

# 面向数字孪生边缘网络的区块链分片及资源自适应优化机制

蒋丽<sup>1</sup>, 谢胜利<sup>2</sup>, 田辉<sup>3</sup>

(1. 广东工业大学自动化学院, 广东 广州 510006; 2. 广东工业大学物联网信息技术广东省重点实验室, 广东 广州 510006;  
3. 北京邮电大学信息与通信工程学院, 北京 100876)

**摘要:** 为了解决数字孪生边缘网络数据共享面临的隐私和安全问题, 提出一种基于区块链分片的数字孪生边缘网络数据安全共享机制。考虑动态时变的数字孪生边缘网络和边缘网络孪生模型与物理网络映射误差, 建立联合多播簇头选择、本地基站 (BS) 共识接入选择、频谱和计算共识资源分配的自适应资源优化模型, 实现最大化区块链分片交易吞吐量的目标。在数字孪生边缘网络环境下, 提出双层近端策略优化 (PPO) 算法, 求解自适应资源优化问题。仿真结果表明, 所提算法可以有效改善区块链分片交易吞吐量。同时, 在适应映射误差方面优于传统深度强化学习算法。

**关键词:** 数字孪生边缘网络; 区块链分片; 近端策略优化

**中图分类号:** TN929.5

**文献标志码:** A

**DOI:** 10.11959/j.issn.1000-436x.2023044

## Adaptive resource optimization mechanism for blockchain sharding in digital twin edge network

JIANG Li<sup>1</sup>, XIE Shengli<sup>2</sup>, TIAN Hui<sup>3</sup>

1. School of Automation, Guangdong University of Technology, Guangzhou 510006, China

2. Guangdong Key Laboratory of IoT Information Technology, Guangdong University of Technology, Guangzhou 510006, China

3. School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

**Abstract:** To address the security and privacy issues for data sharing in digital twin edge networks, a new distributed and secure data sharing mechanism was developed by leveraging blockchain sharding. Considering dynamic and time varying features in digital twin edge networks and mapping errors of physic networks, an adaptive resource allocation model was formulated which jointly optimized cluster head selection, local base station (BS) consensus access selection, spectrum and computation consensus resources allocation, in order to maximize transactional throughput for the blockchain sharding. Moreover, a digital twin edge networks empowered two-layered proximal policy optimization (PPO) algorithm was designed to obtain the optimal solutions. Numerical results illustrate that the proposed algorithm can effectively improve the transactional throughput for the blockchain sharding. Meanwhile, it can well adapt to the mapping errors.

**Keywords:** digital twin edge network, blockchain sharding, PPO

## 0 引言

为了应对未来新兴服务和全新应用场景的通信需求, 全球业界已开启 6G 的研究探索, 通过采

用多频谱协同与动态管理、信息超表面、感知通信一体化、全新信道编码、空天地海一体化接入等技术<sup>[1]</sup>, 实现每秒太比特级峰值速率、支撑每立方千米 0.1~1 亿设备连接数密度, 并提供即时全息连接。

**收稿日期:** 2022-10-15; **修回日期:** 2023-01-10

**基金项目:** 国家重点研发计划基金资助项目 (No.2020YFB1807801); 移动通信教育部工程研究中心开放研究基金资助项目 (No.cqupt-mct-202003)

**Foundation Items:** The National Key Research and Development Program of China (No.2020YFB1807801), Engineering Research Center of Mobile Communications, Ministry of Education (No.cqupt-mct-202003)

随着极致性能、立体覆盖、泛在连接等 6G 愿景达成共识，个性化服务定制、智能网元节点激增、多场景多业务动态叠加、网络规模持续扩大使网络优化和管理越来越复杂<sup>[2]</sup>。将数字孪生技术应用于网络，构建物理网络的虚拟镜像，助力网络实现低成本试错、智能化决策和新技术部署，可更加高效地应对 6G 网络的复杂挑战<sup>[3-6]</sup>。

近年来，数字孪生技术受到了业界的广泛关注，通过充分利用物理实体、传感器更新以及运行历史等数据，生成物理世界的数字化镜像，即数字世界；同时，在数字世界中模拟物理世界的运行状态，并且预测物理世界的发展趋势，根据预测结果设计防御性方案，提前干预物理世界的运行<sup>[7]</sup>。数字孪生技术已在航天航空<sup>[7]</sup>、智能制造<sup>[8]</sup>、医疗健康<sup>[9]</sup>等多个领域成功应用。数字孪生技术在移动通信中的研究和应用还处在起步阶段。Nguyen 等<sup>[10]</sup>提出数字孪生网络架构，采用数据挖掘、可视化和数据分析等技术，在云端生成物理网络及其运行过程的数字镜像。通过数字镜像与物理网络之间的闭环信息交互，实现物理网络性能评估、运行状态监控和主动维护。Yu 等<sup>[11]</sup>提出基于信息孪生的云-边-端协同网络架构，采用信息孪生技术辅助网络通信，生成网络行为日志并管理用户数字资产，以增强网络的可扩展性、安全性和移动性管理等。将数字孪生部署在网络边缘，用户数据将在网络接入侧就近处理并构成数字孪生边缘网络，可提升无线传输效率，保证孪生网络模型与物理网络之间的实时同步<sup>[12-15]</sup>。然而，数字孪生技术在提高无线网络效能、灵活性和自主能力的同时，也存在新的隐私和安全挑战。数字孪生技术需要海量异构数据的实时采集、数据高频应用以及数据共享，这些数据不仅包含用户的个人隐私信息，还包含生物特征识别信息、车辆自动驾驶控制信息、工业生产数据、家庭电器控制信息等敏感信息。传统集中式数据共享管理面临着单点故障等安全风险。此外，数据拥有者的信任问题将影响数字孪生边缘网络中数据共享与融合，从而阻碍数字孪生技术在移动通信中的发展。

区块链作为一种新兴的去中心化分布式账本技术，通过多方共识和交易备份，可为用户提供安全的数据共享环境<sup>[16-17]</sup>。Lu 等<sup>[18]</sup>在数字孪生边缘网络中，提出基于区块链的孪生网络模型关联机制，改进数字孪生边缘网络的可靠性和安全性。

Yaqoob 等<sup>[19]</sup>在数字孪生网络中，设计基于区块链的数据可信共享模型，提高数字孪生网络的安全性。Jiang 等<sup>[20]</sup>设计基于有向无环图（DAG, directed acyclic graph）区块链的数字孪生边缘网络架构，提出基于双边拍卖的网络资源共享机制，以提高构建的孪生网络模型精准性。但上述研究均未考虑面向数字孪生边缘网络的区块链性能优化问题。随着数字孪生边缘网络中的数据共享交易规模不断增大，区块链验证节点需要消耗大量网络资源（如带宽、能量等）实现全网交易的验证和共识。然而，资源有限的验证节点可能无法应对全网急剧增长的交易，这将降低交易处理速度，导致区块链在数字孪生边缘网络中的可扩展性问题。区块链分片技术通过将网络中节点、交易等划分成不同碎片，每个碎片中的节点只需处理相关交易，这样可以极大提高区块链的可扩展性<sup>[21]</sup>。但当前将区块链分片技术应用于数字孪生边缘网络解决可扩展性问题的研究工作较少。此外，考虑实际边缘孪生模型与物理网络之间存在映射误差，并且数字孪生边缘网络的动态时变环境均将增加区块链交易确认时延，这将无法满足数字孪生边缘网络对数据共享低时延需求。

为了解决以上问题，本文主要研究工作如下。

1) 采用区块链分片技术，设计数字孪生边缘网络数据安全共享机制。将不同服务区域的智能终端和接入点（AP, access point）分片，片内智能终端和 AP 采用 DAG 验证本地多播数据共享交易，生成本地数据共享链。同时，片间边缘服务器采用基于委托权益证明（DPoS, delegated proof of stake）的联盟区块链验证全局孪生数据共享交易，生成全局孪生数据共享链。本地数据共享链和全局孪生数据共享链采用跨链通信交互信息，共同维护数字孪生边缘网络数据共享的安全性。

2) 构建最大化区块链分片交易吞吐量的自适应资源优化模型。提出基于数字孪生边缘网络的双层近端策略优化（PPO, proximal policy optimization）算法，通过联合优化多播簇头选择、本地基站（BS）共识接入选择、频谱和计算共识资源分配，实现最大化区块链分片交易吞吐量目标。

3) 仿真结果表明，本文算法可有效改善区块链分片的交易吞吐量。同时，与传统深度强化学习算法相比，本文算法在适应映射误差方面具有明显的优越性。

# 1 系统模型

本文构建的面向数字孪生边缘网络的区块链分片结构如图 1 所示, 考虑有  $K$  个 BS, 用集合  $\forall k \in \mathcal{K} = \{1, 2, \dots, K\}$  表示。BS<sub>*k*</sub> 配备有移动边缘计算服务器且拥有足够的计算资源。同时, BS<sub>*k*</sub> 覆盖的服务区域用  $\mathcal{A}_k$  表示, 对于任意  $k$  和  $k'$ ,  $k \neq k'$ , 满足如下条件:  $\bigcup_{\forall k} \mathcal{A}_k = \mathcal{A}$ ,  $\mathcal{A}_k \cap \mathcal{A}_{k'} = \emptyset$ 。在服务区域  $\mathcal{A}_k$  中, 有  $U$  个智能终端簇, 用集合  $\forall u \in \mathcal{U} = \{1, 2, \dots, U\}$  表示。在簇  $u$  中, 均匀分布  $N^u$  个智能终端, 用集合  $\mathcal{N}^u$  表示。考虑服务区域  $\mathcal{A}_k$  中分布有  $N_k$  个智能终端, 数字孪生边缘网络分布有  $N$  个智能终端, 用集合  $\forall n \in \mathcal{N} = \{1, 2, \dots, N\}$  表示, 且  $N = \sum_{k=1}^K N_k$ 。除此之外, 服务区域  $\mathcal{A}_k$  中分布有  $J_k$  个 AP, 用集合  $\forall j_k \in \mathcal{J}_k = \{1, 2, \dots, J_k\}$  表示。数字孪生边缘网络分布有  $J$  个 AP, 用集合  $\forall j \in \mathcal{J} = \{1, 2, \dots, J\}$  表示, 且  $J = \sum_{k=1}^K J_k$ 。AP 是小蜂窝接入点。智能终端  $n$  的数据表示为  $D_n = \{(x_{n1}, y_{n1}), \dots, (x_{nD_n}, y_{nD_n})\}$ , 且 AP<sub>*j<sub>k</sub>*</sub> 的数据表示为  $D_{j_k} = \{(x_{j_k1}, y_{j_k1}), \dots, (x_{j_kD_{j_k}}, y_{j_kD_{j_k}})\}$ 。D<sub>*n*</sub> 和 D<sub>*j<sub>k</sub>*</sub> 分别表示智能终端  $n$  和 AP<sub>*j<sub>k</sub>*</sub> 的数据大小。BS<sub>*k*</sub> 连续采集服务区域  $\mathcal{A}_k$  内智能终端和 AP 的数据, 并且计算和

分析智能终端和 AP 的行为特征, 在边缘服务器构建服务区域内智能终端、AP、拓扑关系及环境的边缘孪生模型。为了减少资源有限智能终端的数据上传时延, 提高虚实映射准确率, BS<sub>*k*</sub> 雇佣  $J_k$  个 AP 协作传输智能终端的数据。如图 1 所示, AP<sub>*j<sub>k</sub>*</sub> 充当簇头, 在同一时刻能且只能服务一个智能终端簇。用  $\alpha_{j_k}^u$  表示 AP<sub>*j<sub>k</sub>*</sub> 与智能终端簇  $u$  的关联关系, 即当 AP<sub>*j<sub>k</sub>*</sub> 辅助智能终端簇  $u$  上传数据时,  $\alpha_{j_k}^u = 1$ , 否则  $\alpha_{j_k}^u = 0$ 。

智能终端  $n$  和 AP<sub>*j<sub>k</sub>*</sub> 在 BS<sub>*k*</sub> 的边缘孪生模型分别表示为  $DT_n = \{\mathcal{M}_n, D_n, s_n(t), s_n(t+1)\}$  和  $DT_{j_k} = \{\mathcal{M}_{j_k}, D_{j_k}, s_{j_k}(t), s_{j_k}(t+1)\}$ 。其中,  $\mathcal{M}_n$  和  $\mathcal{M}_{j_k}$  分别表示智能终端  $n$  和 AP<sub>*j<sub>k</sub>*</sub> 的行为模型,  $s_n(t)$  和  $s_{j_k}(t)$  分别表示智能终端  $n$  和 AP<sub>*j<sub>k</sub>*</sub> 的实时状态,  $s_n(t+1)$  和  $s_{j_k}(t+1)$  分别表示智能终端  $n$  和 AP<sub>*j<sub>k</sub>*</sub> 的更新状态。与此同时, BS<sub>*k*</sub> 也可生成自身的边缘孪生模型。进而, BS<sub>*k*</sub> 基于生成的边缘孪生模型对服务区域  $\mathcal{A}_k$  内的智能终端、AP 等的运行状态进行模拟和仿真并生成孪生数据, 并且不同服务区域的 BS 之间可以共享孪生数据。BS 通过在构造的边缘孪生模型中使用合适的资源分配算法设计网络策略 (如路由选择、接入选择等), 以满足物理网络性能需求。同时, 在将设计的网络策略部署到物理网络之前, 可在边缘

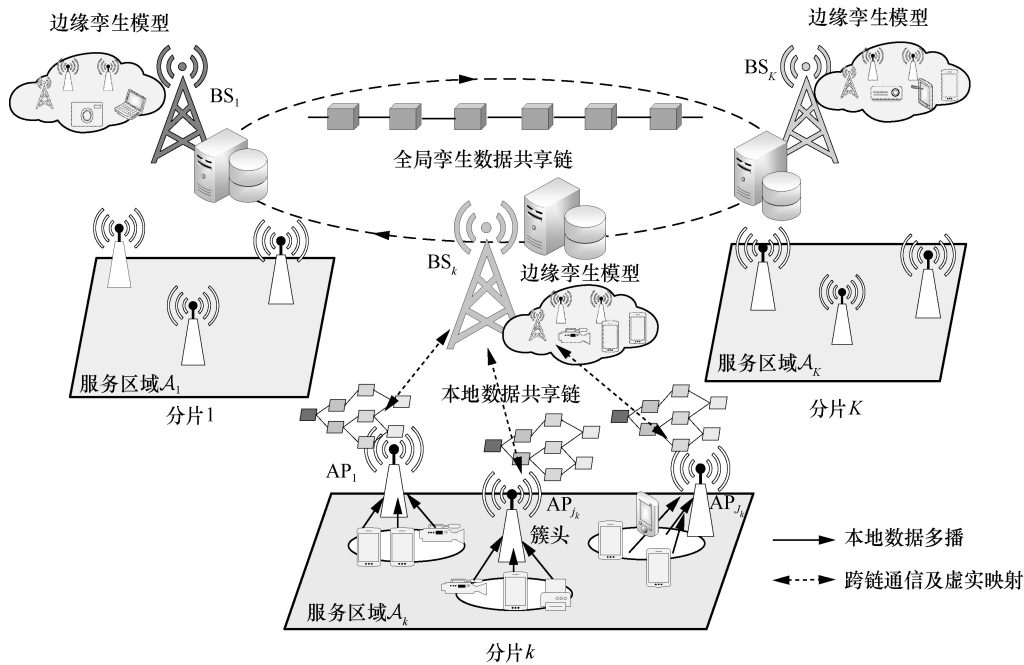


图 1 面向数字孪生边缘网络的区块链分片结构

孪生模型中测试，为物理网络提供优化的网络策略。本文考虑的网络策略为 AP<sub>j<sub>k</sub></sub> 簇头选择、共识接入选择以及共识资源分配策略将在第 2 节详细描述。

然而，数字孪生边缘网络也存在新的隐私和安全挑战。为了保护智能终端、AP 之间本地数据共享安全，以及 BS 之间孪生数据共享安全，本文采用区块链分片技术，将每个服务区域 A<sub>k</sub> 中的智能终端和 AP 划分为独立分片 S<sub>k</sub>，片内智能终端和 AP 采用基于累计权重共识机制的 DAG 验证本地数据共享交易，生成本地数据共享链。片间边缘服务器采用基于 DPoS 共识机制的联盟区块链验证全局孪生数据共享交易，生成全局孪生数据共享链。同时，本地数据共享链和全局孪生数据共享链采用跨链通信交互信息，共同维护数字孪生边缘网络数据共享的安全性。数字孪生边缘网络的无线传输环境具有随机性，且计算、通信等可用资源呈现动态变化特性。此外，物理网络与边缘孪生模型映射过程中，数据传输差错将会导致边缘孪生模型与物理网络存在轻微误差。这均会增加数字孪生边缘网络中区块链分片交易确认时延。为此，本文将构造数字孪生边缘网络中区块链分片资源自适应优化模型，实现最大化区块链分片交易吞吐量目标。

### 1.1 本地数据共享链的资源自适应优化模型

DAG 作为一种新的分布式账本技术，可以增强传统区块链处理交易的效率<sup>[18]</sup>。与传统区块链的单链结构不同，DAG 是由用户发布的小额微交易构成的网络，支持异步并发记账。DAG 已广泛应用于分布式网络以支持高频微交易的验证<sup>[22]</sup>。本文采用 DAG 验证数字孪生边缘网络本地数据共享交易，并生成本地数据共享链。在分片 S<sub>k</sub> 中，为了减少智能终端的资源消耗，AP<sub>j<sub>k</sub></sub> 负责存储本地数据共享链，

并采用多播传输方式与分片中的其余 AP 同步本地数据共享链。为了将本地数据共享交易添加到本地数据共享链上，AP<sub>j<sub>k</sub></sub> 需要先验证本地数据共享链上两笔未被验证过的交易，然后将验证的两笔交易的哈希值添加到新的本地数据共享交易。此时，新的本地数据共享交易已被添加到本地数据共享链上，并被存入 AP<sub>j<sub>k</sub></sub> 的缓存，排队等待多播传输。AP<sub>j<sub>k</sub></sub> 的本地数据共享交易队列遵循先进先出原则，新的本地数据共享交易到达服从泊松点过程，用 λ 表示到达速率。考虑 AP<sub>j<sub>k</sub></sub> 多播传输的最大本地数据共享交

易为 L 个，本地数据共享交易 m<sub>j</sub> 的排队等待时延可表示为

$$T_q(m_{j_k}) = \iota T^{\text{com}} - \frac{L}{2\lambda} \quad (1)$$

其中，ι 表示缓存系数，T<sup>com</sup> 表示 AP<sub>j<sub>k</sub></sub> 平均多播时延。因此，T<sup>com</sup> 为

$$T^{\text{com}} = \frac{1}{J_k + N_k} \left( \sum_{j_k \in \mathcal{J}_k} \left( \max_{n \in \mathcal{N}_k^u} \frac{|m_{j_k}|}{R_{n,j_k}} + \max_{j_k, j'_k \in \mathcal{J}_k, j_k \neq j'_k} \frac{|Lm_{j_k}|}{R_{j_k, j'_k}} \right) \right) \quad (2)$$

其中，|m<sub>j<sub>k</sub></sub>| 表示本地数据共享交易 m<sub>j<sub>k</sub></sub> 的大小。考虑智能终端 n 采用正交频分多址接入（OFDMA, orthogonal frequency division multiple access）技术将本地数据共享交易上传到 AP<sub>j<sub>k</sub></sub>，R<sub>n,j<sub>k</sub></sub> 表示智能终端 n 上传本地数据共享交易到 AP<sub>j<sub>k</sub></sub> 的速率，且 R<sub>n,j<sub>k</sub></sub> = W<sub>n,j<sub>k</sub></sub> W<sub>0</sub> log(1 + Γ<sub>n,j<sub>k</sub></sub>)，W<sub>n,j<sub>k</sub></sub> 表示 AP<sub>j<sub>k</sub></sub> 分配给智能终端 n 的子信道，W<sub>0</sub> 表示子信道带宽，Γ<sub>n,j<sub>k</sub></sub> 表示智能终端 n 到 AP<sub>j<sub>k</sub></sub> 的信干噪比。此外，考虑 AP<sub>j<sub>k</sub></sub> 采用 OFDMA 技术将本地数据共享交易多播到片内 AP<sub>j<sub>k</sub></sub>，R<sub>j<sub>k</sub>,j'<sub>k</sub></sub> 表示 AP<sub>j<sub>k</sub></sub> 多播本地数据共享交易到 AP<sub>j'<sub>k</sub></sub> 的速率，且 R<sub>j<sub>k</sub>,j'<sub>k</sub></sub> = W<sub>j<sub>k</sub>,j'<sub>k</sub></sub> W<sub>0</sub> log(1 + Γ<sub>j<sub>k</sub>,j'<sub>k</sub></sub>)，W<sub>j<sub>k</sub>,j'<sub>k</sub></sub> 表示 AP<sub>j<sub>k</sub></sub> 分配给 AP<sub>j'<sub>k</sub></sub> 的子信道，Γ<sub>j<sub>k</sub>,j'<sub>k</sub></sub> 表示 AP<sub>j<sub>k</sub></sub> 到 AP<sub>j'<sub>k</sub></sub> 的信干噪比。当 AP<sub>j'<sub>k</sub></sub> 收到新的本地数据共享交易并执行验证时，如果收到的交易是正确的，则它将等待本地数据共享链确认。本地数据共享交易 m<sub>j<sub>k</sub></sub> 的权重为

$$W(m_{j_k}) = \frac{\sum_{n=1}^{N_k^u} D_n}{\sum_{n=1}^{N_k} D_n} \mathcal{G}_{j_k}(m_{j_k}) \quad (3)$$

其中，∑<sub>n=1</sub><sup>N<sub>k</sub><sup>u</sup></sup> D<sub>n</sub> 表示 AP<sub>j<sub>k</sub></sub> 服务的第 u 个簇内 N<sub>k</sub><sup>u</sup> 个智能终端的数据，∑<sub>n=1</sub><sup>N<sub>k</sub></sup> D<sub>n</sub> 表示服务区域 A<sub>k</sub> 覆盖范围内所有智能终端的数据，G<sub>j<sub>k</sub></sub>(m<sub>j<sub>k</sub></sub>) 表示 AP<sub>j<sub>k</sub></sub> 的信誉值。本地数据共享交易 m<sub>j<sub>k</sub></sub> 的累计权重 CW(m<sub>j<sub>k</sub></sub>) 为

$$CW(m_{j_k}) = W(m_{j_k}) + \sum_{i=1}^l \frac{W(m_i)}{\Delta \mathcal{G}_i} \quad (4)$$

其中，W(m<sub>i</sub>) 表示本地数据共享交易 m<sub>i</sub> 的权重，本地数据共享交易 m<sub>i</sub> 可直接或间接验证本地数据共

享交易  $m_{j_k}$ ； $\Delta \mathcal{G}_i$  表示信誉估计值，且  $\Delta \mathcal{G}_i = \mathcal{G}_i(m_{j_k}) - \mathcal{G}_i(m_{j_k})$ 。本地数据共享交易的确认取决于其不断增加的累积权重，该累积权重由基于马尔可夫链蒙特卡罗的交易选择算法确定<sup>[19]</sup>。因此，从 AP <sub>$j_k$</sub>  发布本地数据共享交易  $m_{j_k}$  到该交易被确认的时延表示为

$$T^{\text{conf}}(m_{j_k}) = T_q(m_{j_k}) + T_{\text{ac}}(m_{j_k}) \quad (5)$$

其中， $T_q(m_{j_k})$  表示式(1)中本地数据共享交易  $m_{j_k}$  的排队时延， $T_{\text{ac}}(m_{j_k})$  表示本地数据共享交易  $m_{j_k}$  的权重累积时延<sup>[20]</sup>，即

$$T_{\text{ac}}(m_{j_k}) = 2.84 \ln(2T^{\text{com}} \lambda) T^{\text{com}} + \frac{CW_{\text{th}} - CW(m_{j_k})}{\lambda} \quad (6)$$

其中， $CW_{\text{th}}$  表示累计权重确认阈值，即当本地数据共享交易  $m_{j_k}$  的累计权重  $CW(m_{j_k})$  增加到  $CW_{\text{th}}$  时  $m_{j_k}$  被本地数据共享链确认。考虑交易吞吐量作为一项重要指标，用以衡量区块链处理交易的能力。因此，本地数据共享链的交易吞吐量表示为

$$\Omega(T^{\text{com}}, \lambda) = \frac{T^{\text{com}} \lambda}{T^{\text{conf}}(m_{j_k})} \quad (7)$$

其中，分子表示 AP <sub>$j_k$</sub>  的多播时间间隔内到达的新的本地数据共享交易，分母表示本地数据共享交易的确认时延。由式(7)可以看出，AP 的多播时延对本地数据共享链的交易吞吐量有显著影响。如果 AP 贡献更多的通信资源多播本地数据共享交易，则交易吞吐量将会增加。然而，由于 AP 的通信资源有限，额外的资源开销导致 AP 服务自身业务的资源减少。因此，BS <sub>$k$</sub>  需要设计合适的激励策略以补偿 AP 在本地数据共享交易验证过程中的资源消耗。将 AP <sub>$j_k$</sub>  多播传输时延的单价定义为  $\eta_m/s$ ，且 AP <sub>$j_k$</sub>  贡献的通信资源（如带宽资源）的单位成本价格定义为  $\mu_c/\text{Hz}$ 。则 BS <sub>$k$</sub>  的效用函数可表示为

$$U_{\text{BS}_k} = \sum_{j_k=1}^{J_k} \alpha_{j_k}^u \left( \eta_m e^{-T_{j_k}^{\text{com}}} - \mu_c W_{j_k} W_0 \right) \quad (8)$$

其中，第一项表示智能终端簇  $u$  需支付给 BS <sub>$k$</sub>  的本地数据共享交易验证服务费用，且  $T_{j_k}^{\text{com}} = \max_{n \in N_k^u} \frac{|m_{j_k}|}{R_{n,j_k}} + \max_{j_k, j_k' \in \mathcal{J}_k, j_k \neq j_k'} \frac{|Lm_{j_k}|}{R_{j_k, j_k'}}$ ；第二项表示 BS <sub>$k$</sub>  提供给 AP <sub>$j_k$</sub>  的资源消耗补偿，且

$$W_{j_k} = \sum_{n=1}^{N_k^u} W_{n,j_k} + \sum_{j_k=1, j_k \neq j_k'}^{J_k} W_{j_k, j_k}。为了提高带宽使用效$$

率，本文考虑数字孪生边缘网络中的  $J$  个 AP 共享相同带宽资源。因此，通过优化智能终端簇  $u$  和 AP <sub>$j_k$</sub>  的关联变量  $\alpha_{j_k}^u$ ，以及 AP <sub>$j_k$</sub>  的带宽资源分配，设计如下最大化 BS 效用函数的优化问题

$$\begin{aligned} & \max_{\alpha, W_c} \sum_{k \in \mathcal{K}} U_{\text{BS}_k} \\ & \text{s.t. C1: } \sum_{u \in \mathcal{U}} \alpha_{j_k}^u = 1, \forall j \in \mathcal{J}, k \in \mathcal{K} \\ & \text{C2: } \sum_{j_k \in \mathcal{J}_k} \alpha_{j_k}^u = 1, \forall u \in \mathcal{U}, k \in \mathcal{K} \\ & \text{C3: } \sum_{k \in \mathcal{K}} \sum_{j_k \in \mathcal{J}_k} W_{j_k} W_0 \leq W_c^{\text{max}} \end{aligned} \quad (9)$$

其中， $\alpha = \{\alpha_{j_1}^u, \dots, \alpha_{j_k}^u, \dots, \alpha_{j_K}^u\}$  表示  $K$  个 BS 覆盖范围内智能终端簇和 AP 的关联向量，且  $\alpha_{j_k}^u \in \{0, 1\}$ ； $W_c = \{W_{j_1}, \dots, W_{j_k}, \dots, W_{j_K}\}$  表示 AP 的带宽分配向量；约束 C1 表示每个智能终端簇仅能获得一个 AP 的协作服务；约束 C2 表示每个 AP 仅能给一个智能终端簇提供协作服务；约束 C3 表示数字孪生边缘网络中  $J$  个 AP 的带宽分配不能超过总带宽  $W_c^{\text{max}}$ 。当 BS <sub>$k$</sub>  覆盖范围内智能终端簇的本地数据通过本地数据共享链验证后，AP 采用专用信道将通过验证的智能终端簇的本地数据协作传输给 BS <sub>$k$</sub> 。同时，BS <sub>$k$</sub>  计算和分析收集的数据，生成边缘孪生模型  $k$ ，且基于生成的边缘孪生模型  $k$  对服务区域  $\mathcal{A}_k$  内的智能终端、AP 等的运行状态进行模拟和仿真并生成孪生数据。不同服务区域 BS 之间可共享孪生数据。为了保证不同服务区域 BS 之间孪生数据安全共享，本文设计了全局孪生数据共享链。

## 1.2 全局孪生数据共享链的资源自适应优化模型

DPoS 共识机制通过选择部分代表参与区块的验证和共识过程，可有效提高交易处理效率<sup>[23]</sup>。本文采用 DPoS 共识机制验证数字孪生边缘网络全局孪生数据共享交易，生成全局孪生数据共享链。 $V$  个验证节点（即 BS）采用高效的 DPoS 共识机制维护全局孪生数据共享链。其中， $V$  个验证节点由智能终端投票选出，用集合  $\forall v \in \mathcal{V} = \{1, 2, \dots, V\}$  表示，且  $V \leq K$ 。在区块验证时隙，考虑有  $Q$  个本地 BS 需要将产生的区块添加到全局孪生数据共享链，并向验证节点集合  $\mathcal{V}$  发送共识请求，用集合  $\forall q \in \mathcal{Q} = \{1, 2, \dots, Q\}$  表示。在本轮共识过程中，区块管理者  $p$  ( $p \in \mathcal{V}$ ) 管理验证节点并执行区块验

证。区块管理者仅允许一个本地 BS 的区块接入当前共识过程。用  $\gamma_q$  表示本地 BS 接入变量，若第  $q$  个本地 BS 的区块接入当前共识过程则  $\gamma_q = 1$ ，否则  $\gamma_q = 0$ 。同时，每个验证节点  $v$  ( $v \in \{\mathcal{V} \setminus p\}$ )，负责验证接入区块的交易，并将验证和审计结果反馈给区块管理者  $p$ 。

共识过程包括请求、预准备、准备、审计和反馈 5 个阶段。共识过程中本地接入  $BS_q$ 、区块管理者  $p$  以及验证节点  $v$  之间采用 OFDMA 技术多播传输区块。考虑本地接入  $BS_q$ 、区块管理者  $p$  以及验证节点  $v$  分配的子信道分别表示为  $W_{q,p}$ 、 $W_{p,v}$ 、 $W_{v,v'}$  和  $W_{v,p}$ ，且本地  $BS_q$  产生孪生数据共享区块大小表示为  $S_{B,q}$ 。

因此，共识过程的区块多播传输时延表示为

$$T^G = \min \left\{ \frac{S_{B,q}}{R_{q,p}}, T_{th} \right\} + \min \left\{ \max_{v \in \{\mathcal{V} \setminus p\}} \left\{ \frac{S_{B,q}}{R_{p,v}} \right\}, T_{th} \right\} + \min \left\{ \max_{v, v' \in \mathcal{V}, v \neq p} \left\{ \frac{S_{B,q}}{R_{v,v'}} \right\}, T_{th} \right\} + \min \left\{ \max_{v, v' \in \mathcal{V}, v \neq v'} \left\{ \frac{S_{B,q}}{R_{v,v'}} \right\}, T_{th} \right\} + \min \left\{ \max_{v, p \in \mathcal{V}} \left\{ \frac{S_{B,q}}{R_{v,p}} \right\}, T_{th} \right\} \quad (10)$$

其中，每个求和项分别对应共识过程 5 个阶段的区块多播传输时延； $T_{th}$  表示每个阶段区块传输超时阈值； $R_{q,p}$  表示从本地 BS  $q$  到区块管理者  $p$  的传输速率，且  $R_{q,p} = W_{q,p} W_0 \log(1 + \Gamma_{q,p})$ ； $R_{p,v}$  表示区块管理者  $p$  到验证节点  $v \in \mathcal{V}$  的传输速率，且  $R_{p,v} = W_{p,v} W_0 \log(1 + \Gamma_{p,v})$ ； $R_{v,v'}$  表示验证节点  $v$  到验证节点  $v' \in \{\mathcal{V} \setminus v\}$  的传输速率，且  $R_{v,v'} = W_{v,v'} W_0 \log(1 + \Gamma_{v,v'})$ ； $R_{v,p}$  表示从验证节点  $v$  到区块管理者  $p$  的传输速率，且  $R_{v,p} = W_{v,p} W_0 \log(1 + \Gamma_{v,p})$ ； $\Gamma_{q,p}$ 、 $\Gamma_{p,v}$ 、 $\Gamma_{v,v'}$ 、 $\Gamma_{v,p}$  分别表示本地  $BS_q$ 、区块管理者  $p$ 、验证节点  $v$  以及验证节点  $v'$  之间多播传输子信道的信干噪比。

考虑孪生数据共享交易的平均大小为  $\bar{\omega}$ ，因此区块中包含的孪生数据共享交易个数为  $\frac{S_{B,k}}{\bar{\omega}}$ 。在共识过程中，区块管理者  $p$  验证区块以及孪生数据交易的计算负载表示为  $o_p$ ，区块管理者  $p$  能提供的计算资源表示为  $f_p$ 。同时，验证节点  $v$  验证区块以及孪生数据交易的计算负载表示为  $o_v$ ，验证节点  $v$  能提供的计算资源表示为  $f_v$ 。因此，在共识过程验证

区块和孪生数据交易的计算时延表示为

$$T^v = \max_{l \in \mathcal{V}} \left\{ \frac{o_l}{f_l} \right\} \quad (11)$$

由以上分析可知，全局孪生数据共享链的交易吞吐量表示为

$$\Phi(S_{B,q}, T^{B,q}) = \frac{\left\lfloor \frac{S_{B,q}}{\bar{\omega}} \right\rfloor}{T^{B,q}} \quad (12)$$

其中， $T^{B,q}$  表示区块确认时延，它由区块多播传输时延  $T^G$  和区块验证计算时延  $T^v$  组成，即  $T^{B,q} = T^G + T^v$ 。根据式(11)和式(12)可知，在给定区块大小  $S_{B,q}$  的情况下，如果区块管理者和验证节点贡献更多的带宽资源和计算资源用于多播传输区块和验证区块，式(12)中的交易吞吐量将会增加。同时，这会导致区块管理者和验证节点服务其他业务的资源减少，例如基于数字孪生的仿真和测试计算。为了减少区块确认时延和提高交易吞吐量，本地  $BS_q$  需要向区块管理者  $p$  和所有验证节点支付区块验证服务费以补偿其资源消耗。定义区块管理者和验证节点区块确认时延的单价为  $\eta_v$  / s。定义区块管理者和验证节点贡献的带宽资源单位成本价格为  $\lambda_v$  / Hz，以及贡献的计算资源单位成本价格为  $\zeta_v$ 。由此定义区块管理者  $p$  的效用函数为

$$U_p^v = \gamma_k \left( \eta_v \Phi(S_{B,k}, T^{B,k}) - \lambda_v \sum_{l=1}^V W_l - \zeta_v \sum_{l=1}^V f_l \right) \quad (13)$$

在全局孪生数据共享链中，区块管理者  $p$  的目标是通过选择接入的本地  $BS_q$ ，与优化区块验证过程中的带宽资源和计算资源来最大化其效用函数，设计的优化问题为

$$\begin{aligned} & \max_{\gamma, W, f} U_p^v \\ & \text{s.t. C4: } \sum_{q=1}^Q \gamma_q = 1, \forall q \in Q \\ & \text{C5: } \sum_{l=1}^V W_l W_0 \leq W_v^{\max}, \forall l \in \mathcal{V} \\ & \text{C6: } f_l \leq f_l^{\max}, \forall l \in \mathcal{V} \\ & \text{C7: } \eta_v \Phi(S_{B,q}, T^{B,q}) \leq U_{BS_q} \end{aligned} \quad (14)$$

其中， $\gamma = \{\gamma_q\}$ 、 $q \in Q$  表示本地 BS 接入向量， $W_v = \{W_l\}$ 、 $l \in \mathcal{V}$  表示区块管理者和验证节点的带宽资源分配向量， $f_v = \{f_l\}$ 、 $l \in \mathcal{V}$  表示区块管理者和验证节点计算资源分配向量，约束 C4 表示在每轮共

识过程中仅允许一个本地 BS 接入, 约束 C5 表示区块管理者和验证节点的带宽分配不能超过总带宽  $W_v^{\max}$ , 约束 C6 表示区块管理者和验证节点的计算资源分配不能超过最大可用计算资源  $f_l^{\max}$ , 约束 C7 表示本地 BS<sub>q</sub> 支付的验证服务费用不能超过其在本地数据共享链中获得的收益。

## 2 区块链分片的资源自适应优化算法设计

由于数字孪生边缘网络无线传输环境的随机性和计算、通信等资源的动态时变性, 以及边缘孪生模型与物理网络的映射误差。用传统的优化方法较难求解优化问题式(9)和式(14)。由 OpenAI 公司于 2017 年提出的 PPO 算法相较于其他深度强化学习算法具有更好的样本复杂度。通过设置置信区间, PPO 算法可以确保控制策略在置信区间内更新幅度不会过大, 从而使算法更加稳定, 且能够更有效地适应差错<sup>[24]</sup>。Cronrath 等<sup>[25]</sup>提出采用 PPO 算法适应钣金装配线数字孪生模型映射误差, 以获得夹具最优固定位置。目前, 将 PPO 算法用于数字孪生边缘网络的区块链分片, 解决边缘孪生模型与物理网络映射误差条件下的资源优化问题, 还没有相关研究。

本节设计基于数字孪生边缘网络的双层 PPO 算法, 以求解区块链分片的资源自适应优化问题式(9)和式(14), 如图 2 所示。

在设计的优化算法中,  $K$  个 BS 采用多代理 PPO 算法。此外, 区块管理者  $p$  采用单代理 PPO 算法。

双层 PPO 算法的运行环境分别由  $K$  个 BS 和区块管理者  $p$  的边缘服务器中边缘孪生模型组成。如图 2 中步骤①,  $K$  个 BS 和区块管理者  $p$  分别从边缘孪生模型中观测维护本地数据共享链以及全局孪生数据共享链的 AP 和验证节点的状态, 并将观测状态输入 PPO 神经网络, 通过步骤②输出资源优化决策, 求解资源自适应优化问题式(9)和式(14)。然后, 将输出的资源优化决策在步骤③边缘孪生模型中进行仿真测试。最后, 将经过测试的资源优化决策通过步骤④反馈给对应物理节点执行。相比传统 PPO 算法, 本文设计的基于数字孪生边缘网络的双层 PPO 算法,  $K$  个 BS 和区块管理者  $p$  均从边缘孪生模型中获取状态信息, 可减少 PPO 算法训练过程中基站和区块管理者与物理节点交互的资源开销。

### 2.1 基于数字孪生边缘网络的双层 PPO 算法

在本文提出的基于数字孪生边缘网络的双层 PPO 算法中, 将决策问题构造造成双层马尔可夫决策过程, 其中, 状态空间、动作空间和奖励函数的定义分别如下。

1) 状态空间。在决策时刻  $t (t=1, 2, \dots)$ , 第  $k$  个 BS 中维护本地数据共享链的 AP 孪生模型状态包括智能终端  $n$  到 AP <sub>$j_k$</sub>  的信干噪比  $\Gamma_{n, j_k}$ , 以及 AP <sub>$j_k$</sub>  到 AP <sub>$j'_k$</sub>  信干噪比  $\Gamma_{j_k, j'_k}$ , 且  $j_k, j'_k \in \mathcal{J}_k, j_k \neq j'_k, n \in \mathcal{N}_k^u, u \in \mathcal{U}$ , 第  $k$  个 BS 的状态空间表示为

$$\mathcal{S}_{k,t}^s = [\Gamma_{n, j_k}, \Gamma_{j_k, j'_k}] \quad (15)$$

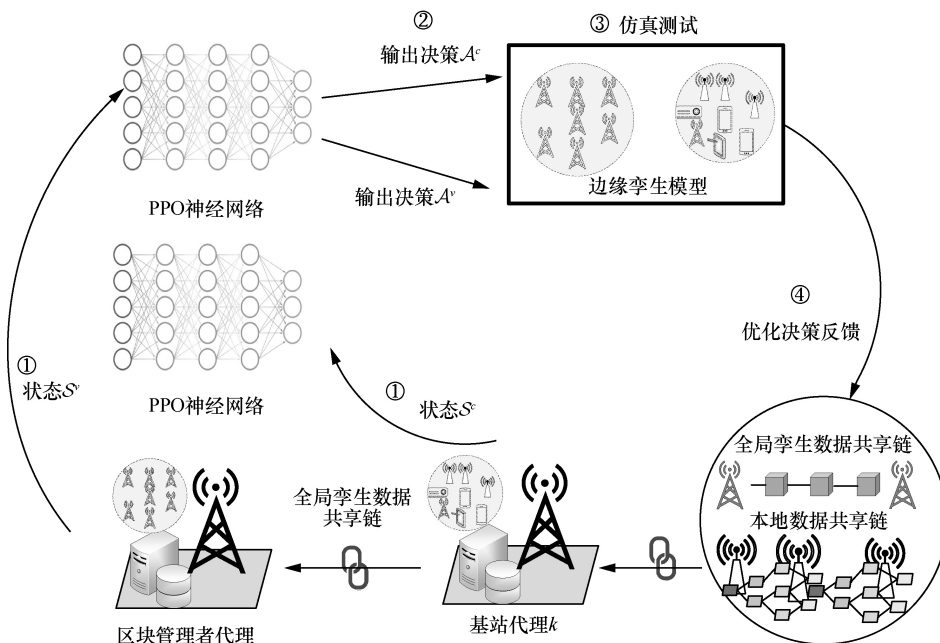


图 2 基于数字孪生边缘网络的双层 PPO 算法

因此， $K$  个 BS 的状态空间表示为

$$\mathcal{S}_i^c = \{\mathcal{S}_{1,t}^c, \dots, \mathcal{S}_{k,t}^c, \dots, \mathcal{S}_{K,t}^c\} \quad (16)$$

同样地，在区块管理者  $p$  中，维护全局孪生数据共享链的验证节点孪生模型状态包括区块大小  $S_{B,q}$ 、本地接入  $BS_q$ 、区块管理者  $p$ 、验证节点  $v$  和验证节点  $v'$  之间多播传输子信道的信干噪比  $\Gamma_{q,p}$ 、 $\Gamma_{p,v}$ 、 $\Gamma_{v,v'}$ 、 $\Gamma_{v,p}$ ，以及区块管理者和验证节点的最大可用计算资源  $f_i^{\max}$ ，且  $p, v, v' \in \mathcal{V}$ ， $p \neq v$ ， $p \neq v'$ ， $v \neq v'$ ，则区块管理者  $p$  的状态空间表示为

$$\mathcal{S}_i^v = [S_{B,q}, \Gamma_{q,p}, \Gamma_{p,v}, \Gamma_{v,v'}, f_i^{\max}] \quad (17)$$

2) 动作空间。为了最大化  $K$  个 BS 和区块管理者  $p$  的效用，需要调整决策变量以适应数字孪生边缘网络动态时变特性。在本地数据共享链中需要调整的决策变量包括智能终端簇和 AP 的关联向量  $\alpha$ ，以及 AP 的带宽资源分配向量  $\mathcal{W}_c$ 。在全局孪生数据共享链中需要调整的决策变量包括本地 BS 接入向量  $\gamma$ 、区块管理者和验证节点的带宽资源分配向量  $\mathcal{W}_v$ ，以及区块管理者和验证节点计算资源分配向量  $f_v$ 。因此，在决策时刻  $t$ ，优化第  $k$  个 BS 效用的动作空间表示为

$$\mathcal{A}_{k,t}^c = [\alpha, \mathcal{W}_c] \quad (18)$$

因此，优化  $K$  个 BS 效用的动作空间表示为

$$\mathcal{A}_i^c = \{\mathcal{A}_{1,t}^c, \dots, \mathcal{A}_{k,t}^c, \dots, \mathcal{A}_{K,t}^c\} \quad (19)$$

此外，优化区块管理者  $p$  效用的动作空间表示为

$$\mathcal{A}_i^v = [\gamma, \mathcal{W}_v, f_v] \quad (20)$$

3) 奖励函数。在双层 PPO 算法执行过程中，需要检验是否满足 C1~C7，因此，设计如下即时奖励函数  $r_t$

$$r_t = \begin{cases} r_t^c + r_t^v, & \text{满足 C1} \sim \text{C7} \\ 0, & \text{不满足 C1} \sim \text{C7} \end{cases} \quad (21)$$

其中， $r_t^c = \sum_{k \in \mathcal{K}} U_{BS_k}$ ， $r_t^v = U_p^v$ 。如果限制条件 C1~C7 不能被满足，则表明当前优化决策的性能较差。在这种情况下，将即时奖励设置为 0，以避免无效决策。

## 2.2 双层 PPO 算法原理

考虑状态空间和动作空间的高维数据，本文提出的双层 PPO 算法采用深度神经网络结合强化学习解决构造的双层马尔可夫决策过程问题。双层

PPO 算法通过构建策略优化模型，寻找最优神经网络参数  $\theta^*$ ，获得最优决策变量，最大化如式(22)所示的平均回报的目标。

$$R(\theta) = \mathbb{E}_{\zeta \sim \pi_{\theta}(c|s)} \left[ \sum_{t=1}^T \ell r_t \right] \quad (22)$$

其中， $0 \leq \ell \leq 1$  表示折扣因子； $\mathbb{E}$  表示根据转移序列  $\zeta$  的随机采样，在给定策略  $\pi_{\theta}$  和状态  $s$  的情况下，计算即时奖励的期望， $\zeta$  表示对应时刻  $t$  的状态和动作的转移序列，即  $\zeta = \{\mathcal{S}_1^c, \mathcal{S}_1^v, \mathcal{A}_1^c, \mathcal{A}_1^v, \dots, \mathcal{S}_t^c, \mathcal{S}_t^v, \mathcal{A}_t^c, \mathcal{A}_t^v\}$ 。PPO 算法是一种基于策略梯度和信任区间的强化学习算法。PPO 的策略优化模型包含 Actor-Critic 这 2 个神经网络。其中，Actor 网络接收当前状态并输出对应决策，Critic 网络对当前状态和 Actor 网络输出的决策进行评估，并反馈给 Actor 网络。Actor-Critic 网络通过策略梯度方法进行迭代训练，以寻找最优的网络参数  $\theta^*$ ，进而优化策略。应用信任区间方法将网络参数更新幅度限制在置信区间范围内，以便提高对复杂环境的适应性，以及收敛的稳定性。

在双层 PPO 算法训练过程中， $K$  个 BS 采用多代理 PPO 算法。此外，区块管理者  $p$  采用单代理 PPO 算法。每个 BS 和区块管理者  $p$  从边缘孪生模型中获取初始状态  $\mathcal{S}_1^c$  和  $\mathcal{S}_1^v$  生成随机策略。通过与边缘孪生模型交互，获取新的状态和动作转移序列。同时，调节各自的 Actor-Critic 网络参数  $\theta_k^a$ 、 $\theta_k^c$ 、 $\theta_p^a$  和  $\theta_p^c$ ， $k \in \mathcal{K}$ ，以得到最优策略。在  $K$  个 BS 的多代理 PPO 算法中，第  $k$  个 BS 的 Actor 网络参数  $\theta_k^a$  更新过程的损失函数表示为

$$L(\theta_k^a) = \min \left( \beta_t(\theta_k^a) A_t, \text{clip}(\beta_t(\theta_k^a), 1-\epsilon, 1+\epsilon) A_t \right) \quad (23)$$

其中， $\beta_t(\theta_k^a)$  表示网络参数更新幅度； $A_t$  表示优势函数，反映从当前网络参数生成的决策  $\mathcal{A}_{k,t}^c$  相对于其他可行决策的优势值；参数  $\epsilon \in (0,1)$  确定了上界和下界（即  $(1-\epsilon, 1+\epsilon)$ ），即 PPO 算法的置信区间； $\text{clip}(\cdot)$  函数将  $\beta_t(\theta_k^a)$  限制在该置信区间范围内，保证 Actor 网络对复杂环境的适应性以及收敛的稳定性。 $\beta_t(\theta_k^a)$  定义为

$$\beta_t(\theta_k^a) := \frac{\pi_{\theta_k^a}(\mathcal{A}_{k,t}^c | \mathcal{S}_{k,t}^c)}{\pi_{\theta_k^a, \text{old}}(\mathcal{A}_{k,t}^c | \mathcal{S}_{k,t}^c)} \quad (24)$$

其中， $\theta_k^a$  表示当前更新的网络参数； $\theta_k^a, \text{old}$  表示更

新前的网络参数；如果生成的决策  $\mathcal{A}_{k,t}^c$  获得更好的预期回报，则  $A_t > 0$ ，否则  $A_t < 0$ 。优势函数  $A_t$  定义为

$$A_t = \delta_t^c + \kappa \delta_{t+1}^c + \dots + \kappa^{T-t+1} \delta_{T-1}^c \quad (25)$$

其中， $\kappa \in [0,1]$  表示折扣因子， $\delta_t^c$  表示单步时间误差，定义为

$$\delta_t^c = r_{t+1}^c + \kappa \chi(\mathcal{S}_{t+1}^c) - \chi(\mathcal{S}_t^c) \quad (26)$$

其中， $\chi(\cdot)$  表示 Critic 网络对决策  $\mathcal{A}_{k,t}^c$  的估计奖励， $r_{t+1}^c + \kappa \chi(\mathcal{S}_{t+1}^c)$  表示决策  $\mathcal{A}_{k,t}^c$  对应的即时奖励  $r_{t+1}^c$  与 Critic 网络对决策  $\mathcal{A}_{k,t}^c$  的估计奖励之和。同时，Critic 网络将  $\delta_t^c$  的均方差作为损失函数，通过更新网络参数  $\theta_k^c$ ，以最小化损失函数，从而使 Critic 网络的估计奖励  $\chi(\cdot)$  更加准确。Actor-Critic 网络均使用自适应学习率优化算法进行随机梯度迭代训练，从而得到最优网络参数  $\theta_k^{a*}$  和  $\theta_k^{c*}$ 。

与  $K$  个 BS 的多代理 PPO 算法不同，区块管理者  $p$  采用单代理 PPO 算法，其中 Actor 网络参数  $\theta_p^a$  更新过程和第  $k$  个 BS 的 Actor 网络参数  $\theta_k^a$  更新过程一样。在第  $t$  时刻，Actor 网络输出的决策为  $\mathcal{A}_t^v$ 。定义单步时间误差  $\delta_t^v$  为

$$\delta_t^v = r_{t+1}^v + \kappa \chi(\mathcal{S}_{t+1}^v) - \chi(\mathcal{S}_t^v) \quad (27)$$

其中， $\chi(\cdot)$  表示 Critic 网络对决策  $\mathcal{A}_t^v$  的估计奖励， $r_{t+1}^v + \kappa \chi(\mathcal{S}_{t+1}^v)$  表示决策  $\mathcal{A}_t^v$  对应的即时奖励  $r_{t+1}^v$  与 Critic 网络对决策  $\mathcal{A}_t^v$  的估计奖励之和。同时，Critic 网络将  $\delta_t^v$  的均方差作为损失函数，通过更新网络参数  $\theta_p^c$ ，以最小化损失函数，从而使 Critic 网络的估计奖励  $\chi(\cdot)$  更加准确，以得到最优网络参数  $\theta_p^{a*}$  和  $\theta_p^{c*}$ 。算法 1 给出了本文提出的  $K$  个 BS 多代理 PPO 算法，算法 2 给出了本文提出的区块管理者  $p$  单代理 PPO 算法。

**算法 1** 基于数字孪生边缘网络的  $K$  个 BS 多代理 PPO 算法

输入 状态  $\mathcal{S}_1^c, \mathcal{S}_2^c, \dots, \mathcal{S}_T^c$

输出 对应时刻的  $K$  个 BS 资源分配决策  $\mathcal{A}_1^c, \mathcal{A}_2^c, \dots, \mathcal{A}_T^c$

- 1) 当 episode=1,2,...时，开始循环 1
- 2) 当  $t=1,2,\dots,T$  时，开始循环 2
- 3)  $K$  个 BS 从边缘孪生模型分别获取当前时刻

$t$  的状态  $\mathcal{S}_{k,t}^c$ ， $k \in \mathcal{K}$ ，并输入各自的 Actor 网络得到决策  $\mathcal{A}_{k,t}^c$ ， $k \in \mathcal{K}$

4)  $K$  个 BS 在各自边缘孪生模型中执行决策，并得到各自的即时报酬与下一时刻的状态

5) 如果  $t \% b=0$

6)  $K$  个 BS 采用随机梯度方法更新网络参数  $\theta_k^a$  和  $\theta_k^c$ ， $k \in \mathcal{K}$ ，以最小化损失函数式(23)和式(26)

7) 结束循环 2

8) 结束循环 1

**算法 2** 基于数字孪生边缘网络的区块管理者  $p$  单代理 PPO 算法

输入 状态  $\mathcal{S}_1^v, \mathcal{S}_2^v, \dots, \mathcal{S}_T^v$

输出 对应时刻区块管理者  $p$  的资源分配决策  $\mathcal{A}_t^v$

1) 当 episode=1,2,...时，开始循环 1

2) 对于  $t=1,2,\dots,T$ ，开始循环 2

3) 区块管理者  $p$  从边缘孪生模型中获取当前时刻  $t$  的状态  $\mathcal{S}_t^v$ ，并输入 Actor 网络得到决策  $\mathcal{A}_t^v$

4) 区块管理者  $p$  在边缘孪生模型中执行决策并得到即时报酬与下一时刻状态

5) 如果  $t \% b=0$

6) 采用随机梯度方法更新网络参数  $\theta_p^a$  和  $\theta_p^c$

7) 结束循环 2

8) 结束循环 1

### 3 仿真分析

本节通过实验验证提出的基于数字孪生边缘网络的双层 PPO 算法的性能。考虑 BS 个数为  $K=50$  个。在第  $k$  个 BS 的覆盖范围内，随机分布  $U=10$  个智能终端簇，以及随机分布  $J_k$  个 AP，且  $J_k \in [10, 15]$ 。每个簇内均匀分布  $N_k^u$  个智能终端，且  $N_k^u \in [5, 10]$ 。在 PPO 算法实现过程中，设置 Actor 网络由一个输入层、2 个隐藏层和一个输出层构成，Critic 网络由一个输入层、一个隐藏层和一个输出层构成。Actor 网络和 Critic 网络的训练参数，以及本地数据共享区块链和全局孪生数据共享区块链的仿真参数如表 1 所示。考虑边缘孪生模型与物理智能终端和 AP 的映射误差，导致孪生模型仿真的状态与物理智能终端和 AP 的状态也存在轻微误差。为了验证本文提出的基于数字孪生边缘网络的双层 PPO 算法在适应映射误差方面的优越性，将基于数字孪生边缘网络的 DQN 和 DDQN 作为对比算法。

表 1 仿真参数	
参数	取值
子信道带宽/MHz	$W_0=30$
子信道个数	50
计算负载/MHz	$\alpha_l=1$
AP $j_k$ 信誉值	$\vartheta_{j_k}=[0.1, 0.8]$
交易到达速率	$\lambda=30$
交易大小/KB	$ m_{j_k} =[1, 30]$
区块大小/MB	$S_{B,q}=[1, 10]$
确认阈值	$CW_{th}=600$
验证节点个数	$V=21$
神经网络训练批次	$b=60$

图 3 对比了 3 种算法在本地数据共享链中的 BS 效用。从图 3 中可以看出，随着训练次数增加，3 种算法的 BS 效用先增加，然后收敛到固定值，证明了本文提出的基于数字孪生边缘网络的 PPO 算法可以通过调节智能终端簇和 AP 的关联变量，以及调节 AP 的带宽分配变量，实现最大化 BS 效用的目标。在没有映射误差的情况下，本文算法在训练 700 次时就可收敛，而对比算法 DQN 和 DDQN 分别训练 500 次和 1 500 次才收敛，且获得的 BS 效用均比提出方案的效用小，证明了对比算法容易陷入局部最优。同时，DDQN 获得的 BS 效用比 DQN 的 BS 效用大。此外，在有映射误差的情况下，本文算法可以较好地适应映射误差，并且在训练 1 000 次时可以收敛，收敛的效用接近无映射误差情况下的效用。然而，对比算法 DQN 和 DDQN 则无法较好地适应映射误差，且实现的效用比无映射误差情况下 DQN 和 DDQN 效用小。

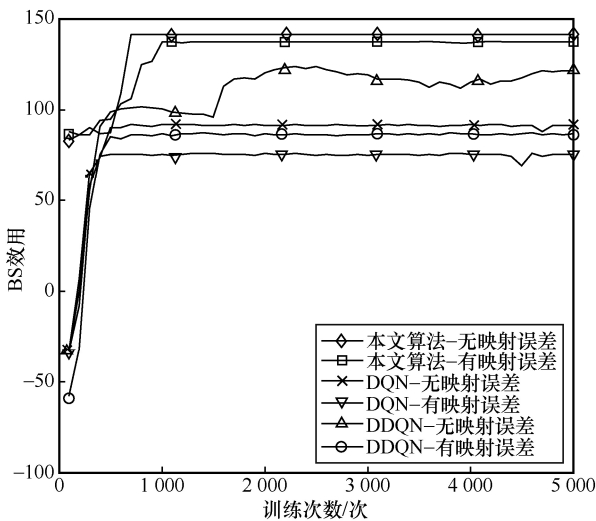


图 3 3 种算法在本地数据共享链中的 BS 效用

图 4 给出了本地数据共享链的交易吞吐量与交易大小的关系。从图 4 中可以看出，随着交易大小增加，3 种算法实现的交易吞吐量均减小。这是因为随着交易大小增加，AP 能够缓存的交易个数减少。同时，交易的多播时延和确认时延均会增加，从而导致交易吞吐量减小。此外，在没有映射误差的情况下，本文算法获得的交易吞吐量对比算法 DQN 和 DDQN 的交易吞吐量大，且 DDQN 实现的交易吞吐量比 DQN 的交易吞吐量大。在有映射误差的情况下，本文算法的交易吞吐量略小于无映射误差情况下的交易吞吐量，但大于对比算法实现的交易吞吐量，证明了本文算法在本地数据共享链中适应映射误差方面的优越性。

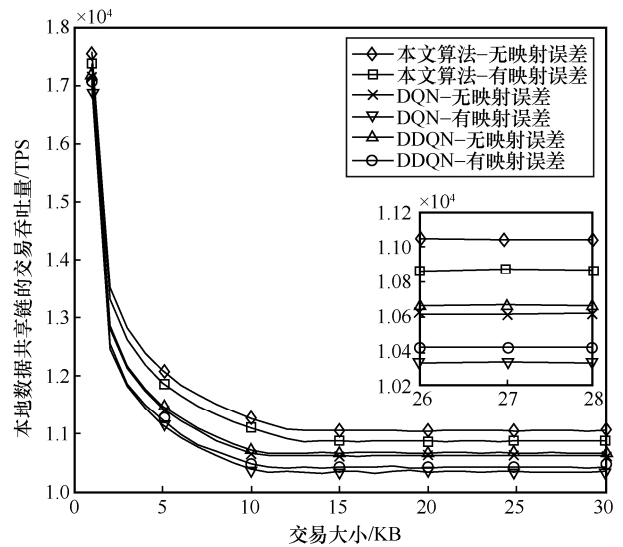


图 4 本地数据共享链的交易吞吐量与交易大小的关系

图 5 对比了 3 种算法在全局孪生数据共享链中的区块验证节点效用。从图 5 中可以看出，随着训练次数增加，3 种算法实现的区块管理者的效用先增加，然后收敛到固定值。因此，证明了本文算法通过调节本地 BS 接入向量、区块管理者和验证节点的带宽资源分配向量，以及区块管理者和验证节点计算资源分配向量，可以实现最大化区块管理者效用的目标。在没有映射误差的情况下，虽然本文算法的收敛速率对比算法 DQN 和 DDQN 收敛速率稍慢，但实现的区块验证节点的效用对比算法的效用明显增加。同时，DDQN 实现的区块验证节点的效用比 DQN 的效用大。同样，在有映射误差的情况下，本文算法可以较好地适应映射误差，获得的区块验证节点效用接近无映射误差情况下的区

块验证节点效用。然而，对比算法 DQN 和 DDQN 则无法较好地适应映射误差，并且与无映射误差情况相比，实现的效用明显减小。

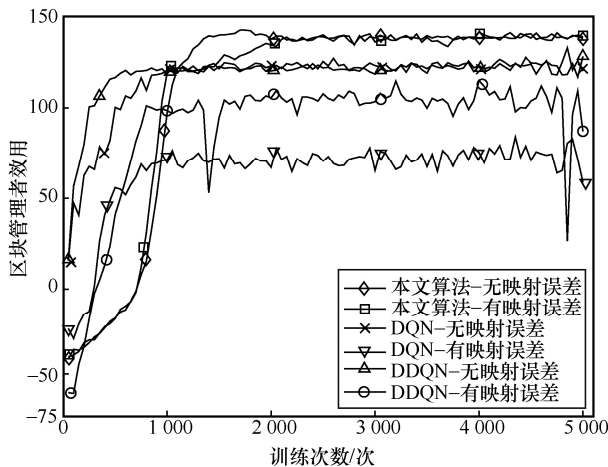


图 5 3 种算法在全局孪生数据共享链中的区块验证节点效用

图 6 给出了全局孪生数据共享链的交易吞吐量与区块大小的关系。从图 6 中可以看出，随着区块大小增加，3 种算法实现的交易吞吐量先增加，然后趋于平稳。这是因为增加区块大小，区块中打包的交易个数也随着增加。然而，随着区块大小增加，区块多播传输时延和区块确认时延也会增加。此外，在没有映射误差的情况下，本文算法获得的全局孪生数据共享链的交易吞吐量对比算法 DQN 和 DDQN 获得的交易吞吐量大。在有映射误差的情况下，本文算法的交易吞吐量比无映射误差情况的交易吞吐量稍微减少，但优于对比算法实现的交易吞吐量，证明了本文算法在全局孪生数据共享链中适应映射误差方面的优越性。

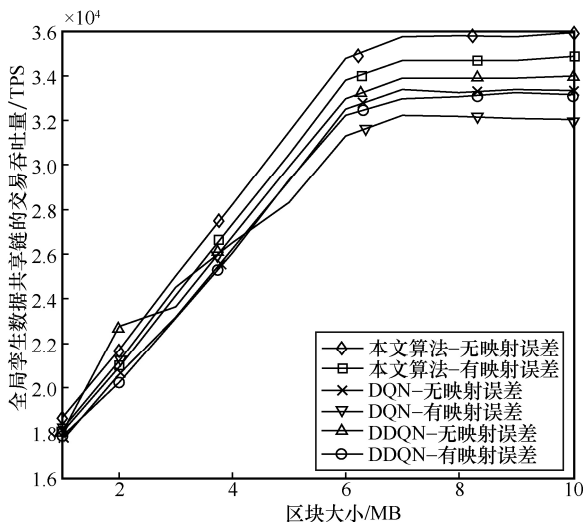


图 6 全局孪生数据共享链的交易吞吐量与区块大小的关系

## 4 结束语

本文采用区块链分片技术，提出一种数字孪生边缘网络本地数据和孪生数据安全共享机制。考虑数字孪生边缘网络环境的动态时变特性，构造区块链分片的资源自适应优化问题，通过优化本地数据共享链和全局孪生数据共享链的资源分配决策，实现最大化区块链分片的交易吞吐量的目标。考虑实际边缘孪生模型与物理网络的映射误差，提出基于数字孪生边缘网络的双层 PPO 资源自适应优化算法，求解区块链分片的资源自适应优化问题。仿真实验表明，本文算法可以显著提高区块链分片的交易吞吐量，且能较好地适应映射误差。

### 参考文献:

- [1] 张平, 牛凯, 田辉, 等. 6G 移动通信技术展望[J]. 通信学报, 2019, 40(1): 141-148.  
ZHANG P, NIU K, TIAN H, et al. Technology prospect of 6G mobile communications[J]. Journal on Communications, 2019, 40(1): 141-148.
- [2] 林奕琳, 陈思柏, 单雨威, 等. 6G 网络潜在关键技术研究综述[J]. 移动通信, 2021, 45(4): 120-127.  
LIN Y L, CHEN S B, SHAN Y W, et al. A survey of potential technologies in 6G network[J]. Mobile Communications, 2021, 45(4): 120-127.
- [3] 孙耀华, 王则予, 袁硕, 等. 智能内生 6G 网络: 架构、用例和挑战[J]. 电子技术应用, 2021, 47(3): 8-13, 17.  
SUN Y H, WANG Z Y, YUAN S, et al. The sixth-generation mobile communication network with endogenous intelligence: architectures, use cases and challenges[J]. Application of Electronic Technique, 2021, 47(3): 8-13, 17.
- [4] 王敬宇, 周铨, 张蕾, 等. 知识定义的意图网络自治[J]. 电信科学, 2021, 37(9): 1-13.  
WANG J Y, ZHOU C, ZHANG L, et al. Knowledge-defined intent-based network autonomy[J]. Telecommunications Science, 2021, 37(9): 1-13.
- [5] 张平, 许晓东, 韩书君, 等. 智简无线网络赋能行业应用[J]. 北京邮电大学学报, 2020, 43(6): 1-9.  
ZHANG P, XU X D, HAN S J, et al. Entropy reduced mobile networks empowering industrial applications[J]. Journal of Beijing University of Posts and Telecommunications, 2020, 43(6): 1-9.
- [6] 孙滔, 周铨, 段晓东, 等. 数字孪生网络(DTN): 概念、架构及关键技术[J]. 自动化学报, 2021, 47(3): 569-582.  
SUN T, ZHOU C, DUAN X D, et al. Digital twin network(DTN): concepts, architecture, and key technologies[J]. Acta Automatica Sinica, 2021, 47(3): 569-582.
- [7] GLAESSGEN E, STARGEL D. The digital twin paradigm for future NASA and U.S. air force vehicles[C]//Proceedings of 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference. Reston: AIAA, 2012: 1818.
- [8] 陶飞, 张萌, 程江峰, 等. 数字孪生车间: 一种未来车间运行新模

- 式[J]. 计算机集成制造系统, 2017, 23(1): 1-9.
- TAO F, ZHANG M, CHENG J F, et al. Digital twin workshop: a new paradigm for future workshop[J]. Computer Integrated Manufacturing Systems, 2017, 23(1): 1-9.
- [9] CROATTI A, GABELLINI M, MONTAGNA S, et al. On the integration of agents and digital twins in healthcare[J]. Journal of Medical Systems, 2020, 44(9): 161.
- [10] NGUYEN H X, TRESTIAN R, TO D, et al. Digital twin for 5G and beyond[J]. IEEE Communications Magazine, 2021, 59(2): 10-15.
- [11] YU Q, REN J, FU Y J, et al. Cybertwin: an origin of next generation network architecture[J]. IEEE Wireless Communications, 2019, 26(6): 111-117.
- [12] LU Y L, HUANG X H, ZHANG K, et al. Communication-efficient federated learning for digital twin edge networks in industrial IoT[J]. IEEE Transactions on Industrial Informatics, 2021, 17(8): 5709-5718.
- [13] WU Y W, ZHANG K, ZHANG Y. Digital twin networks: a survey[J]. IEEE Internet of Things Journal, 2021, 8(18): 13789-13804.
- [14] SUN W, ZHANG H B, WANG R, et al. Reducing offloading latency for digital twin edge networks in 6G[J]. IEEE Transactions on Vehicular Technology, 2020, 69(10): 12240-12251.
- [15] DAI Y Y, ZHANG K, MAHARJAN S, et al. Deep reinforcement learning for stochastic computation offloading in digital twin networks[J]. IEEE Transactions on Industrial Informatics, 2021, 17(7): 4968-4977.
- [16] JIANG L, XIE S L, MAHARJAN S, et al. Blockchain empowered wireless power transfer for green and secure Internet of things[J]. IEEE Network, 2019, 33(6): 164-171.
- [17] JIANG L, XIE S L, MAHARJAN S, et al. Joint transaction relaying and block verification optimization for blockchain empowered D2D communication[J]. IEEE Transactions on Vehicular Technology, 2020, 69(1): 828-841.
- [18] LU Y L, HUANG X H, ZHANG K, et al. Low-latency federated learning and blockchain for edge association in digital twin empowered 6G networks[J]. IEEE Transactions on Industrial Informatics, 2021, 17(7): 5098-5107.
- [19] YAQOOB I, SALAH K, UDDIN M, et al. Blockchain for digital twins: recent advances and future research challenges[J]. IEEE Network, 2020, 34(5): 290-298.
- [20] JIANG L, ZHENG H, TIAN H, et al. Cooperative federated learning and model update verification in blockchain-empowered digital twin edge networks[J]. IEEE Internet of Things Journal, 2022, 9(13): 11154-11167.
- [21] LIU Y Z, LIU J W, WU Q H, et al. SSHC: a secure and scalable hybrid consensus protocol for sharding blockchains with a formal security framework[J]. IEEE Transactions on Dependable and Secure Computing, 2022, 19(3): 2070-2088.
- [22] CAO B, LI M Y, ZHANG L, et al. How does CSMA/CA affect the performance and security in wireless blockchain networks[J]. IEEE Transactions on Industrial Informatics, 2020, 16(6): 4270-4280.
- [23] YUAN Y, WANG F Y. Towards blockchain-based intelligent transportation systems[C]//Proceedings of 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). Piscataway: IEEE Press, 2016: 2663-2668.
- [24] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[J]. arXiv Preprint, arXiv: 1707.06347, 2017.
- [25] CRONRATH C, ADERIANI A R, LENNARTSON B. Enhancing digital twins through reinforcement learning[C]//Proceedings of 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE). Piscataway: IEEE Press, 2019: 293-298.

#### [作者简介]



蒋丽（1986-），女，四川简阳人，博士，广东工业大学讲师，主要研究方向为B5G/6G网络、物联网、移动边缘计算、数字孪生、边缘智能和区块链等。

谢胜利（1956-），男，湖北荆州人，博士，广东工业大学教授，主要研究方向为智能信息处理、无线通信与网络、多媒体传输、射频识别等。

田辉（1963-），女，河南郑州人，博士，北京邮电大学教授，主要研究方向为无线网络资源管理、移动边缘计算、边缘智能、无线自组织网络、移动社交网络等。