车联网中联邦学习模型低时延传输迁移方法研究

王帅[1, 2]，尹宏博[2]，江池[2]，张科[1, 2]，张引[1, 3]
（1. 电子科技大学（深圳）高等研究院，广东 深圳 518110；2. 电子科技大学信息与通信工程学院，四川 成都 611731；
3. 广东省智能机器人研究院，广东 东莞 523830）

摘 要：联邦学习因其分布式与隐私保护特性，在车联网数据安全领域中引起广泛关注。异步联邦学习机制能够更好地适应车辆算力网络状态的动态变化，在提升全局模型更新效率的同时，实现对本地隐私数据的有效保护。然而，恶意车辆在联邦学习训练中可能进行中毒攻击，上传恶意模型至全局模型，进而影响正常车辆的本地训练。在模型下发时，增加候选模型数量虽可提升规避恶意模型的概率，却会显著增加通信时延，影响系统性能。为了平衡安全性与时延，提出一种联邦学习模型传输迁移方法，对城市道路中移动车辆与路边单元（RSU, roadside unit）的交互过程以及模型下发安全性进行建模，通过强化学习优化车辆对 RSU 的传输迁移策略，在保证模型下发安全性的同时有效降低通信时延。仿真结果表明，该方法相较于基线方法平均传输时延降低了约 7%，验证了其在安全性与通信时延方面的优势。

关键词：车联网；联邦学习；时延优化；强化学习；传输迁移
中图分类号：TN915.08
文献标志码：A
doi：10.11959/j.issn.2096−3750.2026.00525

Research on low-latency transmission migration method for federated learning models in the Internet of vehicles

Wang Shuai[1,2], Yin Hongbo[2], Jiang Chi[2], Zhang Ke[1,2], Zhang Yin[1,3]
1. Shenzhen Institute for Advanced Study, UESTC, Shenzhen 518110, China
2. School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
3. Guangdong Intelligent Robot Research Institute, Dongguan 523830, China

Abstract: Federated learning, due to its distributed and privacy-preserving characteristics, has attracted widespread attention in the field of data security in vehicular networks. The asynchronous federated learning mechanism can better adapt to the dynamic changes of vehicle computing power and network conditions, and at the same time improve the efficiency of global model updates and realize effective protection of local privacy data. However, the malicious vehicles in federated learning training may perform poisoning attacks by uploading malicious models to the global model, which in turn affects the local training of normal vehicles. During model dissemination, although increasing the number of candidate models can improve the probability of avoiding malicious models, it will significantly increase communication latency and affect system performance. To balance security and latency, a federated learning model transmission migration method was proposed. The interaction process between moving vehicles and roadside units (RSUs) on urban roads were modeled, as well as the security of model dissemination. Through reinforcement learning, the vehicle-to-RSU transmission migration strategy was optimized, ensuring the security of model dissemination while effectively reducing communication latency. Simulation results show that, compared with baseline methods, the proposed method reduces the average transmission la-

收稿日期：2025−08−28；**修回日期：**2025−09−28
通信作者：张引，zhangyin123@uestc.edu.cn
基金项目：广东省重点研发计划项目（No. 2024B1111060001）
Foundation Item: The Key Research and Development Program of Guangdong Province (No. 2024B1111060001)

tency by about 7%, which verifies its advantages in terms of security and communication latency.

Key words: Internet of vehicles, federated learning, latency optimization, reinforcement learning, transmission migration

0　引言

随着智能交通系统的快速发展，车联网（IoV, Internet of vehicles）已成为推动智能城市建设和自动驾驶技术落地的核心基础设施[1]。车联网通过实现车辆与路边单元（RSU, roadside unit）、云服务器及其他车辆之间的高效通信，赋能协同感知、路径优化与交通调度等关键应用，在提升道路安全性和通行效率方面发挥着重要作用。然而，由于车联网天然具有的开放性、动态性和异构性，其在实际运行中面临着多重挑战[2-4]：一方面，车辆节点高速移动导致通信连接频繁切换，网络环境极度不稳定，进而引发通信时延升高与任务中断风险；另一方面，传统集中式数据处理模式无法兼顾隐私保护与实时性需求，大量敏感数据在传输与处理过程中容易遭受泄露。因此，如何在动态复杂的车联网环境中，实现高效、低时延且具有鲁棒性的模型传输与智能调度，已成为构建下一代智能交通系统亟须解决的重要问题。

联邦学习[5-6]（FL, federated learning）作为一种分布式机器学习范式，被广泛应用于车联网场景中。通过将模型训练任务分发至车辆本地执行，并聚合各方上传的模型参数，联邦学习有效避免了原始数据的集中化传输，在保护数据隐私的同时减轻了中心服务器的负担。然而，传统同步联邦学习[7-9]要求各客户端同步上传参数，这在高动态、多变连接质量的车联网环境下往往会导致大量上传等待与资源浪费，严重影响系统效率。此外，频繁的模型下发过程还可能受到车辆高速运动、信道条件波动以及安全鲁棒性约束等多重因素的干扰，易造成通信中断、任务失败甚至模型污染等问题，进一步加剧模型下发过程中的不确定性与不稳定性[10-11]。

针对上述问题，目前已有许多研究者提出相应的解决方案。Wang等[12]提出面向边缘协同的层次式联邦学习，通过在边缘侧分层聚合与自适应调度，减少全网同步开销与时延，然而对于鲁棒性考虑不足。对此，Nguyen等[13]提出FLAME，通过有针对性的鲁棒聚合与自适应滤除机制显著提升全局模型的抗后门能力，但它的侧重点在于安全聚合，没有结合车联网的高移动性与下行时延联合优化。Zhang等[14]提出考虑车辆移动性与信道动态的联合优化框架，将FL落实到车联网场景中，然而多集中在上行训练参与聚合，对下行模型下发的RSU迁移稳定性与低时延优化涉及较少。

近年来，随着车联网系统复杂度和动态性的不断提升，传统静态优化方法在面对多变的网络状态和不确定的环境因素时逐渐显现出局限性。为此，研究者开始尝试引入强化学习[15-17]（RL, reinforcement learning）等智能决策方法，利用其在高维状态空间中实时学习与自适应策略调整的能力，来优化联邦学习中的通信与迁移决策。强化学习通过与环境的交互学习策略，在无须依赖明确数学建模或先验知识的前提下，能够探索并逼近长期最优回报，在非静态、部分可观测的系统中尤为适用[18]。因此，其在异构设备协同、动态资源分配和鲁棒优化等问题中展现出强大的适应性与泛化能力[19-20]。例如，Yao等[21]用深度强化学习优化IoV中的动态计算卸载，能在变化的网络与计算资源下实时决策，但典型的深度Q网络（DQN, deep Q-network）面临Q值过估计与离散动作粒度受限。对此，Tang等[22]提出DFO-DDQN，通过双网络估计与帧级调度提升训练稳定性，但对于连续控制和多目标权衡有所不足。而Li等[23]引入DDPG用于车载边缘计算的联合卸载与资源分配，能在连续动作空间学习更细粒度功率与资源控制，但由于缺乏剪切更新与熵正则带来的稳定探索，易受超参数敏感性与分布漂移影响。

结合以上思路，本文提出一种基于近端策略优化（PPO, proximal policy optimization）[24]的联邦学习模型低时延传输迁移方法，通过构建多约束优化框架，在保证系统安全性的前提下，动态优化RSU与车辆之间的模型传输迁移策略，从而有效降低系统传输时延。

本文研究的主要贡献如下。

(1) 针对车联网中模型下发过程中对低时延传输效率与系统安全性的需求，构建了一个多约束优化问题模型，综合考虑车辆的移动特性、通信速率、安全性保障等因素，有效降低模型传输时延。

(2) 鉴于传统方法难以在动态环境中同时满足多重约束条件并保障高效模型传输，提出了一种低时延联邦学习模型传输迁移方法，有效实现了在动态车联网环境下的高效策略选择。

(3) 通过在城市道路车辆网络环境中的仿真实验，验证了所提方法的有效性。实验结果表明，该方法能够显著降低模型下发时延，优于传统基线方法。

1　系统模型

本文车联网系统模型如图1所示，在该网络中存在M个RSU，用集合$m \in \mathcal{M} = \{1, 2, \cdots, M\}$表示，$N$辆移动车辆用集合$n \in \mathcal{N} = \{1, 2, \cdots, N\}$表示，以及一个基站（BS, base station）。BS维护全局模型库，并周期性地将更新后的模型同步至各个RSU。RSU接收到BS下发的模型后，形成本地模型池，并每次选择K个模型下发给车辆，每个模型数据量记为D_{model}。车辆根据当前通信状态与信道条件，从RSU中选择一个最优的RSU建立连接并接收模型。由于车辆具有高移动性，在下发过程中可能驶入另一个覆盖范围更优的RSU区域，此时系统会触发关联策略迁移，保证时延最小化。车辆在完成模型接收后，结合本地数据进行模型更新与聚合，并将结果上传至RSU，RSU将聚合结果进一步上传至BS，BS对全局模型库进行更新，并将最新模型再次同步至所有RSU，实现联邦学习的迭代优化。

此外，记$D(t)$为t时刻尚未完成下传至车辆的模型数据量，用于刻画车辆在模型获取过程中的任务进度。为了便于分析，整个任务周期T可以离散为N_t个时隙，时隙持续时间为$\Delta = T/N_t$，时隙索引用$t \in T = \{1, 2, \cdots, N_t\}$表示。

2　系统建模与问题表述

2.1　通信模型

车辆在行驶过程中，不同RSU的带宽资源、信道状态以及车辆的运动状态都会对传输速率产生影响。基于此，本文将RSU m，$m \in \mathcal{M}$到车辆n，$n \in \mathcal{N}$的传输速率表示为

$$R_{m,n}(t) = \frac{B}{b}\text{lb}\left(1 + \text{SINR}_{m,n}(t)\right), \forall m \in \mathcal{M}, n \in \mathcal{N} \quad (1)$$

其中，B为信道总带宽，b为RSU可用信道数量，$\text{SINR}_{m,n}(t)$可以表示为

$$\text{SINR}_{m,n}(t) = \frac{\alpha_{m,n} P_m g_{m,n}(t)}{\sum\limits_{i \neq m} \alpha_{i,n} P_i g_{i,n}(t) + \sigma_0} \quad (2)$$

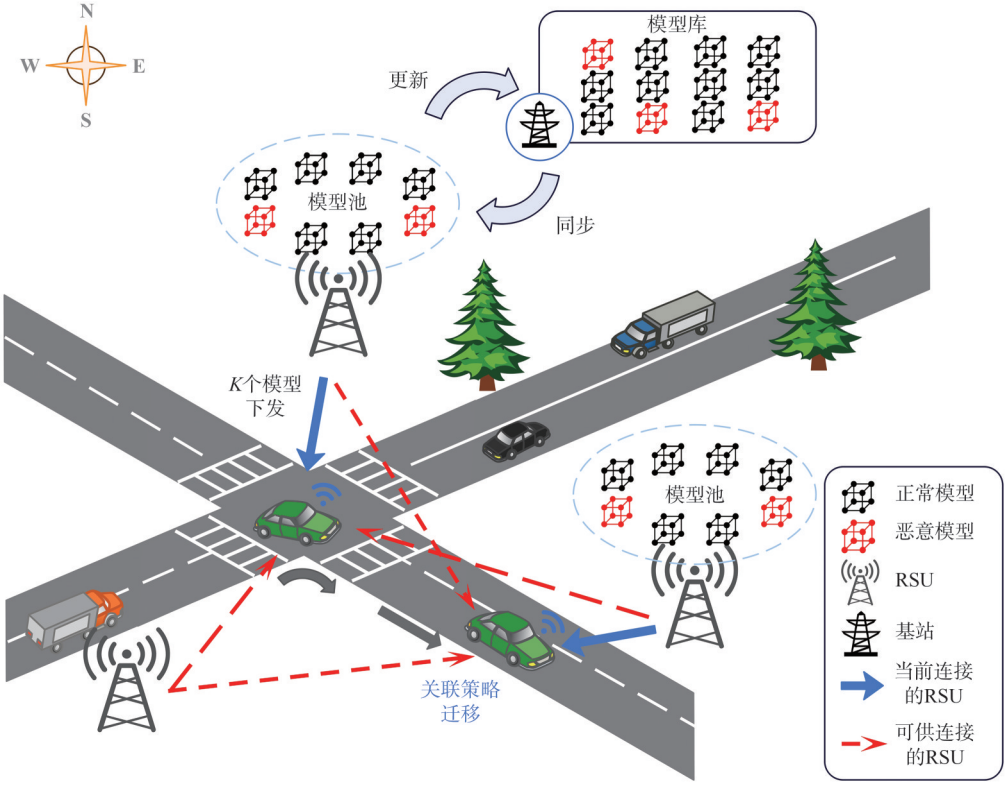

图1　车联网系统模型

其中，P_m表示RSU m的发射功率，σ_0为高斯白噪声，$\alpha_{n,m}(t)\in\{0,1\}$表示RSU m与车辆n之间的关联变量，当RSU m和车辆n在t时刻相关联时$\alpha_{n,m}(t)=1$，否则$\alpha_{n,m}(t)=0$。$g_{m,n}(t)$代表t时刻RSU m的信号增益，根据3GPP Release18[25]标准可以表示为

$$g_{m,n}(t)=10^{-\frac{\mathrm{PL}}{10}} \tag{3}$$

$$\mathrm{PL}=32.4+21\lg\left(d_{m,n}(t)\right)+20\lg\left(f_c\right)+\sigma_{\mathrm{SF}} \tag{4}$$

其中，$d_{m,n}(t)$为RSU m与车辆n在t时刻的欧氏距离，f_c表示通信链路所使用的中心频率，σ_{SF}为阴影衰落项。进一步，为了保证关联决策的唯一性和传输过程的稳定性，假定同一时刻一辆车最多与一个RSU相关联，即满足约束

$$\sum_{m=1}^{M}\alpha_{m,n}(t)\leqslant 1,\quad \forall n\in\mathcal{N},t\in T \tag{5}$$

在RSU与车辆进行数据传输的过程中，由于车辆行驶速度较快且运动状态具有不确定性，可能在传输过程中出现车辆进入另一个信号覆盖更优的RSU范围的情况。为提高传输效率并降低中断风险，引入传输迁移机制：当车辆尚未完成模型接收时，系统可将当前传输任务动态迁移至相邻信道条件更优的RSU继续执行，从而实现更快速稳定的模型下发过程。

该策略不仅要考虑车辆的当前位置、行驶方向与剩余模型数据量，还应结合RSU的信道状态情况，最小化整体模型传输时延。由于RSU中均同步有完整的模型，且车辆状态信息的数据量相较于模型大小可以忽略不计，迁移过程中仅考虑模型下发产生的时延。

为了计算RSU m到车辆n的实时距离，本文设定地图正北方向为y正半轴，正东方向为x正半轴，并假设在t时刻的车辆当前行驶速度v_t与x轴正半轴的夹角为θ_t，车辆t时刻位置与RSU的距离如图2所示，则车辆在t时刻坐标的计算可以分为两种情况。

(1) 车辆当前行驶状态为直线行驶，那么可以将车辆n在t时刻的坐标表示为

$$\begin{aligned}x_n(t)&=x_n(t-1)+v_t\Delta\cos\theta_t\\ y_n(t)&=y_n(t-1)+v_t\Delta\sin\theta_t\end{aligned} \tag{6}$$

(2) 车辆行驶状态为曲线行驶，设定车辆在t时刻行驶曲率为$k(t)$，单个时隙内曲率保持恒定，可以

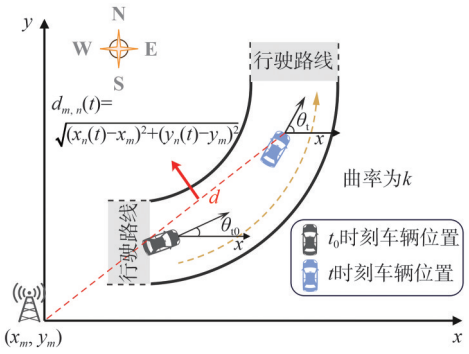

图2　车辆t时刻位置与RSU的距离

得出

$$\begin{cases}k(t)<0,\text{顺时针运动}\\ k(t)>0,\text{逆时针运动}\\ k(t)=0,\text{直线运动}\end{cases} \tag{7}$$

那么车辆速度方向与x轴正半轴夹角角度变化θ_t可以表示为

$$\theta_t=\theta_{t-1}+k(t)v_t\Delta \tag{8}$$

因此，车辆在曲线行驶状态下的坐标为

$$\begin{aligned}x_n(t)&=x_n(t-1)+\int_{t-1}^{t}v_t\cos\theta_t\mathrm{d}t\\ y_n(t)&=y_n(t-1)+\int_{t-1}^{t}v_t\sin\theta_t\mathrm{d}t\end{aligned} \tag{9}$$

可分别表示出RSU m的坐标$l_m^R=\left(x_m,y_m\right)$和车辆$n$的坐标$l_n^V=\left(x_n(t),y_n(t)\right)$，根据上述计算式可以计算出RSU m到车辆n的距离为

$$d_{m,n}(t)=\sqrt{(x_n(t)-x_m)^2+(y_n(t)-y_m)^2} \tag{10}$$

那么，t时刻尚未完成下发的模型数据量$D(t)$可以表示为

$$D(t)=D(t-1)-R_{m,n}(t)\Delta \tag{11}$$

在每个时间步t，通信时间$T_{\mathrm{com}}(t)$会根据剩余需要下传的数据量$D(t)$与当前传输速率$R_{m,n}(t)$进行更新。若在当前时间步内能够完成剩余数据的传输，则通信时间会按照比例递增，否则通信时间直接增加一个时间步。其递推计算式可以表示为

$$T_{\mathrm{com}}(t)=\begin{cases}T_{\mathrm{com}}(t-1)+\dfrac{D(t)}{R_{m,n}(t)\Delta}, & D(t)<R_{m,n}(t)\Delta\\ T_{\mathrm{com}}(t-1)+1, & \text{其他}\end{cases} \tag{12}$$

最终，当完成所有剩余数据传输后，可以得到RSU m下发给车辆n的传输时延T_n^{down}

$$T_n^{\mathrm{down}}=T_{\mathrm{com}}(t)\Delta \tag{13}$$

2.2 安全性建模分析

在动态车联网中，传统同步联邦学习方法因依赖所有客户端同时上传模型，常在异构通信条件与车辆高频移动环境下导致训练中断或效率低下。为此，尹宏博等[26]提出了一种异步鲁棒联邦学习方法，该方法允许车辆在无须全局同步的情况下异步获取候选模型、进行本地更新，并上传训练结果以更新系统模型池。整个过程通过RSU与BS的协同控制，实现了模型的分发、聚合与更新，提升了训练效率与系统灵活性。

在车联网的异步联邦学习场景中，恶意攻击主要指通过恶意车辆提交伪造或污染的模型更新，以干扰全局模型的训练。具体来说，恶意车辆通过实施中毒攻击等手段，在其本地训练过程中刻意扭曲模型参数，从而使得其提交的模型更新偏离真实数据的分布，导致全局模型性能下降，进而通过模型下发影响正常车辆的本地聚合。

为降低RSU下发恶意模型对正常车辆本地训练的影响，针对异步鲁棒联邦学习的安全性进行了建模与分析。结合前期研究，RSU在每轮训练中向车辆发送的候选模型数量K对系统规避恶意模型的概率具有决定性影响。具体而言，若设模型池中模型总数为L，恶意车辆的比例为p^m，则恶意模型数量可近似为

$$L_{\text{mal}} = Lp^m \tag{14}$$

在RSU随机向车辆提供K个候选模型时，候选集中包含x个恶意模型的概率为

$$P_1(X=x) = \frac{C_{L_{\text{mal}}}^{x} C_{L-L_{\text{mal}}}^{K-x}}{C_L^K} \tag{15}$$

进一步的，车辆通过模型选择算法从K个候选模型中选择a个模型进行聚合，仅选择正常模型的概率可表示为

$$P_{\text{normal}} = \sum_{x=0}^{K-a}\left[P_1(X=x)\cdot P_2(x)\right] \tag{16}$$

其中，$P_2(x)$为在给定恶意模型数量x时，车辆选择的a个聚合模型全为正常模型的条件概率。通过前期工作的模型选择算法，第二阶段模型选择过程可近似为直接选择精度最高的k个模型。所以进一步可得

$$P_{\text{normal}} \to \sum_{x=0}^{K-a} \frac{C_{L_{\text{mal}}}^{x} C_{L-L_{\text{mal}}}^{K-x}}{C_L^K} \tag{17}$$

该分析表明，增加K能够显著提高规避恶意模型的概率，从而增加系统安全性。然而，K的增大也意味着更多模型需要在RSU与车辆之间进行传输，必然导致通信时延的增加。因此，单纯追求更大的K并非最优方案。为此，本文将在上述安全性分析的基础上，引入通信时延建模，构建一个兼顾安全性和通信效率的联合优化目标，以确定最优的K值。

2.3 问题建模

在考虑车辆高速移动带来的模型下发中断风险的基础上，并结合车辆与RSU之间的通信速率约束、模型数据量大小D_{model}以及车辆接收模型的安全性约束，优化车辆对RSU的模型传输选择策略$\alpha(t)$和车辆的模型下发决策，最小化整个任务周期内模型下发的总时延。为此，本文将RSU的通信速率、车辆的运动状态以及模型传输的连续性进行联合建模，以确保车辆在有限的通信时隙内能够完成模型数据的高效下传与更新，从而提升联邦学习的系统性能。因此，建立如下优化问题

$$\begin{aligned}
&\mathrm{T} = \min_{(K,\boldsymbol{\alpha})} \sum_{n=1}^{N} T_n^{\text{down}} \\
\text{s.t.}\quad & K_{\text{safe}} \leqslant K,\ K \in \mathbb{Z}, \\
& P_{\text{normal}}(K) \geqslant P_{\text{normal}}(K_{\text{safe}}), \\
& v_{\min} \leqslant v_n(t) \leqslant v_{\max}, \quad \forall n \in \mathcal{N} = \{1,2,\cdots,N\}, \\
& 0 \leqslant D(t) \leqslant KD_{\text{model}}, \\
& -\pi \leqslant \theta \leqslant \pi, \\
& \sum_{m=1}^{M} \alpha_{n,m}(t) \leqslant 1, \quad \alpha_{n,m}(t) \in \{0,1\}
\end{aligned} \tag{18}$$

其中，K_{safe}为模型下发数量的安全阈值，$P_{\text{normal}}(K)$为车辆不受恶意模型影响的概率，$v_n(t)$为车辆n在t时刻的速度。该优化问题面临高维动态性、多约束耦合、不确定性强及长期性能权衡等挑战，使得传统静态优化或贪婪方法难以有效求解。为此，本文将其建模为马尔可夫决策过程，并引入PPO方法，通过剪切更新保证策略稳定性，结合熵正则增强探索能力，从而在动态车联网环境下实现低时延与安全性兼顾的模型传输迁移策略。

3 基于PPO的低时延模型传输迁移算法

3.1 马尔可夫决策过程建模

为了应对车联网场景下车辆高速移动导致的信道质量波动以及传输中断风险，本文提出了一种基

于PPO的低时延联邦学习模型传输迁移算法，将车辆与RSU之间的传输关联和迁移决策建模为一个马尔可夫决策过程[27]（MDP, Markov decision process），记为四元组 $\mathcal{M}=(\mathcal{S},\mathcal{A},\mathcal{P},\mathcal{R})$。

状态空间 $\mathcal{S}$：系统状态由所有车辆的二维坐标、行驶角度、与各RSU的欧氏距离、剩余待发数据量以及上一时刻的关联结果构成。对于单个车辆而言，其局部状态可表示为

$$s_n(t)=[x_n(t),y_n(t),\theta_t,\{d_{m,n}(t)\}_{m=1}^{M},D(t),a_n(t-1)] \tag{19}$$

全局状态为所有车辆的局部状态拼接而成。

动作空间 $\mathcal{A}$：在每一个离散时隙，系统需要为车辆选择目标RSU进行关联，本文采用联合动作的单一离散编码方式来表示车辆与RSU的关联策略。系统中用 M 个RSU和 N 辆车，则整个动作空间大小为 M^N。每一个动作都可以用一整数进行唯一编码，该整数对应所有车辆在当前时刻选择的RSU组合。通过进制转换，实现整数动作与车辆选择向量之间的相互映射，从而在实现上简化动作空间定义。

状态转移 $\mathcal{P}$：状态随着车辆运动和信道传播过程演化。车辆位置每步随机选择的曲率 k 做直线或曲线更新，信道增益依据3GPP标准建模，传输速率 $R_{m,n}(t)$ 的计算过程如式（1）所示，进而更新剩余数据量 $D(t+1)$（如式（11）所示）。

奖励函数 $\mathcal{R}$：奖励设计目标是最小化时延并保持迁移稳定性。本文在奖励中综合考虑高传输速率带来的正向收益、迁移行为的惩罚以及完成任务时的结果奖励，从而引导策略趋向低时延与高稳定性，奖励计算式如下

$$r(t)=\sum_{n=1}^{N}(R_{m,n}(t)\Delta-P_{\text{mig}}\cdot\mathbb{I}[a_n(t)\neq a_n(t-1)]+R_{\text{done}}\cdot\mathbb{I}[D_n(t+1)\leqslant 0]) \tag{20}$$

其中，P_{mig} 为迁移惩罚，$\mathbb{I}[\cdot]$ 为指示函数，R_{done} 为任务完成奖励。

3.2　PPO优化思想

在上述MDP建模的基础上，本文采用近端策略优化算法对迁移策略进行优化。PPO属于基于策略梯度的强化学习方法，其核心思想是通过限制新旧策略之间的差异来实现稳定更新。传统策略梯度方法在更新时容易出现过大的步长，导致策略崩塌；而PPO通过引入剪切目标函数，在保证充分探索的同时抑制过大的策略更新，从而兼顾训练的稳定性与收敛速度。算法流程如下所示。

算法：基于PPO的低时延联邦学习模型传输迁移算法

输入：训练回合数 E；学习率 $\alpha_{\text{actor}},\alpha_{\text{critic}}$；折扣因子 γ；剪切阈值 ε；每次更新迭代数 U；更新间隔步数 T_{update}；车辆数 N；RSU数 M；时隙 Δ；最大步数 $T_{\max}$。

输出：车辆-RSU关联/迁移策略 π_θ。

初始化策略参数 θ 与价值参数 ϕ；初始化轨迹缓存 $\mathcal{D}$；置旧策略 $\theta_{\text{old}}\leftarrow\theta$。

for episode e =1..E do

　环境复位得 s_0，设 $t\leftarrow 0$，done ← False

　While（not done）and $t<T_{\max}$ do

　　按旧策略 $\pi_{\theta_{\text{old}}}$ 采样整数动作 $c_t\in\{0,\cdots,M^N-1\}$

　　执行动作：随机曲率更新车辆位置/航向；按3GPP计算信道与速率；更新 $D(t+1)$；得到奖励 r_t 与新状态 s_{t+1}；

　　将 $(s_t,c_t,r_t,s_{t+1},\text{done})$、$\log\pi_{\theta_{\text{old}}}(c_t|s_t)$、$V_\phi(s_t)$ 存入 $\mathcal{D}$；$t\leftarrow t+1$。

　　if $|\mathcal{D}|\geqslant T_{\text{update}}$ or done then

　　计算优势 $\widehat{A}_t=\widehat{G}_t-V_\phi(s_t)$ 与目标回报 $\widehat{G}_t=\sum_{l=0}^{T-t}\gamma^l r_{t+l}$

　　for iter = 1..U do(从 $\mathcal{D}$ 中按batch_size采样小批量数据)

　　计算策略比率并构造PPO-clip目标；

　　计算价值损失与熵正则，合成为总损失；

　　更新 θ，ϕ（学习率 $\alpha_{\text{actor}},\alpha_{\text{critic}}$）。

　　end for；清空 $\mathcal{D}$，同步旧策略：$\theta_{\text{old}}\leftarrow\theta$。

　　end if

　end while

end for

4　仿真测试

4.1　仿真设置

实验环境与平台：本文实验在本地计算机平台上完成，硬件配置包括Intel Core i7-13620H CPU、NVIDIA GeForce RTX 4060 GPU。在此基础上，实验软件环境基于Python 3.12开发，并使用PyTorch 2.6.0动态迁移仿真平台。

参数设置：仿真设置4个RSU，并将其随机均

匀部署在600 m×600 m的区域内，车辆数量设置为15辆，由于联邦学习过程中，车辆可能处于不同的工作状态，如本地训练或数据传输，并不是所有车辆都会同时参与模型下发，可能只有一部分车辆会在每轮中处于模型下发过程。此外，设置车辆行驶曲率k={0，−0.02，0.02}，模拟3种典型行驶轨迹，反映车辆运动的不确定性。考虑仿真中车辆数量少于实际系统规模，为保持城市通信负载压力的真实性，设定车辆与RSU之间的带宽为2 MHz，若多辆车同时连接单一RSU则采用均等分配策略为各车辆分配带宽资源，以此模拟实际场景中有限无线资源下的带宽分配约束。仿真详细参数设置见表1。

表1　　仿真详细参数设置

参数	参数设置
车辆速度v	15 m/s
θ_0	0
D_{model}	20 M
P_m	0.2 W
σ_0	10^{-13} W
f_c	5.6 GHz
σ_0	4 dB
Δ	1 s
E	2 000轮
R_{done}	100
batch_size	128
剪切系数	0.2
学习率$\alpha_{\text{actor}}/\alpha_{\text{critic}}$	0.000 8/0.001
GAE系数λ	0.9
更新迭代数U	10
更新间隔步数T_{update}	2 048
最大步数$T_{\max}$	300
L	20
L_{mal}	5
a	2

对比方法：为了验证所提基于PPO的低时延联邦学习模型传输迁移方法的有效性，本文选取4种典型迁移策略作为对比基准，包括随机策略[28]、贪婪策略[29]、粒子群优化[30]（PSO, particle swarm optimization）策略以及DQN[31]策略。（1）随机策略在每一步中随机选择是否向目标RSU迁移；（2）贪婪策略始终选择当前距离最近的RSU进行迁移，追求最优的即时收益；（3）PSO策略通过粒子群协同搜索最优迁移策略，具备一定全局优化能力，其中种群数量设置为30、惯性权重为0.7、最大迭代次数100次、个体学习因子以及全局学习因子均为1.5；（4）DQN策略采用强化学习方法进行策略学习，在历史状态与奖励基础上优化长期收益，其中起始探索率为1、结束探索率为0.2、探索率衰减为8 000、经验池大小设置为10 000。上述策略的其他参数设置与本文方法一致，并在统一环境下进行对比评估，重点考察多车辆传输过程中的总传输时延，衡量各策略在迁移调度任务中的整体效率与实用性。

4.2　仿真结果及分析

不同参与下发车辆数量（比例）下单个车辆平均时延对比如图3所示。横坐标表示参与下发车辆数量以及占实验总车辆的比例，包括2辆（13.3%）、3辆（20%）、4辆（26.7%）和5辆（33.3%），纵坐标则为单个车辆平均时延。结果表明，随着下传比例的增加，系统可分配的带宽资源逐渐减少，导致时延呈上升趋势。然而，本文方法的单车平均时延始终保持低于DQN方法，显示出其在不同下传车辆比例下的较优表现。同时，本文方法在不同下传比例条件下均能实现较为平稳的收敛，进一步验证了其在车联网场景中的优势。综合考虑实验效率以及算法能否有效区分多辆车连接同一RSU时带宽分配的劣势，20%下发比例被选定为后续所有实验的标准设置。

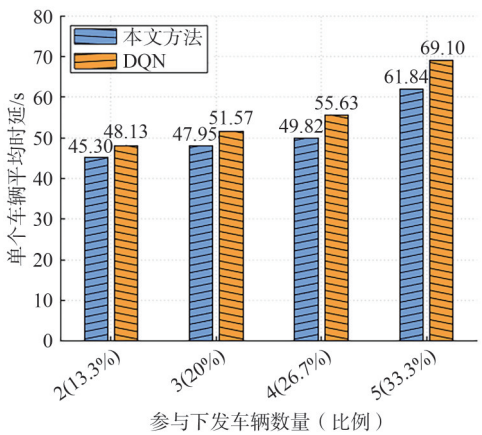

图3　不同参与下发车辆数量（比例）下单个车辆平均时延对比

本文方法与DQN总时延收敛对比曲线如图4所示，对比了本文方法与DQN在多车辆异步鲁棒联邦学习迁移任务中的总传输时延收敛过程。由于本文实验中设定的仿真环境相对理想化（如车辆数量较少，信号干扰因素受控），DQN策略在前期训

练阶段表现出更快的收敛速度，尤其在前200轮内时延快速下降，并于约400轮后趋于稳定。然而，随着训练轮数的增加，本文基于PPO的低时延联邦学习模型传输迁移方法在策略优化稳定性和长期累积奖励方面的优势逐步显现，表现出更优的收敛性能，约1 500轮后实现收敛。在保证模型最终收敛的前提下，本文方法所消耗的收敛时间显著低于基线方法DQN，不同策略下收敛时间与执行时间对比见表2，收敛时间降低了42.8%。此外，为了验证策略在车联网中的实际执行时间，本文对训练好的PPO和DQN模型进行了100次测试，并记录执行所用时间，最终求平均值。结果显示，DQN方法的平均每轮执行时间为0.705 4 s，而本文方法的平均每轮执行时间为0.048 8 s，执行效率显著优于DQN。表明本文方法具备更好的策略稳定性与泛化能力，能够在长期训练过程中获得更优的迁移策略，对多车辆异步传输场景具有更强的适应性和鲁棒性。

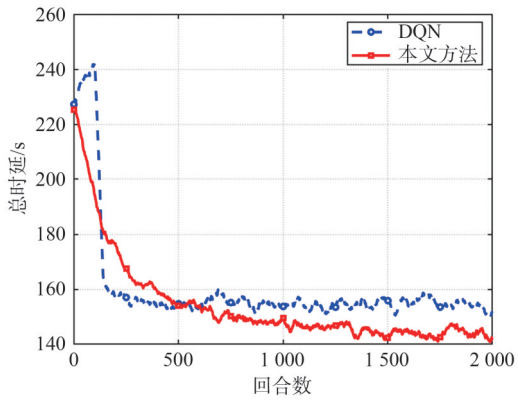

图4　本文方法与DQN总时延收敛对比曲线

表2　　不同策略下收敛时间与执行时间对比

策略	收敛时间/s	测试总执行时间/s	平均每轮执行时间/s
DQN	148.05	70.54	0.705 4
本文方法	**84.69**	**4.88**	**0.048 8**

不同折扣因子γ对本文方法平均时延的影响如图5所示，横轴为折扣因子，纵轴为对应平均总时延，每个点表示该折扣因子下多回合训练后的平均性能。从图5中可以观察到，随着折扣因子的增加，时延先下降后上升。在$\gamma=0.85$时，系统达到最优性能，对应平均总时延为143.84 s，优于其他所有配置。该结果表明，在该仿真环境下，过低的折扣因子（如0.7，0.75）可能导致策略短视，缺乏

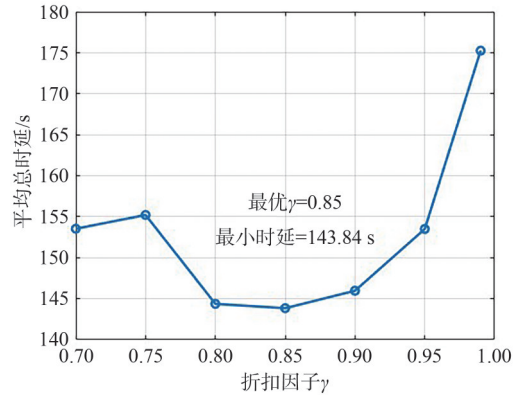

图5　不同折扣因子γ对本文方法平均总时延的影响

对未来收益的考虑；而过高的折扣因子（如0.95，0.99）则可能引入过多未来奖励影响，导致策略收敛速度变慢，甚至在有限训练回合下表现退化。

在本文实验中，测试了迁移惩罚系数P_{mig}（见式（20））对模型下发过程的影响，并分别设置了P_{mig}=0.3、P_{mig}=0.5、P_{mig}=1 3个取值进行实验，迁移惩罚系数对总时延的影响如图6所示。通过实验观察，迁移惩罚系数的适当变化对总时延以及收敛过程的影响几乎没有显著差异，验证了本文算法的鲁棒性，即算法能够适应迁移惩罚系数的变化。鉴于本文设定下迁移时延可以忽略，P_{mig}不对应物理时延，仅作为策略稳定性的正则化常数，在发生迁移时一次性扣减，用于抑制因短期速率波动导致的重复迁移试探。该系数的取值经过实验各部分奖励数值的权衡，最终设定为0.5。

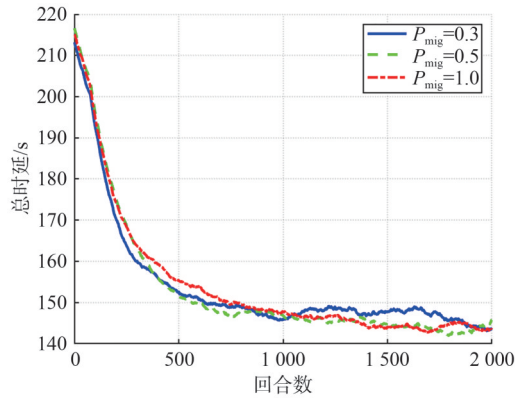

图6　迁移惩罚系数对总时延的影响

不同策略与本文方法的平均总时延对比如图7所示，横轴依次展示了5种策略：随机策略、贪婪策略、PSO、DQN以及本文方法，纵轴表示平均总时延。从图7中可以看出，贪婪策略与随机策略的平均时延最高，分别为235.3 s和226.6 s，说明在复

杂多辆车异步通信任务中，其策略表现出较差的迁移效率。PSO 相较于前两种方法显著降低了总时延，达到了 162.76 s，表现出较好的优化能力。DQN 进一步优化总时延至 154.72 s，但耗费时间过长，训练过程效率不高。本文基于 PPO 的低时延联邦学习模型传输迁移方法达到最优效果，平均总时延为 143.84 s，在保证鲁棒性的同时，实现了更低的时延开销。图 7 验证了本文提出的基于 PPO 的策略在异步场景中的收敛性、稳定性与鲁棒性优势。相比传统和强化学习策略，本文方法更适应动态变化的多车辆通信环境。

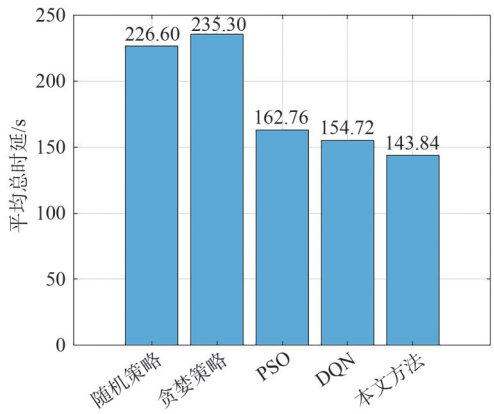

图 7　不同策略与本文方法的平均总时延对比

不同策略下 K 值变化的安全性与时延对比见表 3，展示了在模型数量 K 从 3 到 6 变化时，各策略对应的通信时延以及规避恶意模型的概率 P_{normal}。该概率由式（17）计算得出，衡量车辆经过两次选择后只选择正常模型的概率。本文结合实际车联网场景设定安全性约束阈值 $P_{\text{normal}}(K_{\text{safe}})$=0.99，用于保障模型传输的安全性与可靠性。

从表 3 可以看出，随着模型数量 K 的增加，所有策略的通信时延普遍上升，而 P_{normal} 也相应提高，体现出更小模型的易传输性。综合考虑通信时延与安全性要求，本文最终选取 K=5 为最优模型数量，此时在满足 $P_{\text{normal}}(K)\geqslant P_{\text{normal}}(K_{\text{safe}})$ 的前提下，本文方法相较于随机策略、贪婪策略、PSO、DQN 实现了更低的平均时延，展示出更优的鲁棒性与效率平衡能力。

表 3　不同策略下 K 值变化的安全性与时延对比

模型数量 K	P_{normal}	随机策略/s	贪婪策略/s	PSO/s	DQN/s	本文方法/s
3	0.859 6	135.99	156.74	95.7	93.98	**86.35**
4	0.967 8	180.75	200.44	133.40	125.13	**114.89**
5	0.995 1	226.60	235.30	162.76	154.72	**143.84**
6	0.999 6	273.16	252.55	193.96	185.26	**170.96**

5　结束语

本文针对动态车联网中 RSU 与车辆进行模型下发过程中存在的安全性与时延权衡问题，提出了一种低时延联邦学习模型传输迁移方法。在保证系统安全性满足约束阈值的前提下，通过强化学习优化 RSU 迁移策略，有效提升了传输效率与系统稳定性。实验表明，所提方法在多模型下发过程中展示出良好的安全性与时延平衡，相较于其他基线策略显著降低了通信时延。该研究为构建高效、安全的联邦学习通信系统提供了新思路，具有较强的应用价值与推广前景。

参考文献：

[1] 江恺, 曹越, 周欢, 等. 车联网边缘智能: 概念、架构、问题、实施和展望[J]. 物联网学报, 2023, 7(1): 37-48.
Jiang K, Cao Y, Zhou H, et al. Edge intelligence empowered Internet of vehicles: concept, framework, issues, implementation, and prospect[J]. Chinese Journal on Internet of Things, 2023, 7(1): 37-48.

[2] Alalwany E, Mahgoub I. Security and trust management in the Internet of vehicles (IoV): challenges and machine learning solutions[J]. Sensors, 2024, 24(2): 368.

[3] Wang X J, Zhu H L, Ning Z L, et al. Blockchain intelligence for Internet of vehicles: challenges and solutions[J]. IEEE Communications Surveys & Tutorials, 2023, 25(4): 2325-2355.

[4] Khezri E, Hassanzadeh H, Yahya R O, et al. Security challenges in Internet of vehicles (IoV) for ITS: a survey[J]. Tsinghua Science and Technology, 2025, 30(4): 1700-1723.

[5] 胡海峰, 张熙, 赵海涛, 等. 移动边缘计算中通信高效的联邦学习模型剪枝算法[J]. 物联网学报, 2024, 8(3): 112-126.
Hu H F, Zhang X, Zhao H T, et al. Communication-efficient model pruning for federated learning in mobile edge computing[J]. Chinese Journal on Internet of Things, 2024, 8(3): 112-126.

[6] 李佳恒, 吴钦木. 基于三元联邦学习的车联网数据协同学习与通信优化研究[J]. 现代电子技术, 2024, 47(15): 26-33.
Li J H, Wu Q M. Research on Internet of vehicles data cooperative learning and communication optimization based on tripartite federated learning[J]. Modern Electronics Technique, 2024, 47(15): 26-33.

[7] Liang F Y, Yang Q L, Liu R Q, et al. Semi-synchronous federated learning protocol with dynamic aggregation in Internet of vehicles[J]. IEEE Transactions on Vehicular Technology, 2022, 71(5): 4677-4691.

[8] Sharma I, Sharma A, Gupta S K. Asynchronous and synchronous

federated learning-based UAVs[C]//Proceedings of the 2023 Third International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP). Piscataway: IEEE Press, 2023: 105-109.

[9] Stripelis D, Thompson P M, Ambite J L. Semi-synchronous federated learning for energy-efficient training and accelerated convergence in cross-silo settings[J]. ACM Transactions on Intelligent Systems and Technology, 2022, 13(5): 1-29.

[10] Xu C H, Qu Y Y, Xiang Y, et al. Asynchronous federated learning on heterogeneous devices: a survey[J]. Computer Science Review, 2023, 50: 100595.

[11] Wang Z Y, Zhang Z Y, Wang J. Asynchronous federated learning over wireless communication networks[C]//Proceedings of the ICC 2021-IEEE International Conference on Communications. Piscataway: IEEE Press, 2021: 1-7.

[12] Wang Z Y, Xu H L, Liu J C, et al. Resource-efficient federated learning with hierarchical aggregation in edge computing[C]//Proceedings of the IEEE INFOCOM 2021 IEEE Conference on Computer Communications. Piscataway: IEEE Press, 2021: 1-10.

[13] Nguyen T D, Rieger P, Chen H, et al. FLAME: taming backdoors in federated learning[C]//Proceedings of USENIX Security Symposium. Berkeley: USENIX Association, 2022.

[14] Zhang X R, Chang Z, Hu T, et al. Vehicle selection and resource allocation for federated learning-assisted vehicular network[J]. IEEE Transactions on Mobile Computing, 2024, 23(5): 3817-3829.

[15] 廖岑卉珊, 陈俊彦, 梁观平, 等. 基于深度强化学习的SDN服务质量智能优化算法[J]. 物联网学报, 2023, 7(1): 73-82.
Liao C H S, Chen J Y, Liang G P, et al. Quality of service optimization algorithm based on deep reinforcement learning in software defined network[J]. Chinese Journal on Internet of Things, 2023, 7(1): 73-82.

[16] Jiang K, Cao Y, Song Y J, et al. Asynchronous federated and reinforcement learning for mobility-aware edge caching in IoV[J]. IEEE Internet of Things Journal, 2024, 11(9): 15334-15347.

[17] Wang D, Song B, Lin P, et al. Resource management for edge intelligence (EI)-assisted IoV using quantum-inspired reinforcement learning[J]. IEEE Internet of Things Journal, 2022, 9(14): 12588-12600.

[18] 赵晓焱, 韩威, 张俊娜, 等. 基于异步深度强化学习的车联网协作卸载策略[J]. 计算机应用, 2024, 44(5): 1501-1510.
Zhao X Y, Han W, Zhang J N, et al. Collaborative offloading strategy in Internet of vehicles based on asynchronous deep reinforcement learning[J]. Journal of Computer Applications, 2024, 44(5): 1501-1510.

[19] 刘冰艺, 刘煜昊, 韩玮祯, 等. 边缘智能下基于强化学习的车联网路由协议[J]. 通信学报, 2023, 44(11): 110-119.
Liu B Y, Liu Y H, Han W Z, et al. Edge intelligence-assisted routing protocol for Internet of vehicles via reinforcement learning[J]. Journal on Communications, 2023, 44(11): 110-119.

[20] 王为念, 苏健, 陈勇, 等. 基于多智能体深度强化学习的车联网频谱共享[J]. 电子学报, 2024, 52(5): 1690-1699.
Wang W N, Su J, Chen Y, et al. Multi-agent reinforcement learning enabled spectrum sharing for vehicular networks[J]. Acta Electronica Sinica, 2024, 52(5): 1690-1699.

[21] Yao L, Xu X L, Bilal M, et al. Dynamic edge computation offloading for Internet of vehicles with deep reinforcement learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(11): 12991-12999.

[22] Tang H J, Wu H M, Qu G J, et al. Double deep Q-network based dynamic framing offloading in vehicular edge computing[J]. IEEE Transactions on Network Science and Engineering, 2023, 10(3): 1297-1310.

[23] Li H F, Chen C, Shan H G, et al. Deep deterministic policy gradient-based algorithm for computation offloading in IoV[J]. IEEE Transactions on Intelligent Transportation Systems, 2024.

[24] Zhang H J, Jiang M H, Liu X N, et al. PPO-based PDACB traffic control scheme for massive IoV communications[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(1): 1116-1125.

[25] ETSI. 5G; Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 18.0.0 Release 18)[R]. 2024.

[26] 尹宏博, 王帅, 张科, 等. 车辆算力网络中异步鲁棒联邦学习方法研究[J]. 物联网学报, 2024, 8(4): 14-22.
Yin H B, Wang S, Zhang K, et al. Research on asynchronous robust federated learning method in vehicle computing power network[J]. Chinese Journal on Internet of Things, 2024, 8(4): 14-22.

[27] Steimle L N, Kaufman D L, Denton B T. Multi-model Markov decision processes[J]. IISE Transactions, 2021, 53(10): 1124-1139.

[28] Hazarika B, Singh K, Biswas S, et al. DRL-based resource allocation for computation offloading in IoV networks[J]. IEEE Transactions on Industrial Informatics, 2022, 18(11): 8027-8038.

[29] Attia R, Hassaan A, Rizk R. Advanced greedy hybrid bio-inspired routing protocol to improve IoV[J]. IEEE Access, 2021, 9: 131260-131272.

[30] Wang Y, Hu F J, Xu H C, et al. A multigroups cooperative particle swarm algorithm for optimization of multivehicle path planning in Internet of vehicles[J]. IEEE Internet of Things Journal, 2024, 11(22): 35839-35851.

[31] Yang C Y, Xu X L, Zhou X K, et al. Deep Q network-driven task offloading for efficient multimedia data analysis in edge computing-assisted IoV[J]. 2022, 18(2s): 1-24.

[作者简介]

王帅（2001– ），男，电子科技大学信息与通信工程学院硕士生，主要研究方向为算力网络、边缘智能。

尹宏博(1998–)，男，电子科技大学信息与通信工程学院博士生，主要研究方向为算力网络、联邦学习、边缘计算。

江池(1997–)，女，电子科技大学信息与通信工程学院博士生，主要研究方向为区块链、网络安全、生成模型。

张科(1978–)，男，博士，电子科技大学信息与通信工程学院副教授，主要研究方向为边缘智能网络、智慧车联网、边缘计算。

张引(1986–)，男，博士，电子科技大学（深圳）高等研究院研究员，主要研究方向为移动计算、算力网络、边缘智能。