare with a focus on fairness, taking into account social relationships, training delays, and energy costs. Our approach begins with the design of a reverse auction framework supporting many-to-one matching. This framework transforms the "leader" behavior selection problem into a two-fold challenge: "winner (i.e., leader) selection" and "bidding determination." To address the intricate mathematics involved in "bidding determination," we present a two-stage Stackelberg game approach. Furthermore, we apply the Vickrey auction rule to select the winner and determine their pricing, ensuring a comprehensive and effective economic social behavior selection process.

D. Organization of the paper

The remainder of the paper is structured as follows: In Section II, we present the model of NTN-VNs deployed with SC and outline the design of the NTN-VN-FL framework in Section III. Sections IV and V detail the presented social management system and social behavior selection mechanism for the proposed NTN-VN-FL. The performance of the framework and mechanisms is assessed through simulations in Section VI. Lastly, we conclude the paper in Section VII.

II. System Model

This section introduces the NTN-VN integrated with SC, outlining the computing and communication model adopted during SC codec updating. Additionally, we present the FL-based SC codec updating model, offering a comprehensive overview of the underlying processes and methodologies.

A. System description

We consider a holistic NTN-VN requiring the updating of uplink/downlink SC codecs for vehicles within a terrestrial area. This network is structured into four layers: the vehicles layer, terrestrial facility layer, aerial platform layer, and LEO satellite layer. Vehicles employ SC codecs for the transmission and reception of various types of information, such as images, speech, and videos. The transmitter extracts the meaning of the message, i.e., semantic information, using a semantic encoder and then transmits such information to a receiver deployed with a semantic decoder.

Building on prior research in distributed learning for 3D communication networks [19], [20], we assume that vehicles in the lowest layers necessitate an upper-layer device as a coordinator for cooperative learning. For example, models requiring updates in vehicles require an aerial platform/LEO satellite as a coordinator. However, due to constraints in latency, energy, upper-layer device coverage, etc., the optimal coordinator varies for different vehicles, and there may not be a single optimal coordinator. The vehicles' optimal coordinator may also change during training due to the high mobility of vehicles. These coordinators also may distributed at different layers.

This is grounded in FL, a widely used distributed learning framework. In FL, vehicles train their models locally, uploading only the model (without training data) to a coordinator in each epoch. The coordinator then aggregates the weighted models from participants and returns the aggregated model for the next training epoch. This process ensures rapid and efficient model training while preserving the privacy of training data. Additionally, the trained models are uniform, meeting the requirement for SC codec updating.

B. Computing models

We assume there is a set of $I = \{1, 2, ..., i, ..., I\}$ vehicles that need updating the SC codec in the NTN-VN. The local training delay of vehicle i is

$$T_i = \frac{M_i}{f_i}, \tag{1}$$

where M_i is the required CPU-cycle for updating the semantic codec. Further, f_i is the CPU-cycle at device i with the unit cycles/s. The training energy consumption is, therefore,

$$E_i = \varepsilon f_i^3 \frac{M_i}{f_i} = \varepsilon M_i f_i^2, \tag{2}$$

where εf_i^3 is the computing power and ε represents the impact of underlaying electronics [21].

In addition, cooperative training between multi-vehicles imposes communication costs due to sharing the SC codec model. During model sharing, the communication delay of the device i transmitting the SC codec model to the coordinator device j, is also estimated by

$$T_{i,j} = \frac{D}{r_{i,j}}, \tag{3}$$

where D is the size of the SC codec model and $r_{i,j}$ is the transmission rate between vehicle i and coordinator device j:

$$r_{i,j} = B_{i,j} \log(1 + \frac{p_{i,j} g_{i,j}}{\sigma^2}). \tag{4}$$

Here, $B_{i,j}$, $p_{i,j}$ and $g_{i,j}$ are bandwidth, transmission power, and the channel gain of the device i, respectively. Furthermore, σ^2 is the variance of the Gaussian white noise. The energy consumption of the communication is, therefore,

$$E_{i,j} = p_{i,j} T_{i,j} = \frac{p_{i,j} D}{B \log(1 + \frac{p_{i,j} g_{i,j}}{\sigma^2})}. \tag{5}$$

Furthermore, given the communication occurs within the NTN-VN, the extended propagation distance introduces additional considerations, including the impact of propagation delay. Therefore,

$$T_{i,j}^P = \frac{h_{i,j}}{c}, \tag{6}$$

where $h_{i,j}$ is the distance between vehicle i and device j and c is the speed of light.

C. FL-based SC codec training

In FL, vehicles train a model locally using their respective training data in each epoch. The objective of the training is to minimize the loss between the codec output and the input content:

$$\min \;\; L_i(\theta) \triangleq \frac{1}{m_i} \sum_{v_{i,n} \in m_i} f(\theta; v_{i,n}), \tag{7}$$

where θ is a vector representing the SC codec parameters, L_i is the loss function of device i and $f(\cdot)$ is the user-specified loss function. In the above, $v_{i,n}$ is the training data, m_i is the local dataset of device i and $m_i = \{v_{i,1}, v_{i,2}, ..., v_{i,n}, ..., v_{i,m_i}\}$.

After local training, the corresponding local model parameters, θ, are then submitted to the coordinator for federated aggregation. The aggregation process is based on the weighted average [22], therefore, the objective of aggregation is:

$$\min \;\; L_i(\theta) \triangleq \frac{\sum_{i \in F} m_i L_i(\theta)}{M}, \tag{8}$$

where F is the number of participants that share the model with the coordinator and $M = \sum_{i \in F} m_i$.

For easy reference, the main parameters and their description used throughout this paper are presented in Table I.

TABLE I: NOTATION DEFINITION

Symbol	Definition
I	Set of devices
T_i	Training delay of device i
$T_{i,j}$	Transmission delay between device i and j
$T_{i,j}^P$	Propagation delay between device i and j
E_i	Training energy consumption of device i
L_i	Loss of device i
θ	Global model
ω	Local model
C	Closeness
R_i	Effective bid for device i
S	Satisfaction
B	Bonus
V	Gain from increased closeness

III. NTN-VN-FL: SOCIAL FRAMEWORK DESIGN

In this section, we present the framework design of NTN-VN-FL. We also introduce the three types of social behavior in the proposed framework, including lead, follow, and plagiarize. This is followed by a detailed description of the NTN-VN-FL.

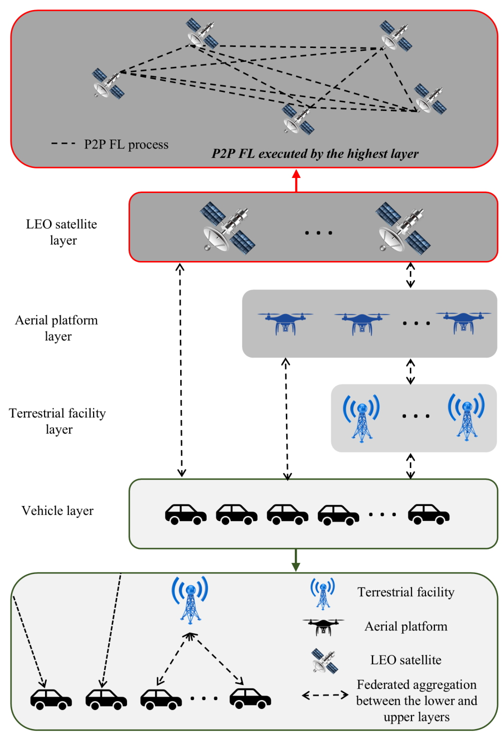


Fig. 1: The NTN-VN-FL framework.

A. Social behaviors

1) Leader: The leadership role is assumed by a device acting as a coordinator in the NTN-VN, and the leader's location dynamically adapts with each training epoch. The leader must set a duration limit for the model parameter collection stage within this epoch, choose the leader for the next epoch, and identify the device available for training. Additionally, it gathers high-volume semantic codec model parameters from followers for information sharing during devices' training, specifically for federated aggregation in FL studies.

2) Followers: Followers refer to devices transmitting SC codec model parameters to the leader during the duration limit of the model parameters collection stage in this epoch, essentially serving as distributed participants. Before transmission, followers are required to locally train/aggregate the SC codec model parameter θ.

3) Plagiarizer: Plagiarizers are devices incapable of transmitting SC model parameters to the leader within the duration limit of the model parameters collection stage in this epoch, i.e., straggling communication devices in federated learning frameworks. Similar to followers, they also need to locally train/aggregate the SC codec model parameter θ. The distinction lies in the fact that their global model θ for this epoch is acquired through plagiarism from their close partners in the same layer, according to sociological definitions. While plagiarism contradicts social values, it is a phenomenon found in society. The impact of plagiarism on the proposed NTN-VN society and measures to ensure its stability will be discussed in Section IV.

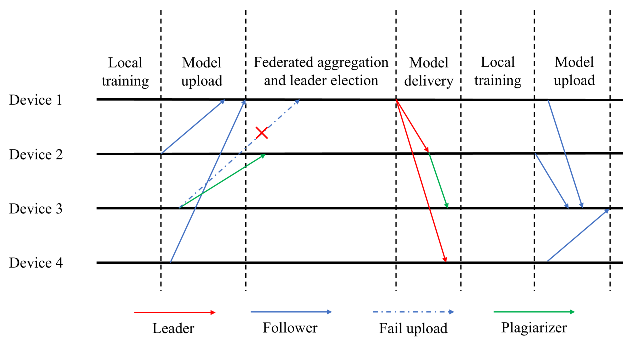


Fig. 2: Proposed NTN-VN-FL in the highest layer for asynchronous updating.

B. The proposed framework

In a concise overview, we assume an extensive SC codec update as an illustrative example, involving the participation of devices on every layer.

1) Updating uplink and downlink SC codecs in a 3D network: For the uplink SC codec, handling new transmission content from vehicles, functioning as followers or plagiarizers. These vehicles look for nodes in the upper layer to serve as coordinators/leaders based on criteria such as speed, coverage, delay, and energy. Eventually, the SC codec models are collected and sent to different satellites for the ultimate aggregation. These satellites collectively form a decentralized peer-to-peer (P2P) cellular network without a master node [23], necessitating the election of a leader for the final aggregation.

For the downlink SC codec, handling new transmission content from the upper participant layer, such as the LEO satellite layer, allows direct training at this layer and aggregates in the LEO satellite layer. Devices in this layer conduct local training and subsequently deliver the trained global SC decoder to vehicles.

Consequently, the challenges of SC codec updating in the NTN-VN for both uplink and downlink are ultimately streamlined into executing P2P FL at the highest layer involved in the updating, namely the LEO satellite layer. The distinction lies in the fact that uplink SC codec updating comprises two parts: 1) FL conducted between the lower and upper layers, and 2) P2P FL executed by the highest layer (refer to Fig. 1). Conversely, downlink SC codec updating only necessitates considering P2P FL performed at the highest layer.

2) Asynchronous updating: Here we first examine a P2P network comprising LEO satellites for FL. Given the heterogeneity among devices, the leader's time to collect aggregated models varies due to asynchronous training. In this context, we assume the leader is determined. For leader election mechanism refer to Section V.

In the asynchronous updating of NTN-VN-FL (see Fig. 2), at the initiation stage, the leader establishes the duration limit for the model parameters collection stage and communicates this to all devices. Notably, devices with more local training data may have longer training times based on Eq. (1), indicating that they carry more information. Therefore, the duration limit for the collection stage should consider these devices.

During the model parameters collection stage, satellites collect the trained SC codec model for initial aggregation and transmit it to the satellite leader where possible. If a satellite fails to transmit the model parameters within the stipulated time, it becomes a plagiarizer, sending a request to its "best friend" in the same layer for the newest global model. However, this comes at the cost of diminished social relations. If a device is sporadically plagiarized without remedy, it faces exclusion from training by society.

In the sharing stage, the leader performs model aggregation and sends it back to followers. Some followers then share this model with accredited plagiarizers.

In the election stage, devices share real-time information with the leader to collectively and fairly elect the next epoch leader, considering both fairness and economic efficiency.

Turning our perspective to FL performed between the lower and upper layers, devices in the lower layer, such as a terrestrial facility, can only choose to be a follower or a plagiarizer. If it fails to upload the model timely to the leader in the aerial platform layer, it may choose to plagiarize other terrestrial facilities. However, it also serves as the leader of the lower layer (i.e., the subscriber layer) and determines the time limit for model collection.

3) Network's dynamic and decentralization: For devices newly added to the training, they also receive the latest model through plagiarism behavior, similar to plagiarizers. Furthermore, we will outline social behavior selection in this decentralized environment in Section V.

C. Discussion

We first consider the uplink SC codec updating. If vehicle i is a follower in training epoch k, the updating objective is the same as in (7). The iteration for local updating of the SC codec model is

$$\theta_{k+1}^{i} = \theta_k - \eta_i \bigtriangledown L_i(\theta_k), \tag{9}$$

where θ_k is the global model sent from the leader in epoch k and η_i is the learning rate of vehicle i. In addition, θ_{k+1}^{i} is the local model parameter of vehicle i in epoch $k+1$.

If subscriber vehicle i is a plagiarizer, the updated parameters ω_{k+1}^{i} of plagiarizer i in epoch $k+1$ is:

$$\omega_{k+1}^{i} = \theta_{k}^{i'} - \eta_i \bigtriangledown L_i(\theta_{k}^{i'}), \tag{10}$$

where $\theta_{k}^{i'} = \frac{\theta_{k}^{'}+\rho_i\omega_{k}^{i}}{1+\rho_i}$. Here, ρ_i is the weight parameter and $\theta_{k}^{'}$ is the SC codec model parameter via weighted average. Since no staleness exists, i.e., vehicles have the latest global model for each epoch, we have $\omega_{k+1}^{i} = \theta_{k+1}^{i}$ if the plagiarizer participates in the epoch $k+1$ aggregation, based on [16].

Therefore, the leader j in the terrestrial facility layer performs the global updating by

$$\theta_{k+1}^{j} = \frac{\sum_{i\in F_{k+1}^{j}} m_i \theta_{k+1}^{j}}{\sum_{i\in F_{k+1}^{j}} m_i}, \tag{11}$$

where F_{k+1}^{j} is the number of subscribers who choose the device j as the leader. The terrestrial facility j can update θ_{k+1}^{j} and upload it to the leader in the aerial platform layer. Similarly, the final global model updated by a satellite leader can be denoted by the same formula, i.e., (10). Different from (7), the aggregation in (10) is partial model aggregation. Nevertheless, partial model aggregation is still rapidly convergent as shown in [24] and [25].

Moreover, our framework presentation is predicated on the involvement of devices across all layers, namely, the subscriber layer, terrestrial facility layer, aerial platform layer, and LEO satellite layer. Nevertheless, it can be implemented across any number of layers. This adaptability stems from the framework's division into two main components: 1) FL executed between the lower and upper layers, and 2) P2P FL performed by the highest layer. For instance, when the downlink SC codec of devices in the aerial platform layer requires updating, they can undergo local training and submit to corresponding leaders in the LEO satellite layer, i.e., FL. Subsequent satellites then execute decentralized aggregation, i.e., P2P FL, of these models.

IV. NTN-VN-FL: SOCIAL MANAGEMENT

We introduce an Individual Social Relation Metric (ISRM) for social management, enabling individual devices to assess the social relationships of other nodes during training. Each device utilizes ISRM to determine its willingness to share information with plagiarizers and to elect a preferred leader. For example, when low-layer devices (e.g., terrestrial devices) seek an upper-layer device (e.g., aerial platform device) as the leader, each low-layer device can solicit information from several surrounding upper-layer devices and calculate its ISRM. By evaluating the ISRM of different upper-layer devices, the low-layer device can elect the optimal upper-layer device as the leader based on the chosen behavior mechanism.

ISRM is also employed by the leader to assess whether a device can participate in training. For instance, when a low-layer device (e.g., terrestrial device) applies to aggregate the training model in an upper-layer leader device (e.g., aerial platform device), the upper-layer device can request information from several surrounding low-layer devices and calculate the device's ISRM. Based on the ISRM, the upper-layer device can determine if the low-layer device is eligible to join the aggregation. Furthermore, ISRM is applicable in the highest-layer P2P network, with inspection devices being replaced by other devices on the same layer.

In alignment with sociological theories, trust plays a pivotal role in measuring the dynamics of social relations [26]. In this paper, we evaluate NTN-VN device social relations based on social trust, categorizing our ISRM into direct trust and indirect trust, akin to conventional trust metrics [27], [28].

A. Direct trust

We utilized three key social trust attributes, drawing inspiration from prior studies on the Internet of Things (IoT) social trust [29], [30]. These attributes include community interest, friendship similarity, and honesty, as these three trust attributes are some of the most salient indicators to characterize social IoT systems.

Community-interest trustworthiness [31] serves as a metric for assessing whether the trustor node and trustee nodes belong to the same social community. In the context of devices, community-interest trustworthiness can be determined by factors such as shared workplaces or similar capabilities. The community-interest trust value (CITV) of trustee node j from trustor node i is expressed as follows:

$$T_C(i,j) = \frac{|C_i \cap C_j|}{|C_i|}, \tag{12}$$

where $|\cdot|$ denotes the cardinality of a set, and C_i and C_j denote the set of community interests of trustor i and trustee j, respectively.

Friendship similarity trustworthiness is a metric used to assess whether the trustor node and trustee nodes share similar social relationships. This measure considers the extent of interaction a device has had with other nodes for specific content or tasks recently. The friendship similarity trust value (FSTV) of trustee node j from trustor node i can therefore be determined by

$$T_F(i,j) = \frac{|F_i \cap F_j|}{|F_i| - 1}, \tag{13}$$

where F_i and F_j denote the set of friends of trustor i and trustee j, respectively.

The honesty trustworthiness indicates whether or not a node is honest. A dishonest node is extremely susceptible as a malicious node and incurs significant damage to the network [29]. We write

$$T_H(i,j) = \begin{cases} \frac{|H_{i,j} - H_u|}{|H_{i,j}|} e^{-\frac{H_u}{H_{i,j}}}, \\ 0, \quad \text{if dishonest} \end{cases} \tag{14}$$

where $H_{i,j} = \sum \rho_{i,j} h_{i,j}$ is the number of total interactions between trustor i and trustee j. Here, $\rho_{i,j}$ is the interaction importance weight factor of the single interaction $h_{i,j}$. Further, $H_u = \sum \rho_u h_u$ is the number of unsuccessful interactions between them.

Therefore, we have the device i's direct trust in device j as:

$$T_D(i,j) = \alpha_1 T_C(i,j) + \alpha_2 T_F(i,j) + \alpha_3 T_H(i,j), \tag{15}$$

where α_1, α_2 and α_3 are weight parameters.

B. Indirect trust

Indirect trust is employed by the trustor to exchange information about the trust relationships with the trustee among other nodes. In conventional IoT social trust mechanisms [29], [30], we have

$$T_I(i,j) = \frac{\sum_{n=1, n\neq i,j}^{I} T_D(n,j)}{I-2}. \tag{16}$$

C. Closeness-associated ISRM

However, whether considering direct or indirect trust, the conventional assessment of social trust in IoT fails to account for the influence of specific social behaviors, such as leading, following, or plagiarizing. Consequently, it struggles to accurately map social relations. Furthermore, within social networks, we contend that not every device involved in exchanging trust information for indirect trust holds the same value.

Fortunately, sociality theories emphasize that closeness is a key indicator for identifying opinion leaders [32]. Information from opinion leaders has a greater impact on influencing the attitudes of follower devices [33]. Additionally, in line with academic plagiarism, instances of plagiarism tend to decrease the closeness among individuals involved [34]. Hence, we firstly and innovatively introduce the sociological concept of closeness into our proposed novel social management system.

Closeness is primarily associated with group type and interaction type [35], [36]. Consequently, we assess the closeness of devices based on the average interaction times at a given moment and the ownership object relationship. Here we express the closeness $C_t(i,j)$ of device i and device j related to average interaction times as:

$$C_t(i,j) = \frac{1}{t}\sum_{h=1}^{H_{i,j}} \omega_h, \tag{17}$$

where t is the assessment time and ω_h is the importance of interaction h. In addition, the ownership object relationship presents among objects belonging to the same owner [31], e.g., whether they belong to the same mobile network operator (MNO). For devices, it can be considered the homogeneity proportion of devices' software or hardware. Hence, we write the closeness $C_s(i,j)$ of device i and device j related to the ownership object relationship as:

$$C_s(i,j) = \frac{|S_i \cap S_j|}{|S_i \cup S_j|}, \tag{18}$$

where S_i and S_j are the set of ownership objects.

Therefore, we denote the closeness between node i and node j as:

$$C(i,j) = \beta_1 C_t(i,j) + \beta_2 C_s(i,j), \tag{19}$$

where β_1 and β_2 are weight parameters.

In addition, when a plagiarizer j wishes to copy training information from device i, the closeness of the device i towards plagiarizer j is inevitably reduced. Because the occurrence of plagiarism decreases the corresponding closeness. Closeness $C(i,j)$ in each training epoch should be rewritten as:

$$C(i,j) = \beta_1 C_t(i,j) + \beta_2 C_s(i,j) - \upsilon\lambda, \tag{20}$$

where λ is the plagiarism penalty factor and $\lambda = 0$, if no plagiarism or the closeness before plagiarism is less than the closeness threshold Γ. Furthermore, υ is the weight factor related to the training epoch. In case all devices' closeness of a plagiarizer device decreases under Γ, plagiarism is also not possible, thus forcing the withdrawal of training. In addition, choosing to lead or follow both can increase closeness due to interaction.

Therefore, we can update the indirect trust value of trustee j from the trustor i by

$$T_I(i,j) = \frac{\sum_{n=1,n\neq i,j}^{I} C(n,j) T_D(n,j)}{\sum_{n=1,n\neq i,j}^{I} C(n,j)}. \tag{21}$$

Combine with direct trust, indirect trust, and [21]–[25], the ISRM for trustor i and trustee j can be expressed by

$$T_{ISRM}(i,j) = \gamma_1 T_D(i,j) + \gamma_2 T_I(i,j), \tag{22}$$

where γ_1 and γ_2 are weight parameters related to closeness $C_s(i,j)$ between node i and node j, and $T_{ISRM}(i,j) = 0$ if $T_{ISRM}(i,j)$ is less than the security threshold Υ.

Therefore, the trustor only needs to be informed about some main friends around the trustee to obtain an accurate social relationship with the trustee, rather than a comprehensive, all-device-involved assessment. Additionally, the decrease in closeness results in a decrease in social relations. After dropping below the threshold, the plagiarizing device will not be included in the training unless behaviors are performed that increase closeness. In our frameworks, being a leader increases the number of socializations with other devices in the same layer, thus improving closeness.

V. NTN-VN-FL: Selecting the Social Behavior of the Nodes

In this section, we explore the optimal selection of social behavior within NTN-VN social networks. For FL performed between the lower and upper layers, the involvement of multiple plagiarizers is precluded from FL aggregation due to the diminished closeness and social relations resulting from multiple instances of plagiarism. While social relations can be utilized to ascertain a device's availability for training, it does not influence the social behavior selection in the lower layer. Additionally, numerous studies have extensively discussed follower selection for coordinators (leaders), as seen in [2], [21] and our previous work [6].

Therefore, we do not delve into that aspect here. This section concentrates on behavior selection in the highest-layer P2P decentralized environment. Individual devices' varied social behaviors incur diverse computational and communication costs, influencing changes in social relationships. Drawing inspiration from economics in the social realm, we quantify computational and communication costs as economic costs resulting from these social behaviors. The optimal social behavior can then be derived based on the economic costs associated with social behaviors and the dynamic changes in the social environment.

In a society where only one leader can be elected, we have designed a reverse auction game to determine the optimal leader in each epoch, with the remaining devices acting as followers. Those who fail to keep up with training in a timely manner are categorized as plagiarists. The social behavior selection aims to optimize social welfare, considering factors such as social relations, fairness, training delays, and energy costs.

In this game, the NTN-VN (i.e., MNO) functions as the buyer seeking to acquire aggregation resources from devices. Devices, acting as bidders, aim to sell their computing and communication resources to maximize revenue and secure the position of a leader. As a result, the game can be divided into a "winner (i.e., leader) selection" problem and a "bidding determination" problem which involves determining the bidding strategy of devices.

A. Winner selection

On the NTN-VNs/MNOs, the sooner a functional semantic codec is trained, the earlier it can be deployed, thereby generating benefits. We can define the effective bid for device j to be a leader as:

$$R_j = S_{j,u}R_u + S_{j,d}R_d + S_{j,c}R_c, \tag{23}$$

where $S_{j,u}$, $S_{j,d}$ and $S_{j,c}$ are the un-aggregated semantic codec transmission (followers to leader) delay satisfaction, aggregated semantic codec transmission (leader to followers) delay satisfaction, and aggregation delay satisfaction of the MNO in this training epoch, respectively. Further, R_u, R_d and R_c are the corresponding greatest revenue from MNOs. Similar to previous studies [37], [38], we set the satisfaction as the same logarithmic function. Here there exist I devices in this society and hence have the $S_{j,u}$, $S_{j,d}$ and $S_{j,c}$ as:

$$S_{j,u} = \ln(1 + \vartheta_u - \max\{T_{i,j} + T^P_{i,j}\}),\ \forall i \in I, j \neq i \tag{24}$$

$$S_{j,d} = \ln(1 + \vartheta_d - \max\{T_{j,i} + T^P_{j,i}\}),\ \forall i \in I, j \neq i \tag{25}$$

$$S_{j,c} = \ln(1 + \vartheta_c - T_j), \tag{26}$$

where ϑ_u, ϑ_d and ϑ_c are parameters to ensure that satisfaction indicators are available in most situations. Moreover, $T_{i,j}$, $T^P_{i,j}$, $T_{j,i}$, $T^P_{i,j}$ and T_j are the transmission delay (i to j), propagation delay (i to j), transmission delay (j to i), propagation delay (j to i), and aggregation computing delay of device j, respectively.

These bids are determined by the T_j and $T_{j,i}$ from device j and $T_{i,j}$ from device i. Mathematically, to maximize social welfare, the "winner selection" problem can be expressed as:

Problem 1:

$$\max_{x_i, R_i} \quad \sum_{i=1}^{I} R_i x_i, \tag{27a}$$

$$s.t. \quad \sum_{i=1}^{I} x_i = 1, \tag{27b}$$

$$x_i = \{0, 1\}, \tag{27c}$$

$$q < I/2, \tag{27d}$$

where x_i is the decision variable and q is the number of plagiarizers. It is observed that the winner can be chosen in descending order, subject to obtaining key elements of bids. e.g., T_j, $T_{i,j}$ and $T_{j,i}$.

B. Bidding determination

We note that the leader's bidding R_j from a device, j is related to $T_{i,j}$ from another device i. How leader j motivates them to generate the $T_{i,j}$ that is optimal for them and hence increasing the bidding has become an issue. To address the bidding problem in a reverse auction, increase the training motivation, and determine $T_{i,j}$, we further proposed a two-stage Stackelberg game approach. Before that here, we first discuss the utility of different devices in case device j bids for leadership.

1) Utilities for followers: For a follower i when the leader is j, participating in training costs them energy. We formulate the energy cost as the followers' monetary cost. Moreover, to motivate the followers to participate in the training with more power and thus increase their auction chips, the leader should give some bonuses to followers. Influenced by a combination of social relations, without loss of generality, the utility function for followers can be expressed as:

$$u_{i,j}(T_{i,j}, b_j) = B_{i,j} + V_{i,j} - \varphi E_{i,j}, \tag{28}$$

where $B_{i,j}$ is the bonus that device i receives from device j and b_j is the unit bonus price decided from device j related to transmission delay. To ensure fairness among devices, a function related to delay satisfaction is set to model $B_{i,j}$, therefore,

$$B_{i,j} = b_j \ln(1 + \vartheta_u - T_{i,j} - T^P_{i,j}). \tag{29}$$

Moreover, $V_{i,j}$ is the gains of follower i due to increased closeness/trust value via this training when the leader is j. During training, the closeness value and trust value increase for each successful interaction. Furthermore, a device with fewer interactions with other devices gets a lower interaction demand. The demand v_i is therefore inversely proportional to total closeness, i.e., $v_i \propto \frac{1}{\sum_{i=1,i\neq j}^{I} C(i,j)}$. We have the follower $V_{i,j}$ as:

$$V_{i,j} = \gamma v_i = \frac{\gamma}{\sum_{i=1,i\neq j}^{I} C(i,j)}, \tag{30}$$

where γ is the monetary factor.

Therefore, we write the optimization problem for device i as:

Problem 2:

$$\max_{T_{i,j}, b_j} b_j \ln(1 + \vartheta_u - T_{i,j} - T^P_{i,j}) + \frac{\gamma}{\sum_{i=1,i\neq j}^{I} C(i,j)} - \varphi E_{i,j}, \tag{31a}$$

$$s.t. \quad T_i^{min} \le T_{i,j}, \tag{31b}$$

where T_i^{min} is the minimum transmission delay related to maximum transmission power.

2) Utility for the leader: Similar to the utility function for followers, we denote the leader's utility function by

$$U_j(T_{i,j}, T_{j,i}, T_j, b_j) = R_j + V_j - \varphi E_j - \varphi E_{j,i} - B_j, \tag{32}$$

where V_j is the gains due to increased closeness. Furthermore, $E_{j,i} = \sum_{i=1,i\neq j}^{I} p_{i,j} T_{j,i}$ and E_j are the total communication energy cost and aggregation computing energy cost of a device j, respectively. The factor φ is the monetary factor to convert energy consumption into monetary cost and $B_j = \sum_{i=1,i\neq j}^{I} B_{i,j}$ is the total bonus for encouraging non-aggregating nodes to transmit speedily to the device j. Note that, R_j (in (23)) is the network/MNO revenue achieved by device j if bidding is successful.

Similar to Eq. (30), we can express the leader V_j by

$$V_j = \frac{(I-1)\gamma}{\sum_{i=1,i\neq j}^{I} C(j,i)}, \tag{33}$$

The optimization problem for an aggregation node is, therefore,

Problem 3:

$$\max_{T_{i,j},T_{j,i},T_j,b_j} R_j + \frac{(I-1)\gamma}{\sum_{i=1,i\neq j}^{I} C(i,j)} - \varphi E_j - \varphi E_{j,i} - B_j, \tag{34a}$$

$$s.t. \quad T_j^{min} \leq T_j, \tag{34b}$$

$$T_{j,i}^{min} \leq T_{j,i}, \tag{34c}$$

where T_j^{min} is the minimum computing delay related to the maximum available CPU-cycle frequency f_j and $T_{j,i}^{min}$ is the minimum transmission delay related to the maximum available transmission power. Moreover, for simplicity and fairness, we assume that $T_{j,i}$ for each device i are the same.

3) Two-stage Stackelberg game for bidding determination: We note that the strategies of the leader and followers are all related to $T_{i,j}$ and b_j. Therefore, we construct a two-stage Stackelberg game. First, the optimization problem of aggregation devices is divided into three sub-problems. The two optimization problems are then solved by convex optimization methods in the first stage. We then formulate a Stackelberg game which is treated in Stage 2.

In the first step, we can decompose three sub-optimization problems from Problem 3, i.e.,

Problem 4:

$$\max_{T_j} S_{j,c}R_c - \varphi E_j, \tag{35a}$$

$$s.t. \quad T_j^{min} \leq T_j, \tag{35b}$$

Problem 5:

$$\max_{T_{i,j},b_j} S_{j,u}R_u + V_j - B_j, \tag{36a}$$

$$s.t. \quad T_i^{min} \leq T_{i,j}, \tag{36b}$$

Problem 6:

$$\max_{T_{j,i}} S_{j,d}R_d - \varphi E_{j,i}, \tag{37a}$$

$$s.t. \quad T_j^{min} \leq T_{j,i}, \tag{37b}$$

For problem 4, the first-order derivative equation of Eq. (35) can be written as $\frac{2M_j^3\varphi\varepsilon}{T_j^3} - \frac{1}{\vartheta_c - T_j + 1}$. In the domain of definition, this problem can be easily solved by basic convex optimization methods.

For problem 5, $T_{i,j}$ and b_j need to be achieved from the joint strategies of the leader and followers. Based on the above, we design a Stackelberg game as the second step to address this problem. We have

$$U_j(T_{i,j}^*, b_j^*) \geq U_j(T_{i,j}^*, b_j), \tag{38}$$

Algorithm 1 Social behavior selection mechanism

Initialization: $T_{i,j}$, $T_{j,i}$, T_j, b_j, the training epochs K, the maximum number of iterations Z, the stopping criterion threshold $\xi > 0$, step length τ, and the leader position, first winner bid $R_{j_0} = R_1$, second winner bid $R_{j_1} = R_1$, winner $n = 1$

for each device $j = 1, 2, ..., I$:
  $T_j \leftarrow$ Eq.(35) & $T_{j,i} \leftarrow$ Eq.(37)
  while $z < Z$:
    $T_{i,j} \leftarrow$ SMO
    $b_j = b_j + \tau$
    Adjust τ, reduce the value of τ
    Until $U_j(z) - U_j(z-1) < \xi$
  end while
  $R_j \leftarrow$ Eq.(20)
end for
for each device $j = 1, 2, ..., I$:
  if $R_j \geq R_{j-1}$:
    $R_{j_0} = R_j \& R_{j_1} = R_{j_0} \& n = j$
  end if
end for
Leader $\leftarrow$ device n
Pricing $\leftarrow R_{j_0}^{'} = R_{j_1}$

$$u_{i,j}(T_{i,j}^*, b_j^*) \geq u_{i,j}(T_{i,j}, b_j^*), \tag{39}$$

where $T_{i,j}^*$, b_j^* are the maximum social welfare solution.

<u>**NE Existence:**</u> According to the Debreu-Glicksberg-Fan theorem [39], a pure Nash Equilibrium (NE) exists when the strategy set of followers is both compact and convex. Further, the $u_{i,j}$ should be continuous and concave in $T_{i,j}$. The second-order partial derivative of Eq. (28) with respect to $T_{i,j}$ thus is:

$$\frac{\partial^2 \mu_{i,j}}{\partial T_{i,j}^2} = -\frac{b_j}{(T_{i,j} + T_{i,j}^P - \vartheta_u - 1)} - \frac{2^{\frac{D}{BT_{i,j}}} D^2 \varphi \sigma^2 ln^2 2}{B^2 T_{i,j}^3 g_{i,j}}. \tag{40}$$

Combining practical communication systems, $\frac{\partial^2 \mu_{i,j}}{\partial T_{i,j}^2} < 0$ and Eq. (28) is concave in $T_{i,j}$. Since the strategy set of followers is also compact and convex, a pure NE exists.

To find the NE point, the followers' strategies need to be derived first followed by the leaders' strategy based on the backward induction. Specifically, we should obtain the relational $T_{i,j} \leftrightarrow f(b_j)$ via $\frac{\partial u_{i,j}}{\partial T_{i,j}} = 0$, where $f(b_j)$ is b_j related function. The $T_{i,j}$ in U_j should be replaced by $f(b_j)$ and b_j^* can be obtained via $\frac{\partial U_j}{\partial b_j} = 0$.

However, the solution of $T_{i,j}$ cannot be obtained in a closed form. As $u_{i,j}$ is concave, for such a non-linear equilibrium problem, sequential minimal optimal (SMO) is a candidate powerful tool to achieve the optimal solution with constraints. The similar problem 6 can be derived from the same approach. We first give a fixed value b_j. Following the SMO, the $p_{i,j}$ could be tightened iteratively until convergence. The b_j^* is subsequently obtained by the step length acceleration method.

Algorithm 2 NTN-VN-FL

Initialization: dataset m_i, leader i, global SC codec model θ

for each epoch $k = 1, 2, ..., K$:
  for each follower device $a = 1, 2, ..., A$:
    $\theta^a_{k+1} = \theta_k - \eta_a \bigtriangledown L_a(\theta_k)$
    Upload θ^a_{k+1} and bidding R_a to the leader i
  end for
  for each plagiarizer device $b = 1, 2, ..., B$:
    Plagiarize a close friend and get punished
    $\theta^{b'}_k = \frac{\theta_k + \rho_b \omega^b_k}{1+\rho_b}$
    $\omega^b_{k+1} = \theta^{b'}_k - \eta_b \bigtriangledown L_b(\theta^{b'}_k)$
    Upload ω^b_{k+1} and bidding R_b to the leader i
  end for
  Leader i receives weights and verifies that the device can participate in the training
  Leader i: $\theta_{k+1} \leftarrow$ Eq.(10)
  Leader i delivery the θ_{k+1} and elect the leader according to the bidding
end for

4) Pricing: Determining the leader j_0 and corresponding R_{j_0} we then obtain the allocation prices according to the Vickrey auction rule. By the Vickrey auction rule, the actual delivery bid $R^{'}_{j_0}$ is determined according to the next winner j_1's bid, i.e., R_{j_1}. If leader j_0 is not achievable, the social trust value is reduced to zero according to Eq. (14). The leader j_0 thus should set a follower training deadline and if followers fail to achieve the requirement, they will become plagiarizers. The procedure of the social behavior selection mechanism is demonstrated in Algorithm 1. Further, the full procedure of the NTN-VN-FL is presented in Algorithm 2 (downlink as an example)

VI. Simulation Results

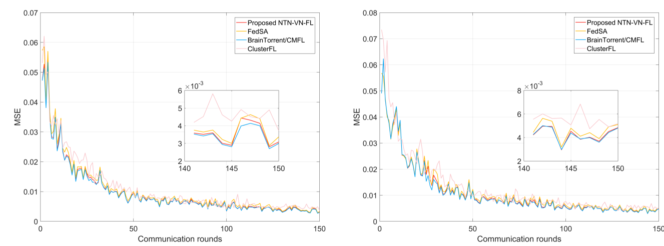
In this section, we present simulation results to evaluate the performance of the proposed NTN-VN-FL framework, conducting a primary comparison with four state-of-the-art learning frameworks: BrainTorrent [12], CMFL [13], ClusterFL [15], and FedSA [16]. The details of these methods could refer to related works in the Introduction.

Our focus is on devices in the highest layer, specifically the satellite layer, utilizing the SC codec for transmission images with vehicles on the terrestrial. The semantic codecs from vehicles and satellites are aggregated on the satellite layer. As satellites in space can be likened to a cellular network over a terrestrial area [23], we model a satellite network over a terrestrial area as a 1-tier cellular edge network. The choice of 1-tier is influenced by the limited number of satellites over a terrestrial area simultaneously [2]. All satellite devices collaborate to train an image semantic transmission codec model. Due to asynchrony, one device consistently takes longer to train than others.

TABLE II: The setting of the semantic codec.

Encoder	Neuron num	Decoder	Neuron num
Conv+PReLU	32	transConv+PReLU	32
Conv+PReLU	32	transConv+PReLU	32
Conv+PReLU	32	transConv+PReLU	32
Conv+PReLU	32	transConv+PReLU	32
Conv+PReLU	32	transConv+PReLU	16

Unless stated otherwise, the plagiarizer attempts to become a leader twice in every 10 epochs. Consequently, in ClusterFL settings, devices are divided into two clusters. We retrain an efficiency-designed end-to-end SC codec model following the previous SC codec study [40] in such an environment. Consistent with [40], the training data are sourced from the CIFAR-10 [41] and CIFAR-100 [42] image datasets. We assume all satellite devices are in the same state and have a few training images. The convolutional autoencoder (CAE) is considered the semantic encoder and semantic decoder [6], [41]. The specific settings of the CAE are detailed in Table II.



(a) CIFAR 10 dataset (b) CIFAR 100 dataset

Fig. 3: Accuracy and convergence speed of various learning frameworks in a static environment.

In Fig. 3, we illustrate the convergence speed and accuracy of various frameworks. In this paper, accuracy is defined as the mean square error between the input image and the transmitted image. Peak Signal-to-Noise Ratio (PSNR) serves as a crucial metric for evaluating the semantic communication model's performance [40], [43]. A higher PSNR signifies more accurate semantic communication encoder transmission.

$$PSNR = 10 \log \frac{MAX^2}{MSE} (dB), \tag{41}$$

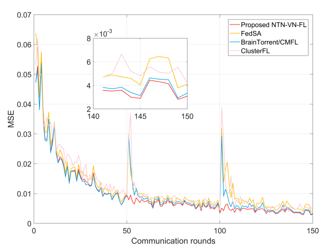
where MAX is the maximum value for a pixel and MSE is the mean square error. The accuracy increases with a lower mean square deviation for the same dataset. For simplicity, we, therefore, denote the performance metric as MSE, the same as [6]. It is evident that all frameworks exhibit nearly identical convergence speeds. Our framework demonstrates comparable accuracy to BrainTorrent and CMFL. However, these latter frameworks neglect the impact of asynchrony, requiring an extended period for federated aggregation. In contrast, NTN-VN-FL achieves superior accuracy compared to ClusterFL and FedSA. The latter frameworks struggle in decentralized scenarios due to ineffective measures and the presence of a staleness model, resulting in the loss of training information and global weights from laggard devices in the last epoch. The proposed social management system in NTN-VN-FL mitigates model staleness and the constant sharing and aggregation of

TABLE III: Accuracy of different models.

	CIFAR-10	CIFAR-100	CIFAR-10 Dynamic	CIFAR-100 Dynamic
BrainTorrent/CMFL	100%	100%	100%	100%
NTN-VN-FL(1)	102.95%	101.45%	92.20%	91.66%
NTN-VN-FL(2)	102.43%	101.21%	91.73%	91.44%
NTN-VN-FL(3)	101.39%	100.73%	90.79%	91.01%
FedSA	112.53%	106.21%	122.51%	132.07%
ClusterFL	124.56%	107.72%	125.92%	136.48%

information by plagiarists aspiring to become leaders further contribute to its success.

Fig. 4 depicts the accuracy and convergence speed of various frameworks in dynamic environments, considering the introduction of a new participant in epochs 50 and 100. Notably, NTN-VN-FL consistently attains the highest accuracy, irrespective of whether the dataset is CIFAR-10 or CIFAR-100. This consistent performance is attributed to the presence of the proposed social management system. The integration of a new participant into the training process, facilitated by plagiarism, allows the newcomer to obtain the latest global model without disrupting the accuracy of the overall model.

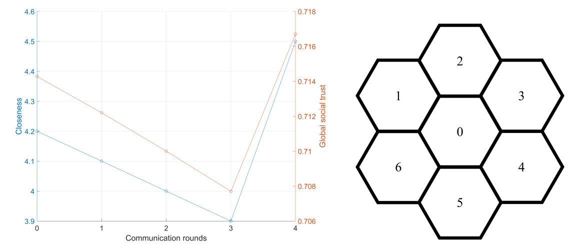


(a) CIFAR 10 dataset (b) CIFAR 100 dataset

Fig. 4: Accuracy and convergence speed of various learning frameworks in a dynamic environment.

In Table III, we present the impact of social behavior on accuracy. Assuming a trusted device resorts to repeated plagiarism due to unforeseen circumstances, it seeks leadership roles through the social behavior selection mechanism to restore closeness, social trust, etc. We explore scenarios where the device applies to become a leader 1, 2, and 3 times in every 10 epochs, denoted as NTN-VN-FL(1), NTN-VN-FL(2), and NTN-VN-FL(3). Employing BrainTorrent/CMFL as a benchmark with 100% accuracy in Mean Squared Error (MSE), lower percentages in our results indicate more accurate image transmission.

It is seen that there is a positive correlation between the frequency of leadership roles and accuracy. Increased leadership instances contribute to higher accuracy as the sharing of training results and data information becomes more frequent. Notably, in non-dynamic environments, NTN-VN-FL demonstrates accuracy comparable to BrainTorrent/CMFL but outperforms FedSA and ClusterFL by approximately 5% to 12%. In dynamic environments, NTN-VN-FL exhibits the highest accuracy, surpassing all frameworks by 9% to 36%, showcasing the superior performance of our NTN-VN-FL model.



(a) Relations variation (b) Labelling of satellites

Fig. 5: Closeness and social relations variation of the plagiarizer.

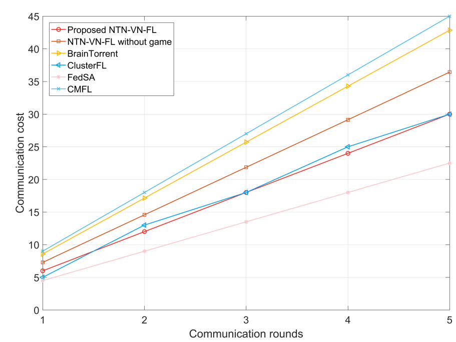


Fig. 6: Communication cost of various learning frameworks.

In Fig. 5, we depict the fluctuation in closeness and social relations of a plagiarizer. The determination of social relations is based on the leader's assessment. For the illustration, we designate satellite "6" (see Fig. 5 (b)) as the plagiarizer. To set the initial conditions, we assign a closeness and social relations value of 0.8 to satellites "0", "1", and "5", while satellites "2", "3", and "4" are assigned values of 0.6. The observed trend reveals a decrease in the social relations of satellite "7" as the training commences. This decline is a consequence of its plagiarism from the closest devices, namely "0", "1", and "5," resulting in reduced aggregation weights despite high social relations. Both closeness and social relations approach precarious thresholds. Recognizing this, in communication round 4, satellite "6" undergoes an auction to become the leader, leading to an increase in closeness and threshold values.

Fig. 6 illustrates the fluctuation in communication costs during the training of various frameworks. We assume uniform trust values and closeness for all devices. To enhance clarity, we express communication costs in terms of the number of communication hops per training epoch. The initial aggregation node is designated as device "6," and half of the participants operate asynchronously.

Observing the results, CMFL incurs the highest communication costs, primarily due to its social behavior selection relying on the device's trust value. This results in an aggregation process akin to traditional FL when the leader's trust value remains constant. Conversely, FedSA exhibits the lowest communication costs, as only half of the devices participate in the aggregation process. However, this efficiency comes at the expense of losing some training information.

The communication cost of our proposed NTN-VN-FL, integrated with a game, falls between ClusterFL and FedSA, attributed to the inclusion of a plagiarism mechanism. Notably, our framework is tailored for decentralized, dynamic, and 3D networks, outperforming these counterparts in achieving higher transmission accuracy. Furthermore, NTN-VN-FL, enhanced by economic game theory, not only improves accuracy but also achieves this with reduced communication costs.

VII. Conclusion

In this paper, we explored the challenges associated with SC codec updating in 3D NTN-VN networks and introduced a pioneering framework called NTN-VN-FL. Leveraging sociological concepts such as closeness, NTN-VN-FL categorizes vehicles and devices as "leaders," "followers," and "plagiarizers," optimizing various social and training behaviors. Technical challenges within NTN-VN-FL, including dynamic, training latency, and energy costs, were scrutinized. To address these challenges, a social trust management system was introduced to uphold the stability of NTN-VN-FL. Furthermore, a social behavior selection mechanism, based on economic game theories like the reverse auction game and Stackelberg game, and aligned with the designed social trust management system, was proposed. This mechanism reduces training communication costs, taking into account the trust value and closeness of devices. Simulation results showcased the superiority of our proposed NTN-VN-FL frameworks and related approaches, outperforming baseline frameworks by 5% to 12%. In dynamic environments, NTN-VN-FL exhibited the highest accuracy, surpassing all frameworks by 9% to 36%. Moreover, in the 3D NTN-VN dynamic environment, transmission accuracy saw an improvement of 9% to 36% relative to baseline frameworks.

References

[1] Y. Liu, K. Xiong, Y. Lu, Q. Ni, P. Fan and K. B. Letaief, "UAV-Aided Wireless Power Transfer and Data Collection in Rician Fading," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 10, pp. 3097-3113, Oct. 2021.

[2] R. Deng, B. Di, S. Chen, S. Sun and L. Song, "Ultra-Dense LEO Satellite Offloading for Terrestrial Networks: How Much to Pay the Satellite Operator?," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6240-6254, Oct. 2020.

[3] E. C. Strinati, S. Barbarossa, T. Choi, A. Pietrabissa, A. Giuseppi, E. De Santis, J. Vidal, Z. Becvar, T. Haustein, N. Cassiau, F. Costanzo, J. Kim, and I. Kim, "6G in the sky: On-demand intelligence at the edge of 3D networks (Invited paper)," *ETRI J.*, vol. 42, no. 5, pp. 643–657, Oct. 2020.

[4] Y. Shi, Y. Zhou, D. Wen, Y. Wu, C. Jiang and K. B. Letaief, "Task-oriented communications for 6G: Vision principles and technologies", arXiv:2303.10920, 2023, [online] Available: http://arxiv.org/abs/2303.10920.

[5] X. Luo, H. -H. Chen and Q. Guo, "Semantic Communications: Overview, Open Issues, and Future Research Directions," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 210-219, February 2022.

[6] G. Zheng, Q. Ni, K. Navaie, and H. Pervaiz, "Semantic Communication in Satellite-borne Edge Cloud Network for Computation Offloading," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 5, pp. 1145-1158, May 2024.

[7] Z. Qin, X. Tao, J. Lu, W. Tong and G. Y. Li, "Semantic communications: Principles and challenges", arXiv:2201.01389, 2021.

[8] G. Zheng, Q. Ni, K. Navaie, H. Pervaiz and C. Zarakovitis, "A Distributed Learning Architecture for Semantic Communication in Autonomous Driving Networks for Task Offloading," *IEEE Communications Magazine*, vol. 61, no. 11, pp. 64-68, November 2023.

[9] Z. Qin, G. Y. Li and H. Ye, "Federated Learning and Wireless Communications," *IEEE Wireless Communications*, vol. 28, no. 5, pp. 134-140, October 2021.

[10] H. Xie and Z. Qin, "A Lite Distributed Semantic Communication System for Internet of Things," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 142-153, Jan. 2021.

[11] G. Shi, Y. Xiao, Y. Li and X. Xie, "From Semantic Communication to Semantic-Aware Networking: Model, Architecture, and Open Problems," *IEEE Communications Magazine*, vol. 59, no. 8, pp. 44-50, August 2021.

[12] A. Guha Roy, S. Siddiqui, S. Pölsterl, N. Navab and C. Wachinger, "BrainTorrent: A Peer-to-Peer environment for decentralized federated learning", arXiv:1905.06731, 2019, [online] Available: http://arxiv.org/abs/1905.06731.

[13] CMFL: C. Che, X. Li, C. Chen, X. He and Z. Zheng, "A Decentralized Federated Learning Framework via Committee Mechanism with Convergence Guarantee," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 12, pp. 4783-4800, 1 Dec. 2022.

[14] C. Xu, Y. Qu, Y. Xiang, and L. Gao, "Asynchronous Federated Learning on Heterogeneous Devices: A Survey", arXiv:2109.04269, 2023, [online] Available: http://arxiv.org/abs/2109.04269.

[15] A. Ghosh, J. Chung, D. Yin and K. Ramchandran, "An Efficient Framework for Clustered Federated Learning," *IEEE Transactions on Information Theory*, vol. 68, no. 12, pp. 8076-8091, Dec. 2022.

[16] Q. Ma, Y. Xu, H. Xu, Z. Jiang, L. Huang and H. Huang, "FedSA: A Semi-Asynchronous Federated Learning Mechanism in Heterogeneous Edge Computing," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 12, pp. 3654-3672, Dec. 2021.

[17] J. Bian, C. Shen, and J. Xu, "Joint Client Assignment and UAV Route Planning for Indirect-Communication Federated Learning", arXiv:2304.10744, 2023, [online] Available: http://arxiv.org/abs/2304.10744.

[18] Z., S. Shi, B. Li and X. Chu, "GossipFL: A Decentralized Federated Learning Framework with Sparsified and Adaptive Communication," *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, no. 3, pp. 909-922, 1 March 2023.

[19] N. Razmi, B. Matthiesen, A. Dekorsy and P. Popovski, "Ground-Assisted Federated Learning in LEO Satellite Constellations," *IEEE Wireless Communications Letters*, vol. 11, no. 4, pp. 717-721, April 2022.

[20] H. Chen, M. Xiao and Z. Pang, "Satellite-Based Computing Networks with Federated Learning," *IEEE Wireless Communications*, vol. 29, no. 1, pp. 78-84, February 2022.

[21] Z. Song, Y. Hao, Y. Liu and X. Sun, "Energy-Efficient Multiaccess Edge Computing for Terrestrial-Satellite Internet of Things," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 14202-14218, 15 Sept.15, 2021.

[22] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," *Proc. Int. Conf. Artif. Intell. Stat. (AISTATS)*, vol. 54, 2017, pp. 1273–1282.

[23] L. D. Earley, "Communication in Challenging Environments: Application of LEO/MEO Satellite Constellation to Emerging Aviation Networks," *2021 Integrated Communications Navigation and Surveillance Conference (ICNS)*, 2021, pp. 1-8.

[24] H. Wu and P. Wang, "Node Selection Toward Faster Convergence for Federated Learning on Non-IID Data," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 5, pp. 3099-3111, 1 Sept.-Oct. 2022.

[25] X. Li, K. Huang, W. Yang, S. Wang and Z. Zhang, "On the convergence of FedAvg on non-IID data", *Proc. Int. Conf. Learn. Represent. (ICLR)*, pp. 1-26, 2020.

[26] O. Schilke, M. Reimann, and K. S. Cook, "Trust in Social Relations," *Annual Review of Sociology*, vol. 47, no. 1, Apr. 2021.

[27] O. Samuel, N. Javaid, A. Khalid, M. Imrarn and N. Nasser, "A Trust Management System for Multi-agent System in Smart Grids using Blockchain Technology," *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Taipei, Taiwan, 2020, pp. 1-6.

[28] H. El-Sayed, H. Alexander, P. Kulkarni, M. A. Khan, R. M. Noor and Z. Trabelsi, "A Novel Multifaceted Trust Management Framework for

Vehicular Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 20084-20097, Nov. 2022.

[29] I. -R. Chen, F. Bao and J. Guo, "Trust-Based Service Management for Social Internet of Things Systems," *IEEE Transactions on Dependable and Secure Computing*, vol. 13, no. 6, pp. 684-696, 1 Nov.-Dec. 2016.

[30] U. Jayasinghe, G. M. Lee, T. -W. Um and Q. Shi, "Machine Learning Based Trust Computational Model for IoT Services," *IEEE Transactions on Sustainable Computing*, vol. 4, no. 1, pp. 39-52, 1 Jan.-March 2019.

[31] L. Atzori, A. Iera, G. Morabito and M. Nitti, "The social internet of things (SIoT)-When social networks meet the internet of things: Concept architecture and network characterization", *Comput. Netw.*, vol. 56, no. 16, pp. 3594-3608, Nov. 2012.

[32] S. Wassermann, K. Faust, "Social Network Analysis – Methods and Applications", *Cambridge University Press*, Cambridge, 1999.

[33] F. Bodendorf and C. Kaiser, "Detecting opinion leaders and trends in online social networks", *Proceedings of the 2nd ACM workshop on social web search and mining*, pp. 65-68, 2009.

[34] S. Senders. "Academic plagiarism and the limits of theft." *Originality, imitation, and plagiarism: Teaching writing in the digital age*, 195-207, 2008.

[35] L. Yang, Y. Qiao, Z. Liu, J. Ma and X. Li, "Identifying opinion leader nodes in online social networks with a new closeness evaluation algorithm", *Soft Comput.*, vol. 22, no. 2, pp. 453-464, Jan. 2018.

[36] A. Aron, E.N. Aron, D. Smollan, "Inclusion of other in the self scale and the structure of interpersonal closeness", *Journal of Personality and Social Psychology*, pp. 596-612, 1992.

[37] Y. Wang et al., "Task Offloading for Post-Disaster Rescue in Unmanned Aerial Vehicles Networks," *IEEE/ACM Transactions on Networking*, vol. 30, no. 4, pp. 1525-1539, Aug. 2022.

[38] H. Zhou, Z. Wang, G. Min and H. Zhang, "UAV-Aided Computation Offloading in Mobile-Edge Computing Networks: A Stackelberg Game Approach," *IEEE Internet of Things Journal*, vol. 10, no. 8, pp. 6622-6633, 15 April15, 2023.

[39] D. Fudenberg and J. Tirole, Game Theory, *MIT Press*, 1993.

[40] D. B. Kurka and D. Gündüz, "DeepJSCC-f: Deep Joint Source-Channel Coding of Images with Feedback," *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 1, pp. 178-193, May 2020.

[41] A. Krizhevsky, V. Nair, and G. Hinton, "Cifar-10 (Canadian Institute for Advanced Research)." [Online]. Available: http://www.cs.toronto.edu/kriz/cifar.html

[42] A. Krizhevsky, "Learning multiple layers of features from tiny images," Univ. Toronto, Toronto, ON, Canada, Tech. Rep. TR-2009, 2009.

[43] Y. Shao and D. Gunduz, "Semantic Communications with Discrete-Time Analog Transmission: A PAPR Perspective," *IEEE Wireless Communications Letters*, vol. 12, no. 3, pp. 510-514, March 2023.