

# 基于区块链的无人机辅助铁路通信系统低能耗传输方法

代海波<sup>1</sup>, 吴天奇<sup>1</sup>, 梁轶群<sup>2</sup>, 张哲<sup>2</sup>, 李春国<sup>3</sup>

(1. 南京邮电大学物联网学院, 南京 210003; 2. 中国铁道科学研究院集团有限公司通信信号研究所, 北京 100081;  
3. 东南大学信息科学与工程学院, 南京 210096)

**摘要:** 无人机基站具有快速部署、机动覆盖等特性, 使其成为铁路应急通信的一种有效解决方案。然而, 无人机组成的低空通信网络既有储能有限的约束又存在数据被窃听或篡改的风险。将区块链技术引入到无人机辅助的铁路无线通信系统中以保障数据安全。在传输时延和数据队列稳定性约束下, 提出了一种最小化通信系统能耗和区块链系统时延的联合优化问题。为了求解这个非凸混合整数且时变的随机优化问题, 提出基于Lyapunov的drift-plus-penalty方法将长期的随机优化问题转化为多个时隙的子问题, 并设计基于D3QN-TD3的深度强化学习算法求解, 得到每个时隙内的最优关联策略和功率控制。实验结果表明, 所提方法在减少能耗和时延方面取得了显著的效果。

**关键词:** 无人机通信网络; 区块链; 关联策略; 深度强化学习; Lyapunov优化

**中图分类号:** U285.21 **文献标志码:** A

## Low-Energy Transmission Method for Blockchain-Based UAV-Assisted Railway Communication System

DAI Haibo<sup>1</sup>, WU Tianqi<sup>1</sup>, LIANG Yiqun<sup>2</sup>, ZHANG Zhe<sup>2</sup>, LI Chunguo<sup>3</sup>

(1. School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; 2. Signal and Communication Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China; 3. School of Information Science and Engineering, Southeast University, Nanjing 210096, China)

**Abstract:** Unmanned aerial vehicle (UAV)-mounted base stations possess characteristics, such as rapid deployment and flexible coverage, making them an effective solution for emergency communication in railways. However, the low-altitude communication network formed by UAVs faces the constraint of the limited energy storage and the risk of data eavesdropping or tampering. This paper introduces blockchain technology into a UAV-assisted railway wireless communication system to ensure data security. Considering the constraints on transmission delay and data queue stability, this paper proposes a joint optimization problem aimed at minimizing the energy consumption of the UAV-assisted communication system and the latency of the blockchain. To solve this non-convex, mixed-integer, and time-varying stochastic optimization problem, a Lyapunov-based drift-plus-penalty method is proposed to transform the long-term stochastic optimization problem into sub-problems of multiple time slots. Deep reinforcement learning based on D3QN-TD3 is designed to solve these sub-problems, and then the optimal association strategies and power control for each time slot are obtained. Experimental results demonstrate the

significant effectiveness of the proposed method in reducing the energy consumption and delay.

**Key words:** UAV-enabled communication network; blockchain; association strategy; deep reinforcement learning; Lyapunov optimization

## 引 言

当高铁运输系统受到自然灾害影响或者出现故障,特别是路边基站等通信基础设施遭到破坏时,亟需保障高铁数据的安全有效传输<sup>[1]</sup>。然而,现有的方法受环境和空间的限制,缺乏灵活性和及时性<sup>[2-3]</sup>。随着技术进步,无人机在无线通信方面的应用愈发广泛,特别是在提供临时通信支持方面的能力。由于其部署便捷、成本效益高以及高度的灵活性特点,无人机能够在没有地面基站的恶劣或危险条件下发挥作用,来执行一些无法完成的任务<sup>[4-5]</sup>。例如,在自然灾害发生或地形受限的情况下,无人机能迅速被派遣至影响区域,为地面通信设备提供必要的无线网络覆盖<sup>[6-7]</sup>,充当临时的通信节点,从而保障通信链路的连续性和稳定性。因此,研究无人机在辅助铁路无线通信系统中的传输策略不仅具有深远的理论意义,也有着显著的实践价值。

确保铁路通信数据,尤其是列车控制数据的传输及时性至关重要,其中,延迟成为评估系统性能的一个核心指标。目前,大多数研究通过计算数据传输大小与速率的比例来估算物理层的传输延时。例如,文献[8]在车辆网络环境下,探讨了如何最小化传输延时的问题,然而,这种度量时延方法中传输速率是时间平均的,甚至假设信道是慢衰落、平稳的,这种时延定义不贴合实际需求,特别是在铁路通信场景中。另一方面,文献[9]通过引入排队理论,提出了一种基于马尔可夫决策过程的数据排队时延模型,并利用 Little 定律将其转化为队列长度的问题,从而在车辆网络系统中探讨了如何最大化传输容量问题。进一步地,文献[10]采用随机网络演算法,将连续到达的数据和传输数据从比特域转换到信噪比域,推导出在低延迟高可靠通信中的延迟违反概率的上限。而文献[11]则基于等效带宽和容量理论,将排队时延的限制转换为最小传输速率的限制,并在低延迟高可靠的无线接入网络中研究了功率控制与带宽分配的联合优化问题。

同时,无人机对能耗敏感,能量的高效利用有利于提升无人机基站的服务时间和质量。在文献[12]中,通过联合优化无人机基站的位置部署和信道分配,最大化整个网络的能效性能。文献[13]研究了基于无人机多播系统中任务完成时间最小化和无人机能量最小化问题,设计了最优的无人机飞行轨迹和天线波束宽度。在文献[14]中,利用部署在车顶的多个接收点克服列车车身外壳严重的穿透损失问题,同时部署无人机作为中继,通过调度不同信道质量的通信链路以提高系统的吞吐量。文献[15]提出了一种深度强化学习算法来对多用户的无线蜂窝网络进行功率分配,并证明了基于数据驱动的学习方法优于传统的基于模型的功率分配方法。

利用无人机基站辅助数据传输面临着严峻的安全性问题。无人机网络中容易存在不信任节点的窃听、篡改和恶意攻击的威胁,可能导致敏感信息泄露或通信系统被未授权方接管。文献[16]使用区块链技术确保数据的可靠传输和交互,提出了基于计算和缓存资源优化方案,通过协同考虑任务卸载决策、缓存决策、卸载共识节点数量、区块间隔和区块大小,优化能源消耗和计算开销。文献[17]通过联合优化用户关联、数据速率分配、区块生产者调度和计算资源分配,保证安全的同时减少了计算复杂度。提出的算法展示了良好的收敛性,并能在最小化能源消耗和区块链系统计算延迟之间实现最佳折中。文献[18]探讨了将区块链技术融入移动边缘计算的创新应用,开发出了一种稳健的计费数据传输策略。该策略通过协调优化主节点与副本节点的数据卸载行为、调整区块大小及其生成时间间隔,致力于应对系统状态的动态变化和持续性问题,从而显著提升了网络的整体效能。

铁路应急通信场景下,通信网络的高动态性以及高速导致通信链路稳定性变差,处理不可预测的传输堵塞需满足严格的时延。同时,无人机的部署所导致能耗敏感、数据安全性问题不容忽视,引入区块链技术保证数据安全是值得研究的问题,现有研究未考虑网络的动态性,且区块链系统和数据传输系统的设计与优化相互独立,这样会导致网络资源的不均衡分配,增大时延,降低系统性能<sup>[19-20]</sup>。为了实现铁路应急通信数据的低时延安全传输,本文将部署有计算设备的多个无人机基站组成空中通信网络为高铁数据提供可靠的传输服务,同时,提出区块链技术实现数据的安全传输与处理。通过联合优化关联策略和功率控制,在保证传输时延约束前提下,拟实现通信系统能耗最小化和区块链系统时延最小化之间的最优折中。

## 1 系统模型

如图1所示,考虑在一个受到自然灾害或紧急情况下,某些高铁路边基站陷入瘫痪无法使用时,部署无人机辅助的高铁数据传输的空地网络,高铁顶部部署 $K$ 个车顶接入点(Vehicular station, VS),空中部署 $N_u$ 个无人机基站形成无线接入网络,所有基站与接入点都部署单天线,无人机的回程链路通过地面接收终端利用光纤网络接入数据处理中心。每个无人机基站安装有边缘计算设备,对接收到VS发射的数据预处理后通过地面接收端传输给数据处理中心。接入点集合表示为 $\mathcal{K}=\{1, 2, \dots, K\}$ 和无人机基站集合表示为 $\mathcal{N}_u=\{1, 2, \dots, N_u\}$ 。本文考虑 $T$ 个时隙内的数据传输问题,每个时隙的长度为 $\tau$ ,并设置为充分小的,以确保收发端间的信道状态信息和位置是近似未变的,时隙集合为 $\mathcal{T}=\{1, 2, \dots, T\}$ 。在每个时隙 $t \in \mathcal{T}$ 中,假设VS只接入一个无人机基站且可接入的基站集合为 $\mathcal{N}_u$ ,VS  $k \in \mathcal{K}$ 的3D坐标位置为 $L_k(t)$ ,无人机 $u \in \mathcal{N}_u$ 的3D坐标位置为 $L_u(t)$ 。本文不考虑无人机的轨迹设计问题,为了简化处理,在 $T$ 个时隙内,假设所有无人机飞行的方向和速度与高铁的运行一致,无人机之间以及与高铁VS间形成相对静态的网络。

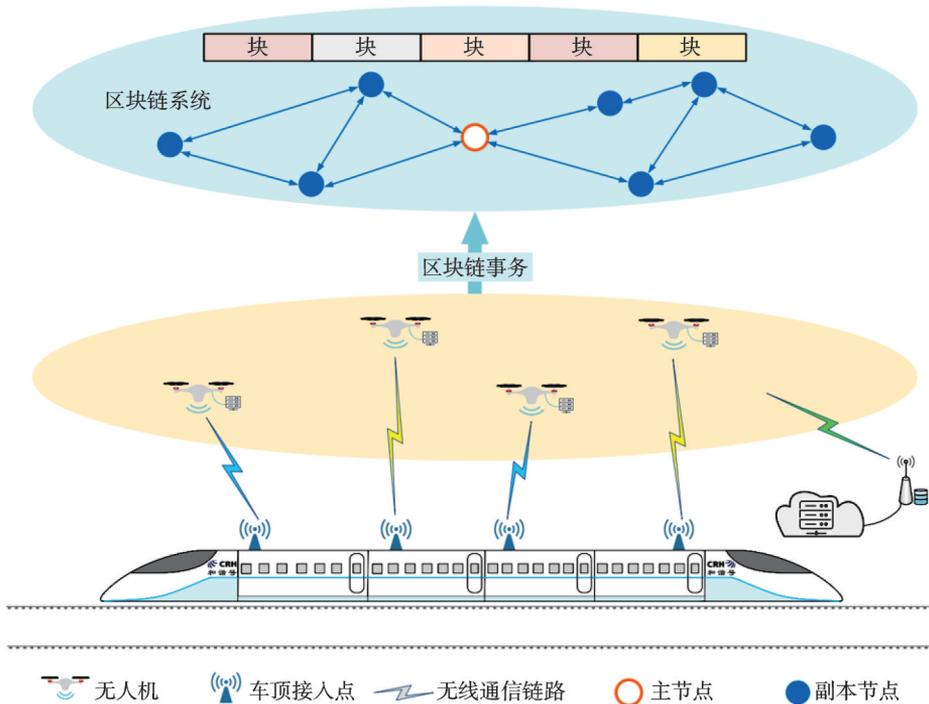


图1 系统模型

Fig.1 System model

### 1.1 传输模型

无人机与地面终端之间的通信模型中,在时隙  $t$  从终端  $k$  到无人机的信道增益服从自由空间路径损耗模型,即

$$h_{ku}(t) = \frac{\mu_0}{\|L_u(t) - L_k(t)\|^2} \quad (1)$$

式中:  $\mu_0 = \left(\frac{4\pi f_c}{c}\right)^{-2}$  表示 1 m 距离处的信道增益,  $f_c$  为载波频率,  $c$  表示光速。

由于高铁的高移性,多普勒效应引起的载波间干扰被认为是白噪声的一部分,多普勒干扰因子  $P_{b-ICI}$  表示为

$$P_{b-ICI} = 1 - \int_{-1}^1 (1 - |\xi|) J_0(2\pi f_d T_s \xi) d\xi \quad (2)$$

式中:  $J_0(\cdot)$  表示第一类 0 阶贝塞尔函数,  $f_d$  为最大的多普勒频率,  $T_s$  为符号时长。

假设无人机与 VS 间采用相同的频谱且存在干扰,带宽为  $B_u$ 。因此,在时隙  $t$  中,高铁端 VS  $k$  关联无人机基站  $u$  的前提下基站  $u$  接收到的 SINR 表示为

$$r_{ku}(t) = \log_2 \left( 1 + \frac{p_{ku}(t) h_{ku}(t)}{\sum_{u' \in \mathcal{N}_u} \sum_{k' \in \mathcal{K} \setminus \{k\}} p_{k'u}(t) h_{k'u}(t) x_{k'u}(t) + P_{u-ICI}(t) h_{ku}(t) + \sigma_{ku}^2(t)} \right) \quad (3)$$

式中:  $p_{ku}(t)$  表示 VS  $k$  到无人机基站  $u$  的发射功率;  $p_{k'u}(t)$  表示其他 VS  $k'$  到无人机基站  $u'$  的发射功率;  $h_{ku}(t)$  和  $h_{k'u}(t)$  分别表示 VS  $k$  和 VS  $k'$  到无人机基站  $u$  的信道增益; 指示函数  $x_{k'u}(t)$  表示 VS  $k'$  的关联策略,若 VS  $k'$  传输数据给无人机基站  $u'$ ,则  $x_{k'u}(t) = 1$ , 否则,  $x_{k'u}(t) = 0$ 。

无人机  $u$  地面接收端收到的 SINR 表示为

$$r_u^{\text{back}}(t) = \log_2 \left( 1 + \frac{p_u^{\text{back}}(t) h_u^{\text{back}}(t)}{P_{u-ICI}^{\text{back}}(t) h_u^{\text{back}}(t) + \sigma_{u-\text{back}}^2(t)} \right) \quad (4)$$

式中:  $p_u^{\text{back}}(t)$  表示无人机回程发射端的发射功率,  $h_u^{\text{back}}(t)$  表示无人机基站  $u$  到地面发射端的信道增益。

因此, VS  $k$  端数据传输时间为

$$t_k^{\text{tr}}(t) = \frac{D_k(t)}{r_{ku}(t)} \quad (5)$$

式中:  $D_k(t)$  为传输的数据大小。

VS  $k$  端数据传输所消耗的能量为

$$E_k^{\text{tr}}(t) = p_{ku}(t) t_k^{\text{tr}}(t) + E^{\text{fix}} + E^{\text{back}} \quad (6)$$

式中:  $E^{\text{fix}}$  表示天线电路的消耗,  $E^{\text{back}}$  表示回程链路的消耗。

在无人机辅助的空地传输网络架构中,无人机基站端维护用于存储网络层数据的内部队列集合。在时隙  $t$  内,假设  $a_k(t)$  表示在 VS  $k$  端更新包的数据大小,且平均到达率  $\mathbb{E}[a_k(t)] = \lambda_k$ ,  $Q_k(t)$  表示 VS  $k$  的当前队列缓冲,  $b_k(t)$  表示无人机基站收到 VS  $k$  传输的数据量,且有

$$b_k(t) = B_u \sum_{u \in \mathcal{N}_u} r_{ku}(t) x_{ku}(t) \quad (7)$$

未来系统状态是由随机到达和数据传输能力决定的,因此队列  $Q_k(t)$  的动态方程为

$$Q_k(t+1) = \max\{Q_k(t) - b_k(t) + a_k(t), 0\} \quad (8)$$

在时隙  $t$  内, VS  $k$  的包传输后,  $E_k(t)$  定义为处理收到包后的队列缓冲, 接收到的数据处理是基于 CPU 计算的能力, 因此, 数据积压的动态方程为

$$E_k(t+1) = \max\{E_k(t) - c_k(t) + b_k(t), 0\} \quad (9)$$

式中:  $c_k(t) = \sum_{u \in \mathcal{N}_u} \frac{x_{ku}(t) f_u(t) \tau}{\sum_{k' \in \mathcal{K}} x_{k'u}(t) \rho_c}$ ,  $f_u(t)$  表示无人机  $u$  的 CPU 周期频率,  $\rho_c$  表示处理 1 比特数据需要的 CPU 周期数。

## 1.2 计算模型

在无人机基站端, 接收到 VS 传输的数据后需要预处理, 令  $f_{ku}(t)$  表示无人机  $u$  处理 VS  $k$  传输数据使用的 CPU 周期频率,  $\rho_c$  表示处理 1 比特数据需要的 CPU 周期数, 那么, 处理 VS  $k$  传输数据的时间为

$$t_k^c(t) = \sum_{u \in \mathcal{N}_u} \frac{D_k(t) \rho_c x_{ku}(t)}{f_{ku}(t)} \quad (10)$$

处理数据所消耗的能量为

$$E_k^c(t) = \sum_{u \in \mathcal{N}_u} D_k(t) \rho_c x_{ku}(t) \bar{\omega}_u (f_{ku}(t))^2 \quad (11)$$

式中:  $\bar{\omega}_u$  表示与芯片架构有关的有效电容系数。

因此, 数据传输到处理完成所需要的时间和消耗的能量分别为

$$\begin{cases} T_k^{\text{tc}}(t) = t_k^{\text{tr}}(t) + t_k^c(t) \\ E_k^{\text{tc}}(t) = E_k^{\text{tr}}(t) + E_k^c(t) \end{cases} \quad (12)$$

## 1.3 区块链模型

本文利用 VSs 和无人机基站构建区块链系统, VS 充当普通节点, 主节点和共识节点由无人机基站充当, 且负责产生区块和参与共识过程。采用共识算法实用拜占庭容错算法 (Practical Byzantine fault tolerance, PBFT), 其可以容忍的恶意节点数量为  $\frac{N_u - 1}{3}$ 。假设  $\theta$  和  $\vartheta$  分别为产生或验证一个签名和一个消息认证码 (Message authentication code, MAC) 所需要的 CPU 周期<sup>[21]</sup>。算法流程如下:

(1) Request 阶段。在时隙  $t$  内, VS 发送传输请求到无人机基站, 区块链系统中每个无人机基站广播每个 VS 发送的事务, 即传输的数据记录。区块链系统随机分配一个主节点, 用于打包事务成一个新的区块。当主节点接收到事务时, 需要验证事务的签名和 MAC, 相应的计算开销为

$$c_1^{\text{p}}(t) = \frac{\kappa(t)}{\varrho(t)} (\theta + \vartheta) \quad (13)$$

式中:  $\kappa(t)$  为时隙  $t$  内事务的数量,  $\varrho(t)$  为正确事务的比例。

(2) Pre-prepare 阶段。主节点验证事务之后, 产生一个签名和  $N_u - 1$  个 MAC, 与产生的新区块一起发送给除主节点外的共识节点。收到的共识节点先验证区块的签名和 MAC, 再验证事务的签名和 MAC。因此, 主节点和共识节点的计算开销分别为

$$\begin{cases} c_2^{\text{p}}(t) = \theta + (N_u - 1) \vartheta \\ c_2^{\text{s}}(t) = (1 + \kappa(t)) (\theta + \vartheta) \end{cases} \quad (14)$$

(3) Prepare 阶段。验证后的共识节点产生一个签名和  $N_u - 1$  个 MAC, 并发送给其他节点, 这样, 每个节点 (包含主节点) 需要验证来自其他共识节点的  $\frac{2(N_u - 1)}{3}$  个签名和 MAC。因此, 主节点和共识节点的计算开销分别为

$$\begin{cases} c_3^p(t) = \frac{2(N_u - 1)(\theta + \vartheta)}{3} \\ c_3^r(t) = \theta + (N_u - 1)\vartheta + \frac{2(N_u - 1)(\theta + \vartheta)}{3} \end{cases} \quad (15)$$

(4)Commit阶段。如果接收正确信息的验证节点超过 $\frac{2(N_u - 1)}{3}$ 个,那么发送一个签名和 $N_u - 1$ 个MAC给所有其他的节点。然后,每个节点要验证 $\frac{2(N_u - 1)}{3}$ 个签名和MAC。因此,主节点和共识节点的计算开销都为

$$c_4(t) = \theta + (N_u - 1)\vartheta + \frac{2(N_u - 1)(\theta + \vartheta)}{3} \quad (16)$$

(5)Reply阶段。收集到 $\frac{2(N_u - 1)}{3}$ 个确认信息后,表明关于该事务达成共识,区块加入到区块链中。同时,每个共识节点产生一个签名和一个MAC,并发送给主节点。主节点需要验证 $\frac{2(N_u - 1)}{3}$ 个签名和MAC。因此,主节点和共识节点的计算开销分别为

$$\begin{cases} c_5^p(t) = \frac{2(N_u - 1)(\theta + \vartheta)}{3} \\ c_5^r(t) = \kappa(t)(\theta + \vartheta) \end{cases} \quad (17)$$

综上所述,在所有节点都本地计算的前提下,一个共识过程所需要总的计算开销为

$$C(t) = \left(4 + 2\kappa(t) + 2(N_u - 1) + \frac{\kappa(t)}{\varrho(t)}\right)\theta + \left(1 + 2\kappa(t) + 5(N_u - 1) + \frac{\kappa(t)}{\varrho(t)}\right)\vartheta \quad (18)$$

主节点和共识节点的计算开销分别为

$$\begin{cases} C^p(t) = \left(2 + 2(N_u - 1) + \frac{\kappa(t)}{\varrho(t)}\right)\theta + \left(4(N_u - 1) + \frac{\kappa(t)}{\varrho(t)}\right)\vartheta \\ C^r(t) = \left(3 + 2\kappa(t) + \frac{4(N_u - 1)}{3}\right)\theta + \left(1 + 2\kappa(t) + \frac{10(N_u - 1)}{3}\right)\vartheta \end{cases} \quad (19)$$

假设 $f_u^{bc}(t)$ 表示节点 $u$ 的CPU周期频率, $k$ 为系数,那么区块链系统的时延和能耗分别为

$$\begin{cases} T^{bc}(t) = \max_{u \in \mathcal{N}_s} \left\{ \frac{C(t)}{f_u^{bc}(t)} \right\} + T^b(t) + T^i(t) \\ E^{bc}(t) = k(f_u^{bc}(t))^2 C^p(t) + \sum_{u' \in \mathcal{N}_s \setminus \{u\}} k(f_{u'}^{bc}(t))^2 C^r(t) \end{cases} \quad (20)$$

式中: $T^b(t)$ 表示共识过程中节点间消息广播的时延, $T^i(t)$ 表示区块间隔,即区块产生者产生一个新的区块所需的平均时间。

## 2 问题设计与求解

### 2.1 优化问题设计

本文的目标是通过优化VS关联策略和发射功率,最小化数据传输系统的能耗和区块链系统的时延,同时满足区块链系统能耗、传输时延、CPU周期频率和队列稳定性的约束。优化问题如下

$$\min_{\{X,P\}} \mathbb{E} \left[ \sum_{t=1}^T \left( \omega_1 \sum_{k \in \mathcal{K}} E_k^{ic}(t) + (1 - \omega_1) \omega_2 T^{bc}(t) \right) \right] \quad (21a)$$

$$\text{s.t. } T_k^{ic}(t) \leq T_k^{\max} \quad t \in \mathcal{T}, k \in \mathcal{K} \quad (21b)$$

$$E^{bc}(t) \leq E^{\max} \quad t \in \mathcal{T} \quad (21c)$$

$$\sum_{k \in \mathcal{K}} x_{ku}(t) f_{ku}(t) + f_u^{bc}(t) \leq F^{\max} \quad t \in \mathcal{T} \quad (21d)$$

$$x_{ku}(t) \in \{0, 1\} \quad u \in \mathcal{N}_u, k \in \mathcal{K} \quad (21e)$$

$$\sum_{u \in \mathcal{N}_u} x_{ku}(t) = 1 \quad t \in \mathcal{T}, k \in \mathcal{K} \quad (21f)$$

$$0 \leq p_{ku}(t) \leq p_k^{\max} \quad k \in \mathcal{K}, t \in \mathcal{T} \quad (21g)$$

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}[Q_k(t)]}{t} = 0 \quad t \in \mathcal{T}, k \in \mathcal{K} \quad (21h)$$

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}[E_k(t)]}{t} = 0 \quad t \in \mathcal{T}, k \in \mathcal{K} \quad (21i)$$

式中:  $\omega_1 \in [0, 1]$  表示权重因子用于归一化目标函数,  $\omega_2$  表示映射因子用于确保两类目标的量级处于同一标准。约束式(21b)保证了数据传输处理时延低于最大容忍时延, 约束式(21c)为区块链的能耗约束, 约束式(21d)保证了相关VS数据处理的计算能力与区块链的计算容量之和不大于无人机的处理能力, 约束式(21e~21g)为关于关联策略的约束和传输功率约束, 式(21h~21i)为队列  $Q_k(t)$  和队列  $E_k(t)$  的稳定性约束。

## 2.2 基于 Lyapunov 理论的动态控制算法

首先, 为了确保队列  $Q_k(t)$  和队列  $E_k(t)$  是稳定的, 定义一个二次 Lyapunov 函数如下

$$\mathcal{L}(\Theta(t)) = \frac{1}{2} \sum_{k \in \mathcal{K}} (Q_k^2(t) + E_k^2(t)) \quad (22)$$

式中  $\Theta(t) = [Q_1(t), Q_2(t), \dots, Q_K(t), E_1(t), E_2(t), \dots, E_K(t)]^T$ 。

基于此, Lyapunov 漂移表示为

$$\Delta \mathcal{L}(\Theta(t)) = \mathcal{L}(\Theta(t+1)) - \mathcal{L}(\Theta(t)) \quad (23)$$

引入两个不等式定理, 对于任意的  $\alpha \geq 0, \beta \geq 0, \gamma \geq 0$ , 有不等式  $(\max(\alpha - \beta, 0) + \gamma)^2 - \alpha^2 \leq \beta^2 + \gamma^2 + 2\alpha(\gamma - \beta)$  成立和不等式  $(\max(\alpha - \beta, 0))^2 \leq (\alpha - \beta)^2$  成立。

根据队列的动态方程(8,9)以及上述不等式定理, 可知

$$\begin{cases} Q_k^2(t+1) - Q_k^2(t) \leq 2Q_k(t)(a_k(t) - b_k(t)) + (a_k(t) - b_k(t))^2 \\ E_k^2(t+1) - E_k^2(t) \leq 2E_k(t)(b_k(t) - c_k(t)) + (b_k(t) - c_k(t))^2 \end{cases} \quad (24)$$

简单化简可得

$$\begin{aligned} \Delta \mathcal{L}(\Theta(t)) = \mathcal{L}(\Theta(t+1)) - \mathcal{L}(\Theta(t)) &= \frac{1}{2} \left( \sum_{k \in \mathcal{K}} (Q_k^2(t+1) + E_k^2(t+1)) - \sum_{k \in \mathcal{K}} (Q_k^2(t) + E_k^2(t)) \right) \leq \\ & B(t) + \sum_{k \in \mathcal{K}} Q_k(t)(a_k(t) - b_k(t)) + \sum_{k \in \mathcal{K}} E_k(t)(b_k(t) - c_k(t)) \end{aligned} \quad (25)$$

式中  $B(t) = \frac{1}{2} \sum_{k \in \mathcal{K}} ((a_k(t) - b_k(t))^2 + (b_k(t) - c_k(t))^2)$ 。

假设每个队列的到达和服务这两项是有界的,因此存在一个有限常数  $\mathbf{r} > 0$ ,使得  $\mathbf{r} \geq B(t)$ ,并对式(25)两边取条件期望,得到一个上界,即

$$\begin{aligned} \mathbb{E}[\mathcal{L}(\Theta(t+1)) - \mathcal{L}(\Theta(t)) | \Theta(t)] = & \\ \mathbb{E} \left[ \left( \frac{1}{2} \sum_{k \in \mathcal{K}} (Q_k^2(t+1) + E_k^2(t+1)) - \frac{1}{2} \sum_{k \in \mathcal{K}} (Q_k^2(t) + E_k^2(t)) \right) \middle| \Theta(t) \right] \leq & \mathbf{r} + \\ \sum_{k \in \mathcal{K}} Q_k(t) \mathbb{E}[a_k(t) - b_k(t) | \Theta(t)] + \sum_{k \in \mathcal{K}} E_k(t) \mathbb{E}[b_k(t) - c_k(t) | \Theta(t)] & \end{aligned} \quad (26)$$

$$\text{式中: } \mathbf{r} \geq \frac{1}{2} \mathbb{E} \left[ \sum_{k \in \mathcal{K}} \left( (a_k(t) - b_k(t))^2 + (b_k(t) - c_k(t))^2 \right) \middle| \Theta(t) \right].$$

根据 Lyapunov 优化理论可知,最小化 Lyapunov 漂移能够保证队列的稳定。本文的目标是在降低传输系统的能耗和区块链时延的同时,保持队列缓冲在低拥塞以满足数据传输时延要求,优化问题就是效用函数,利用 drift-plus-penalty 方法,在时隙  $t$  内,可以将优化目标转化为 Lyapunov 漂移加惩罚函数<sup>[22-23]</sup>,即

$$\Delta \mathcal{L}(\Theta(t)) - V \mathbb{E} \left[ \left( \omega_1 \sum_{k \in \mathcal{K}} E_k^{fc}(t) + (1 - \omega_1) \omega_2 T^{bc}(t) \right) \middle| \Theta(t) \right] \quad (27)$$

式中:  $V \mathbb{E} \left[ \left( \omega_1 \sum_{k \in \mathcal{K}} E_k^{fc}(t) + (1 - \omega_1) \omega_2 T^{bc}(t) \right) \middle| \Theta(t) \right]$  为惩罚项,  $V$  为折中因子。

因此,优化问题分解为每个时隙上的子问题,本文提出一个动态控制算法如算法 1 所示。

**算法 1** 求解原问题的动态控制算法

- (1) 初始化:设置控制参数  $V$ , 权重因子  $\omega_1$  和  $\omega_2$
- (2) For  $t = 1, 2, \dots, T$
- (3) 得到当前队列状态  $Q_k(t)$  和  $E_k(t)$ 。求解关联策略和功率优化联合优化问题

$$\begin{aligned} \min_{\mathbf{X}(t), \mathbf{P}(t)} \quad & \Phi(\mathbf{X}(t), \mathbf{P}(t)) \\ \text{s.t. 式(21a~21f)} & \end{aligned} \quad (28)$$

其中,  $\Phi(\mathbf{X}(t), \mathbf{P}(t)) = \sum_{k \in \mathcal{K}} Q_k(t)(a_k(t) - b_k(t)) + \sum_{k \in \mathcal{K}} E_k(t)(b_k(t) - c_k(t)) - V \left( \omega_1 \sum_{k \in \mathcal{K}} E_k^{fc}(t) + (1 - \omega_1) \omega_2 T^{bc}(t) \right)$

- (4) 得到解  $(\mathbf{X}^*(t), \mathbf{P}^*(t))$
- (5)  $t = t + 1$
- (6) 根据式(8,9),更新队列  $Q_k(t)$  和  $E_k(t)$
- (7) End for
- (8) 输出:  $(\mathbf{X}^*, \mathbf{P}^*)$

### 2.3 基于 D3QN-TD3 的深度强化学习求解方法

对于每个时隙上的子问题,本文采用 D3QN-TD3 的深度强化学习框架求解,通过 D3QN 模型处理 VS 的关联策略,TD3 模型分配连续的功率控制。具体的算法框架如图 2 所示。

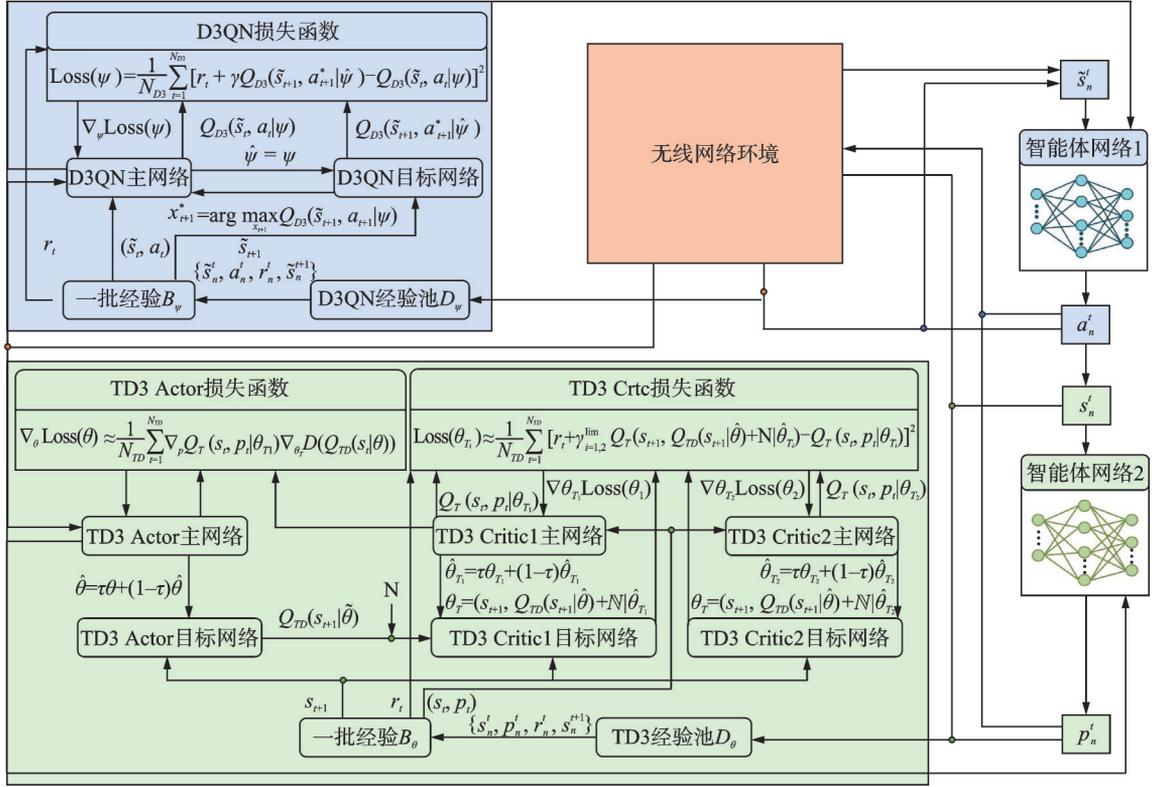


图2 D3QN-TD3混合框架求解的流程图

Fig.2 Flowchart for solving D3QN-TD3 hybrid framework

为了更好地提高网络的整体性能,同时满足各个基站无法在当前时隙内获取完整的CSI要求,需要对D3QN和TD3基本元素的状态、动作和奖励进行合理的设计,其定义分别如下:

(1) 状态。在无线通信中,信道增益反映了无线信号在传播过程中的衰减程度,对于调整传输策略、功率控制有重要意义。因此将终端到无人机的信道增益的数值直接作为状态空间的一部分,即第 $k$ 个VS在时隙 $t$ 的状态空间表示为

$$s_k(t) = \{h_{k,1}^t, h_{k,2}^t, \dots, h_{k,u}^t\} \quad (29)$$

式中: $h_{k,u}^t$ 表示第 $k$ 个终端到第 $u$ 个无人机在时隙 $t$ 的信道增益。

(2) 动作。动作空间是智能体采取动作的集合。智能体需要先选择每个VS该如何关联无人机,随后根据所选择的关联策略进行功率分配。由于功率的选择范围是连续的,理论上存在无数种选择。因此将第 $k$ 个VS在时隙 $t$ 的动作空间表示为

$$\begin{cases} a_k(t) = \{a_{k,1}^t, a_{k,2}^t, \dots, a_{k,u}^t, p_k^t\} \\ \text{s.t. C1: } 0 \leq p_k^t \leq p_{\max} \\ \text{C2: } a_{k,u}^t \in \{0, 1\} \\ \text{C3: } \sum_i^u a_{k,i}^t = 1 \end{cases} \quad (30)$$

式中: $a_{k,u}^t$ 表示第 $k$ 个终端在时隙 $t$ 是否关联第 $u$ 个无人机, $a_{k,u}^t = 1$ 表示关联, $a_{k,u}^t = 0$ 表示不关联; $p_k^t$ 表示第 $k$ 个终端在时隙 $t$ 的功率;约束C1表示VS的发射功率受到限制,可以在0到 $p_{\max}$ 之间的连续空间内进行分配;约束C2、C3表示对于某个无人机 $u$ 来说,VS只能选择关联或不关联两种选择。

(3) 奖励。在强化学习中,目标是通过增加对高 reward 的选择,来最大化总的 reward。因此根据问题式(27),奖励函数定义为优化问题的相反数,即

$$r(t) = -\Phi(X(t), P(t)) \quad (31)$$

所提算法具体步骤如算法 2 所示。

**算法 2** 基于 TD3 和 D3QN 的关联策略与功率联合分配算法

- (1) 初始化 D3QN 算法和 TD3 算法内各个神经网络参数,并将其主网络参数分别记为  $\psi$  和  $\theta$
- (2) 构建经验池  $D_\psi$  和  $D_\theta$
- (3) **while**  $episode < Episode$  **do**
- (4)     初始化系统状态,得到系统状态  $s^t$
- (5)     **while**  $t < Tepisode$  **do**
- (6)         执行 D3QN,得到网络关联策略  $a_t^n$
- (7)         据关联策略  $a_t^n$ ,执行 TD3,得到功率  $p_t^n$
- (8)         根据  $a_t^n$  和  $p_t^n$  执行分配方案,并观察下一阶段状态  $s^{t+1}$
- (9)         为两个 agent 生成经验  $(s_n^t, a_n^t, r_n^t, s_n^{t+1}), (s_n^t, p_n^t, r_n^t, s_n^{t+1})$  放入经验池
- (10)        **if**  $t \bmod Ttrain$  **then**
- (11)            从  $D_\psi$  和  $D_\theta$ , 中分别随机选取数量为  $|B_\psi|$  和  $|B_\theta|$  的一批经验  $B_\psi$  和  $B_\theta$
- (12)            分别根据  $B_\psi$  和  $B_\theta$  训练 D3QN 与 TD3 算法对应神经网络
- (13)            更新神经网络参数  $\psi$  和  $\theta$ , 并将  $\psi$  和  $\theta$  传输给智能体
- (14)            智能体每过一定时间更新  $\psi$  和  $\theta$  替代原本的参数
- (15)         **end if**
- (16)         移动  $Travel$  个时隙
- (17)          $t = t + 1$
- (18)         **end while**
- (19)          $episode = episode + 1$
- (20) **end while**

### 3 仿真结果与分析

假设高铁顶端部署有 4 个 VS,以速度 80 m/s 从点  $(0, 0, 0)$  行驶到点  $(1600, 0, 0)$ ,空中部署 4 个无人机基站,所有无人机飞行的方向和速度与高铁的运行一致,均匀分布在高铁正上方 50 m 处。载波频率  $f_c = 1.8$  GHz,无人机 CPU 最大处理频率  $F^{\max} = 2$  GHz,最大传输功率  $P_k^{\max} = 50$  W,一个签名和一个 MAC 的 CPU 周期为 1 M 个周期和 10 M 周期。共识过程中节点间消息广播的时延  $T_b = 0.1$ ,块间隔  $T_i = 0.5$ 。无人机传输带宽  $B_u = 800$  MHz,处理 1 比特数据需要的 CPU 周期数为 250。相关多普勒干扰因子  $P_{b-ICI} = -19.4$  dB,高斯白噪声频谱密度  $\sigma_{u-back}^2 = \sigma_{ku}^2 = -174$  dBm/Hz。传输数据大小  $D_k(t) = 100$  kb/slot。

接下来开始设定算法中的各项参数。为了兼顾性能和效率,TD3 和 D3QN 神经网络的隐藏层数量均设为 2,其中 TD3 隐藏层神经元数量分别为 400 和 300,而 D3QN 隐藏层的神经元数量都设置为 256;折扣因子  $\gamma$  均设为 0.99;学习率处于 0.001~0.000 1,通过 Adam 动态调整。系统训练时分别随机从经验池中抽取的经验数都为 32,经验池的大小都设置为 256。

在不同权重下,目标函数与迭代轮次的关系曲线如图3所示。从图中可看出,当权重 $\omega_1=0.6$ 时,目标函数在220代左右开始收敛,当 $\omega_1=0.4$ 时,目标函数在200代左右开始收敛,当 $\omega_1=0.2$ 时,目标函数在170代左右开始收敛。这表示对于目标函数而言,当 $\omega_1$ 变大时,意味着在优化过程中对数据传输系统的能耗给予了更多的权重,而相对减少了对区块链系统时延的关注。随着 $\omega_1$ 的减小,整体目标函数的值也逐渐降低,这说明传输系统的能耗优化潜力大于区块链系统时延的优化潜力。当 $\omega_1$ 增加时,目标函数收敛速度变慢,这说明强化学习算法接收到的关于能耗优化的奖励信号变得更加显著,而时延优化的奖励信号减弱,这会导致学习过程需要更多时间来适应新的奖励。在能耗优化中寻找全局最优解会比时延优化更困难,尤其是在高维和复杂的决策空间中。这会导致算法在局部最优解处停滞,减缓收敛速度。

在不同折中因子 $V$ 的影响下,目标函数与迭代轮次的关系曲线如图4所示。从图中可看出,在不同的折中因子 $V$ 的影响下,随着迭代轮次的增加,目标函数最终都是收敛的。折中因子 $V$ 越大,目标函数最终收敛的值越小,这是因为随着折中因子 $V$ 的增加,优化目标更多地关注于降低能耗和时延,Lyapunov 漂移在整体优化中的权重相对减小,这会促使算法更快地调整策略以降低这些指标。

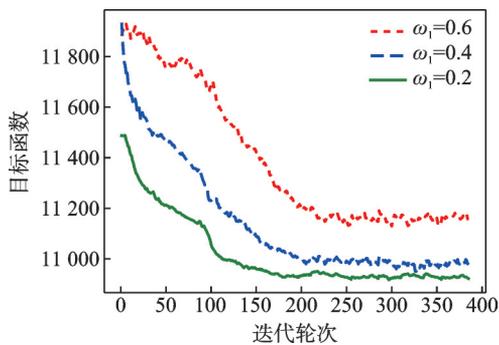


图3 不同权重下目标函数与迭代轮次的关系曲线

Fig.3 Relationship curve between objective function and iteration rounds under different weights

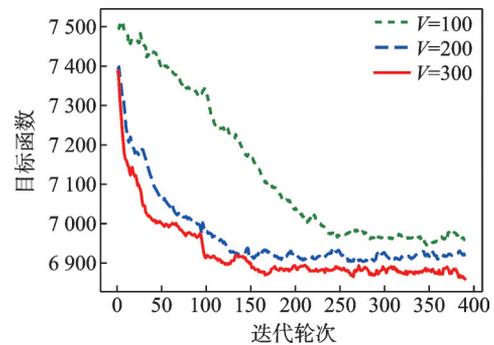


图4 不同折中因子 $V$ 下目标函数和迭代轮次的关系曲线

Fig.4 Relationship curve between objective function and iteration rounds under different compromise factors  $V$

在不同的到达数据包大小 $\lambda_k$ 下,目标函数与迭代轮次的关系曲线如图5所示,可以看出,在到达数据包大小 $\lambda_k$ 的影响下,随着迭代轮次的增加,目标函数最终都是收敛的。 $\lambda_k$ 越大,目标函数最终收敛的值越大。这是因为当上行传输数据包增大时,需要更长的传输时间,也会增加处理队列中数据的积压,为了保证较大的数据包的可靠传输,需要增加发射功率,从而增加了能耗。

图6展示了一个使用强化学习方法优化无人机辅助的高铁数据上行传输的网络中,数据传输系统的能耗和区块链系统的时延的平均目标函数随迭代次数变化的情况。图中包含3条曲线,分别代表以下3种情景:onlyBS,只使用高铁边基站进行数据传输;lose1BS,失去一个高铁基站,即基站数从4降到3;useUAV,使用无人机基站进行辅助数据传输。初始时,3种情景的平均目标函数值均较高,这可能是因为在初期尚未找到更优的策略,因此能耗和时延都未得到有效优化。可以观察到,onlyBS这条线的平均目标函数值在整个迭代过程中变化不大,可能是因为最短距离关联策略相对稳定,而且没有动态优化过程。失去一个高铁基站在最初的几个迭代轮次中,平均目标函数值迅速上升,这可能是因为在基

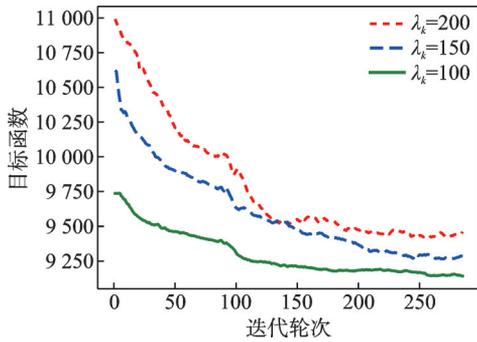


图5 不同到达数据包大小下目标函数和迭代轮次的关系曲线

Fig.5 Relationship curve between objective function and iteration rounds under different packet sizes

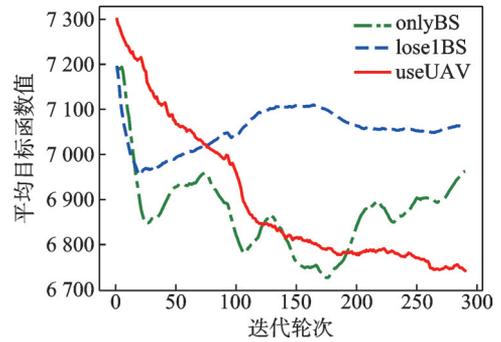


图6 不同场景下平均目标函数值和迭代轮次的关系曲线

Fig.6 Relationship curve between average objective function value and iteration rounds in different scenarios

站数的减少导致 SINR 的增大,使传输速率减小,从而增加了系统的能耗和时延。在迭代后,使用无人机基站进行辅助数据传输的平均目标函数值都有所下降,表明强化学习算法随着迭代学习开始发现并应用更优策略,从而减少能耗和时延,显示出使用无人机基站对于优化系统性能最为有效。

图7展现了不同强化学习策略在训练过程中的平均目标值随训练周期变化的情况。图8为不同强化学习策略的数据队列函数累计分布函数图(Cumulative distribution function, CDF)。图7、8中的4条曲线分别代表:

(1) proposed(所提策略)。蓝色实线表示提出的新策略或算法。

(2) dqn-ddpg: 橙色点划线表示将深度 Q 网络(Deep Q-network, DQN)和深度确定性策略梯度(Deep deterministic policy gradient, DDPG)结合的算法。其中,利用 DQN 进行网络关联策略优化,利用 DDPG 进行功率优化。

(3) ddpg: 绿色点线表示只使用 DDPG 算法进行功率优化,每个时隙开始时每个 VS 等概率随机关联无人机。

(4) dqn: 红色短划线表示只使用 DQN 算法进行网络关联策略优化,将功率划分为 5 个等级,每个时隙开始时智能体从 5 个等级中随机选择一个功率等级。

图7仿真结果显示,所提算法在第 160 代时收敛,且开始时收敛速度快。dqn-ddpg 算法的收敛速度相对所提算法来说慢一些,且在 280 代时收敛,二者的表现在后续过程中逐渐趋于近似。ddpg 算法和 dqn 算法收敛的值比所提算法和 dqn-ddpg 收敛的值大,且 dqn 算法收敛的值最大。

图8仿真结果表示的是所提算法与 3 种基准算法在数据队列长度性能上的比较。通过分析图表数据,观察到所提算法能够显著降低用户的数据队列长度。具体而言,约 94% 的用户享受到了零队列长度,且最长队列长度不会超过 300 b。相比之下,其他 3 种对照算法在实现零数据队列长度的用户比例分别为 91.5%、87.5% 和 80%,显示出较低的 CDF。此外,所提算法还能确保 99% 的用户数据队列长度低于 99 b,即时延违反概率仅为 1%。而对比算法时延违反概率分别为 2.5%、5.8% 和 10.4%,表明所提算法在控制用户的平均数据队列长度方面具有显著的优势,能有效减少数据积压。

综上所述,所提算法效果优于 dqn-ddpg 算法,dqn 算法的效果最差。这是因为问题环境同时包含离散和连续的决策要素,单一策略不足以同时处理两个优化任务,所以不如结合两者的新策略。

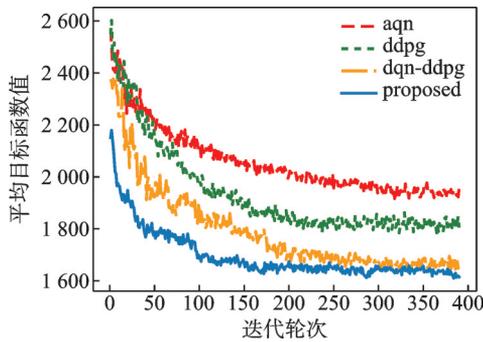


图7 不同算法的平均目标函数值和迭代轮次的关系曲线  
Fig.7 Relationship curve between average objective function value and iteration rounds of different algorithms

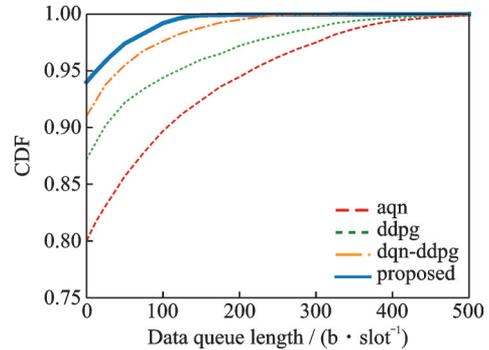


图8 不同算法的数据队列函数累计分布函数图  
Fig.8 Accumulated distribution function diagram of data queue functions for different algorithms

#### 4 结束语

在无人机辅助的高铁应急通信网络架构下,为了实现满足高铁数据低时延传输要求的低能耗和安全通信,通过部署具有计算能力的无人机基站组成空中通信网络,采用区块链技术以保障数据安全性。利用基于Lyapunov理论和机器学习方法得到最优的联合关联策略和功率控制解,在满足传输时延约束下,得到了最小化通信系统能耗和区块链系统时延的最优折中。仿真结果显示,在使用无人机基站进行辅助数据传输时,在不同权重、折中因子和到达数据包大小影响下,目标函数数值最终收敛,且与只有地面基站以及因为自然灾害而导致地面基站缺失的情况相比,性能分别提高了约3%和6%。在参数设置为 $V = 300, \lambda_k = 100$  b/slot时,所提方法在减少能耗和时延方面最为有效。

#### 参考文献:

- [1] ZHOU T, YANG Y, LIU L, TAO C, et al. A dynamic 3-D wideband GBSM for cooperative massive MIMO channels in intelligent high-speed railway communication systems[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(4): 2237-2250.
- [2] 魏昌辛, 吕欣. 西部陆海新通道铁路主通道实施方案研究[J]. *铁道工程学报*, 2022, 39(2): 1-6.  
WEI Changxin, LV Xin. Research on the implementation plan of the main railway channel of the western land sea new corridor [J]. *Journal of Railway Engineering*, 2022, 39(2): 1-6.
- [3] GAO M, AI B, NIU Y, et al. Efficient hybrid beamforming with anti-blockage design for high-speed railway communications [J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(9): 9643-9655.
- [4] PARK Y M, TUN Y K, HAN Z, et al. Trajectory optimization and phase-shift design in IRS-assisted UAV network for smart railway[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(10): 11317-11321.
- [5] 雷阳琦, 段海滨. 基于强化竞争学习鸽群优化的多无人机博弈决策[J]. *中国科学:技术科学*, 2024, 54(1):136-148.  
LEI Yangqi, DUAN Haibin. Multi drone game decision-making based on reinforcement competitive learning pigeon swarm optimization[J]. *Chinese Science: Technical Science*, 2024, 54(1): 136-148.
- [6] ZENG Y, ZHANG R, LIM T J. Wireless communications with unmanned aerial vehicles: Opportunities and challenges[J]. *IEEE Communications Magazine*, 2016, 54(5): 36-42.
- [7] 方晶. 无人机中继系统轨迹优化与资源分配研究[D]. 呼和浩特:内蒙古大学, 2023.  
FANG Jing. Research on trajectory optimization and resource allocation of drone relay system[D]. Hohhot: Inner Mongolia University, 2023.
- [8] HOU L, LEI L, ZHENG K, et al. A Q-learning based proactive caching strategy for non-safety related services in vehicular networks[J]. *IEEE Internet of Things Journal*, 2018, 6(3): 4512-4520.

- [9] CHEN Y, WANG Y, LIU M, et al. Network slicing enabled resource management for service-oriented ultra-reliable and low-latency vehicular networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(7): 7847-7862.
- [10] ROY A, PACHUAU J L, SAHA A K. An overview of queuing delay and various delay based algorithms in networks[J]. Computing, 2021, 103(10): 2361-2399.
- [11] SHE C, YANG C, QUEK T Q S. Cross-layer optimization for ultra-reliable and low-latency radio access networks[J]. IEEE Transactions on Wireless Communications, 2017, 17(1): 127-141.
- [12] CICEK C T, GULTEKIN H, TAVLI B, et al. Backhaul-aware optimization of UAV base station location and bandwidth allocation for profit maximization[J]. IEEE Access, 2020, 8: 154573-154588.
- [13] SONG Q, JIN S, ZHENG F C. Completion time and energy consumption minimization for UAV-enabled multicasting[J]. IEEE Wireless Communication Letters, 2019, 8(3): 821-824.
- [14] WANG Y. Scheduling of UAV-assisted millimeter wave communications for high-speed railway[J]. IEEE Transactions on Vehicular Technology, 2022, 71(8): 8756-8767.
- [15] MENG F, CHEN P, WU L, et al. Power allocation in multi-user cellular networks: Deep reinforcement learning approaches [J]. IEEE Transactions on Wireless Communications, 2020, 19(10): 6255-6267.
- [16] YE X, LI M, YU F R, et al. MEC and blockchain-enabled energy-efficient internet of vehicles based on A3C approach[C]// 2021 IEEE Global Communications Conference (GLOBECOM).[S.l.]: IEEE, 2021.
- [17] FENG J, YU F R, PEI Q, et al. Joint optimization of radio and computational resources allocation in blockchain-enabled mobile edge computing systems[J]. IEEE Transactions on Wireless Communications, 2020, 19(6): 4321-4334.
- [18] YE X, LI M, YU F R, et al. Reliable data transmission over energy-efficient vehicular network based on blockchain and MEC [C]//Proceedings of IEEE International Conference on Communications. [S.l.]: IEEE, 2021.
- [19] 杨明堂, 胡春明, 徐胤泽, 等. 多旋翼混合动力无人机自适应能量管理策略仿真[J]. 南京航空航天大学学报, 2023, 55(6): 1004-1015.  
YANG Mingtang, HU Chunming, XU Yinze, et al. Simulation of adaptive energy management strategy for multi rotor hybrid unmanned aerial vehicles[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2023, 55(6): 1004-1015.
- [20] 郭永安, 王宇翔, 周沂, 等. 边缘网络下多无人机协同计算和资源分配联合优化策略[J]. 南京航空航天大学学报, 2023, 55(5): 757-767.  
GUO Yong'an, WANG Yuao, ZHOU Yi, et al. Joint optimization strategy for multi UAV collaborative computing and resource allocation in edge networks[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2023, 55(5): 757-767.
- [21] 金永光, 叶方伟, 卢晓珍, 等. 区块链赋能的低空智能网[J]. 数据采集与处理, 2024, 39(1): 2-14.  
JIN Yongguang, YE Fangwei, LU Xiaozhen, et al. Low altitude intelligent networking empowered by blockchain[J]. Journal of Data Acquisition and Processing, 2024, 39(1):2-14.
- [22] BI S, HUANG L, WANG H, et al. Lyapunov-guided deep reinforcement learning for stable online computation offloading in mobile-edge computing networks[J]. IEEE Transactions on Wireless Communications, 2021, 20(11): 7519-7537.
- [23] CHANG S, DENG S, WU Y, et al. Online energy balancing strategy based on Lyapunov optimization in mobile crowdsensing [J]. IEEE Transactions on Industrial Informatics, 2023, 19(9): 9266-9279.

## 作者简介:



代海波(1988-),通信作者,男,副教授,硕士生导师,研究方向:无人机辅助通信、无线资源管理、人工智能, E-mail:hbdai@seu.edu.cn.



吴天奇(1998-),男,硕士研究生,研究方向:高速移动场景下智能无线接入与安全传输。



梁轶群(1982-),男,研究员,研究方向:铁路通信技术。



张哲(1988-),男,助理研究员,研究方向:铁路通信技术。



李春国(1983-),男,教授,博士生导师,研究方向:6G蜂窝通信、网络空间安全、人工智能、计算机视觉等。

(编辑:夏道家)