

基于深度强化学习的以太坊 MEV 交易防护与交易排序优化

严彦胜, 李 京

(中国科学技术大学 计算机科学与技术学院, 安徽 合肥 230026)

摘要: 以太坊交易排序中的最大可提取价值 (MEV) 问题, 使恶意交易者通过操控交易顺序获利, 破坏了网络公平性并推高 Gas 费用。为抑制 MEV 行为和优化公平性, 同时提升系统效率, 提出了一种基于深度 Q 网络 (DQN) 的交易排序优化方法。通过设计合理的状态空间、动作空间与奖励函数, 智能体能够自主学习最优排序策略。结合 Geth 私有链、Flashbots MEV-Explore 与 Ethereum Mempool 数据, 系统性验证了该方法的有效性。实验结果表明, DQN 排序策略可将 MEV 提取率降至 13% 以下, 平均 Gas 费用较传统策略降低约 33.1%, 公平性指数提升至 0.78, 显著优于现有方法。同时搭建的闭环实验系统为区块链交易公平性和效率优化提供了可行方案。

关键词: 以太坊; 最大可提取价值 (MEV); 交易排序优化; 深度 Q 网络 (DQN)

中图分类号: TP393.0; TP309

文献标识码: A

DOI: 10.19358/j.issn.2097-1788.2025.07.004

引用格式: 严彦胜, 李京. 基于深度强化学习的以太坊 MEV 交易防护与交易排序优化 [J]. 网络安全与数据治理, 2025, 44(7): 20-26.

Ethereum MEV transaction protection and transaction ordering optimization based on deep reinforcement learning

Yan Yansheng, Li Jing

(School of Computer Science and Technology, University of Science and Technology of China, Hefei 230026, China)

Abstract: The problem of Maximal Extractable Value (MEV) in Ethereum transaction ordering allows malicious actors to profit by manipulating transaction sequences, undermining network fairness and increasing Gas fees. To suppress MEV behavior and optimize fairness while enhancing system efficiency, this paper proposes a transaction ordering optimization method based on Deep Q-Network (DQN). By designing appropriate state space, action space, and reward function, the agent can autonomously learn optimal ordering strategies. The effectiveness of the proposed method is systematically validated using a Geth private chain, along with Flashbots MEV-Explore and Ethereum Mempool data. Experimental results show that the DQN-based ordering strategy reduces the MEV extraction rate to below 13%, decreases average Gas fees by about 33.1% compared to traditional strategies, and raises the fairness index to 0.78, significantly outperforming existing methods. The closed-loop experimental system built in this paper provides a feasible solution for optimizing fairness and efficiency in blockchain transaction ordering.

Key words: Ethereum; Maximal Extractable Value (MEV); transaction ordering optimization; Deep Q-Network (DQN)

0 引言

随着区块链技术的广泛应用, 尤其是以太坊在去中心化金融 (DeFi) 领域的快速发展, 最大可提取价值 (Maximal Extractable Value, MEV) 交易问题逐渐成为影响网络公平性、安全性和效率的重要挑战^[1]。MEV 由矿工或排序者通过操控交易顺序获得额外利润, 表现形式包括前置攻击 (Front-Running)、三明治攻击 (Sandwich Attack) 等^[2]。这些交易模式不仅对普通用户造成经济损失, 还推高整体交易成本, 加剧网络拥堵, 破坏区块链

生态的稳定性与公正性^[3]。

MEV 的产生主要源于区块链交易顺序的可操控性、交易信息的公开性、Gas 竞价机制以及 DeFi 交易的复杂性。现有的交易排序方法主要包括先来先服务 (FIFO) 排序和 Gas 竞价排序。FIFO 排序虽然具有较高的公平性, 但缺乏 Gas 费用优化, 无法有效防范 MEV 交易^[4]。Gas 竞价排序根据交易支付的 Gas 费用决定执行顺序, 导致 Gas 竞价竞争加剧, 使得 MEV 交易通过支付高额 Gas 费用获得优先执行权, 从而进一步破坏交易公平性^[5]。

与现有采用启发式排序策略或基于静态规则的排序机制相比,本文方法引入强化学习框架,具有更强的自适应能力,能够在不同交易负载和攻击场景下动态调整排序策略。同时,相较于现有基于强化学习的研究^[6-8]主要聚焦于合约交互或交易分类问题,本文首次将深度Q网络(Deep Q-Network, DQN)应用于交易排序优化,并构建专门的状态-动作空间及奖励函数,直接面向MEV抑制目标进行学习。

本文的主要贡献如下:

(1) 提出基于DQN的交易排序优化方法,动态调整交易执行顺序,以降低MEV交易的影响,提高交易公平性。

(2) 优化Gas费用,通过强化学习减少Gas竞价排序带来的交易成本膨胀问题。

(3) 构建完整的实验框架,在Geth(Go-Ethereum)本地私有链环境下复现真实交易处理流程,并结合Flashbots MEV-Explore数据集进行实验验证。

实验结果表明,所提出的DQN排序策略相比传统方法将MEV提取率降低至13%以下,平均Gas费用较基准策略降低约33.1%,并显著提升低Gas交易的执行公平性,公平性指数达到0.78,远高于传统策略。

1 相关研究

随着MEV问题的日益严重,区块链交易排序机制成为影响交易执行顺序、公平性及系统效率的关键因素。为解决这一问题,学者们提出了多种技术路线,主要包括透明化交易池、去中心化交易所(DEX)优化以及应用强化学习等。

1.1 透明化与信息公开

为了解决MEV问题,研究者提出了透明化交易池(Mempool)的概念,旨在减少操控交易顺序的空间。例如,Flashbots推出的透明Mempool系统允许矿工和交易者公开访问交易池的数据,从而减少前置交易等MEV行为的发生^[8]。透明Mempool通过确保所有参与者在相同条件下访问交易池数据,显著减少了恶意操控交易顺序的机会。

1.2 去中心化交易所优化

去中心化交易所是解决MEV问题的另一个重要方向。一些研究提出,在DEX中引入时间加权机制或调整交易执行优先级,可以有效减少MEV的影响。例如,在Uniswap等去中心化交易所中,通过优化交易匹配算法,使得所有交易都能够公平地获得执行机会,从而减少了MEV的提取^[9-11]。通过优化DEX的交易匹配机制,避免了恶意用户利用市场波动和订单优先级进行操控,减少了MEV攻击对普通用户的影响。

1.3 强化学习在区块链优化中的应用

传统Q-learning依赖离散状态和动作空间,难以处理交易排序优化中的高维特征,且存储和更新Q表的计算开销较大,难以适应复杂的交易环境。DQN结合深度学习与强化学习,能够在高维状态空间下优化决策。DQN采用神经网络逼近Q值函数,并利用经验回放(Experience Replay)和目标网络(Target Network)等优化策略,提高交易排序优化的稳定性和泛化能力。

1.4 小结

针对MEV问题,现有研究主要集中在透明化交易池、去中心化交易所优化和强化学习方法等方面。透明化交易池和去中心化交易所优化通过提高交易数据的公开性和优化交易匹配机制,减少了MEV攻击的影响。然而,关于MEV防护的强化学习方法仍需进一步研究。本文提出的基于DQN的交易排序优化方法,通过强化学习动态优化交易排序,能够有效减少MEV交易、提升交易公平性并优化Gas费用,为解决MEV问题提供了新的技术方案。

2 基于DQN的交易排序优化方法

本节介绍基于DQN的以太坊交易排序优化方法,排序流程图如图1所示。该方法通过构建状态空间、动作空间及奖励函数,将交易排序建模为强化学习任务,使智能代理能够学习最优交易排序策略,以减少MEV交易影响,提高交易公平性,并优化Gas费用。

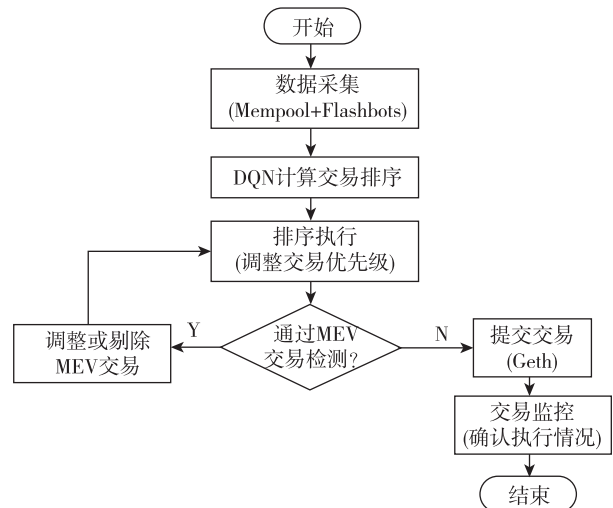


图1 基于DQN算法排序流程图

2.1 交易排序优化建模

以太坊的交易排序问题可以视作一个强化学习任务,其中智能代理通过与环境交互,学习最优交易排序策略,以提高交易公平性并降低MEV交易影响。本研究采用DQN作为强化学习模型,以Mempool状态作为输入,学

习优化交易排序的方法。

在该优化框架中,环境包括交易池中的待处理交易、Gas 费用、时间戳以及潜在的 MEV 交易特征。智能代理观察环境状态,并选择一个交易排序策略,随后计算相应的奖励值,并更新 Q 网络参数,以优化排序决策。

2.2 强化学习框架

DQN 交易排序优化方法的核心包括状态空间、动作空间和奖励函数的设计。

2.2.1 状态空间

交易池中的状态由多个特征构成,其中 T_i 表示当前交易池中待处理交易集合,每笔交易包含其唯一标识; T_i 表示每笔交易的到达时间戳,用于反映先后顺序; G_i 表示交易的 Gas Price (即单位计算成本),是传统排序方法的核心依据; M_i 表示 MEV 风险评分,衡量当前交易被前置、三明治攻击等利用的可能性; B_i 表示当前区块的剩余 Gas 容量,影响可打包交易数; P_i 表示该类交易在历史排序中的平均执行成功率,反映其执行优先级的长期表现。状态空间数学表示如下:

$$S_i = \{T_i, T_i, G_i, M_i, B_i, P_i\} \quad (1)$$

为提升模型泛化能力,所有特征将进行归一化处理,并构建为固定维度的张量输入 Q 网络。通过对交易池状态的高维建模,模型可捕捉交易间的关联性与排序决策对系统性能的影响。

其中风险评分 M_i 本文参考 Flashbots 提出的相关特征因子,构建如下的风险评分函数:

$$M_i = \lambda_1 \cdot G + \lambda_2 \cdot S + \lambda_3 \cdot H \quad (2)$$

其中, G 为该交易相较于历史平均水平的 Gas 价格突增指标, S 为交易执行对链上资产价格的滑点影响因子, H 为历史上该地址被前置攻击的命中率。

2.2.2 动作空间

在本研究的交易排序优化任务中, DQN 智能体的动作空间被定义为可供选择的若干交易排序策略。每一类动作对应一种具体的排序规则,结合以太坊主流排序方式和强化学习优化策略,本文将动作空间 A 设定为如下三种核心操作:

(1) 动作 A_1 : 维持当前顺序 (FIFO)。该操作代表以太坊默认的先来先服务排序方式,按交易到达时间执行,无需额外计算。该方式具有公平性强的优势,但在 MEV 抑制和资源优化方面存在不足,本文将作为性能基准进行对照。

$$A_1 = \text{FIFO}(T) \text{ where } T = \{t_1, t_2, \dots, t_n\} \quad (3)$$

其中, T 是交易池中的交易,按交易到达时间顺序 $t_1, t_2, \dots, t_n$ 进行排序。

(2) 动作 A_2 : 基于 Gas 费用排序。该操作模拟主网上矿工打包交易的常见逻辑,即按照交易的 Gas Price 从高到低进行排序。此方法鼓励用户通过支付高额手续费获取优先执行权,虽然提高了确认效率,但容易造成交易排序不公并放大 MEV 行为。

$$A_2 = \text{GasSort}(T) \text{ where } T = \{t_1, t_2, \dots, t_n\}, \text{ sort by GasPrice}(t) \quad (4)$$

其中, $\text{GasPrice}(t)$ 为交易 t 的 Gas 费用,按 GasPrice 排序。

(3) 动作 A_3 : DQN 智能排序。该操作由训练完成的 DQN 策略网络 Q_θ 输出,通过评估当前交易池状态 S_i 中所有可选排序动作的 Q 值,选择期望回报最大的排序方案。该策略综合考虑交易公平性、Gas 成本与 MEV 风险,是本研究提出的主要优化方案。

$$A_3 = \arg \max_{A_i} Q_\theta(S_i, A_i) \text{ where } S_i = \{T_i, T_i, \dots\} \quad (5)$$

其中, S_i 是当前交易池的状态, $Q_\theta(S_i, A_i)$ 是策略网络根据状态 S_i 和动作 A_i 计算得到的 Q 值,智能体选择 Q 值最大的动作进行排序优化。

2.2.3 奖励函数

为了衡量排序策略的优劣,设计了综合考虑 F (交易公平性)、 E (交易执行效率)、 M (MEV 交易的减少) 和 N (区块链网络健康性) 的奖励函数:

$$R_i = \alpha \cdot F + \beta \cdot E - \gamma \cdot M + \delta \cdot N \quad (6)$$

其中, α 为公平性权重系数,用于控制 DQN 智能体在排序决策中对公平性指标的重视程度。 α 越大,模型将更倾向于提高低 Gas 费用交易的执行概率,增强普通用户的交易公平性保障。

β 为执行效率权重系数,用于衡量优化后系统的吞吐量 (TPS),主要通过减少交易确认延迟和提升 TPS 实现。 β 越大,模型越注重提高交易执行速度与打包效率。

γ 为 MEV 抑制惩罚系数,用于对识别出的潜在 MEV 攻击行为进行惩罚。 γ 越大,智能体越倾向于规避 MEV 风险高的排序策略,从而有效抑制前置攻击与套利行为。

δ 为网络健康度奖励系数,用于度量优化排序策略对以太坊网络整体运行稳定性的积极影响。该项不仅反映了区块资源 (如 Gas 限额) 的合理使用情况,还考虑了网络中交易拥堵程度、Gas 价格波动范围及平均打包成功率等因素。具体而言,当排序策略在降低平均 Gas 价格剧烈波动、提升区块填充率 (Gas 利用率)、避免交易拥堵方面表现良好时,智能体将获得更高奖励;反之,若排序策略导致网络负载不均、Gas 消耗无效或区块利用不足,则该项贡献降低甚至为负,从而引导智能体做出更

“生态友好型”的排序选择。

2.3 DQN 训练过程

为了实现以太坊交易排序策略的自动优化,本文采用 DQN 模型对排序策略进行强化学习训练。DQN 能够在高维状态空间中进行决策学习,输出在当前交易池状态下的最优排序动作,从而动态调整交易顺序以抑制 MEV 行为并提升公平性。

(1) Q 值更新机制

DQN 的核心是学习状态-动作值函数 $Q(s, a)$, 其更新规则如下:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_t + \gamma \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t)] \quad (7)$$

其中, s_t 表示时间步 t 时的状态; a_t 表示在状态 s_t 下采取的动作; r_t 是智能体在 t 时刻执行动作后从环境获得的即时奖励; α 是学习率, 用于控制 Q 值更新的步长; γ 是折扣因子, 用于衡量未来奖励的影响; $\max_{a'} Q(s_{t+1}, a')$ 表示在下一个状态 s_{t+1} 中, 所有可能动作中的最大 Q 值。

(2) 模型结构与训练策略

DQN 模型采用神经网络近似 Q 函数, 输入为当前交易池状态特征 s_t , 输出为所有排序动作对应的 Q 值。为增强训练稳定性, 引入两项机制:

①经验回放: 将每轮交互获得的四元组 (s_t, a_t, r_t, s_{t+1}) 存储于经验池中, 训练阶段从中随机采样, 避免样本间相关性干扰;

②目标网络: 设置与主 Q 网络结构相同但更新滞后的目标网络 Q_0 , 生成训练目标值, 缓解训练震荡。

(3) 探索策略

在动作选择方面采用 ϵ -greedy 策略, 平衡探索与利用的关系:

$$a_t = \begin{cases} \arg \max Q(s_t, a), & \text{以概率 } 1 - \epsilon \text{ 选择} \\ \text{随机动作}, & \text{以概率 } \epsilon \text{ 选择} \end{cases} \quad (8)$$

其中 ϵ 逐轮衰减, 有助于模型在训练前期尝试更多策略, 后期趋于稳定收敛。

(4) 排序策略优化流程

为实现交易排序策略的智能优化, 本文基于 DQN 模型构建了完整的执行流程, 涵盖从环境初始化、状态感知到策略决策与更新的全过程, 以适应动态交易场景下的实时排序需求。

系统首先初始化本地以太坊交易池, 将 Mempool 中待打包交易作为环境输入, 提取交易的 Gas 费用、时间戳、发送者等特征构成状态向量 S_t 。DQN 模型在每一时间步根据当前状态计算动作价值函数 $Q(s_t,$

$a_t)$, 选择最优排序策略 a_t , 并对交易池中的交易执行排序调整。

排序动作执行后, 环境反馈即时奖励 r_t , 该奖励综合考虑 MEV 交易抑制、低 Gas 费用交易执行率、吞吐量变化等因素, 引导模型向公平性与效率兼顾的方向优化。随后, 智能体通过经验回放机制, 从历史交互中采样状态转移元组 (s_t, a_t, r_t, s_{t+1}) 用于更新 Q 网络, 降低训练过程中的数据相关性。为提高学习稳定性, 训练过程中采用了目标网络协助生成稳定 Q 值目标, 并定期与主网络同步参数。

通过持续迭代, DQN 模型能够自适应学习最优排序策略, 在动态链上环境中显著提升交易排序的公平性与系统效率。算法的伪代码如下:

算法 1 基于 DQN 的以太坊交易实时排序与执行

Input: 训练完成的策略网络 Q_0 , 当前交易池状态 S_t ;

Output: 优化后的交易排序列表 T'_t 并提交执行。

初始化 Flashbots 接口与 Geth 节点监听服务

while 系统运行中 do

 从 Mempool 获取最新交易状态 S_t , 构建状态向量, 包括 Gas 费用、交易时间戳、合约类型、MEV 风险评分;

 使用策略网络 Q_0 进行动作推理, 选择最优排序策略 A_t :

$$A_t \leftarrow \arg \max_a Q_0(A_t, a)$$

 根据排序动作 A_t , 执行交易排序优化:

$$T'_t \leftarrow \text{apply_sorting}(A_t, S_t)$$

 将优化排序后的交易序列 T'_t 提交至 Geth 节点或 Flashbots API;

 监听交易打包结果, 记录执行成功率, 确认时间与 Gas 消耗等反馈信息;

 将反馈数据送入策略评估模块, 用于后续模型微调或强化学习再训练;

end

3 实验设计与结果分析

本节围绕基于深度强化学习的交易排序系统展开实验验证, 旨在评估 DQN 排序策略在抑制 MEV、提升交易公平性与优化交易成本方面的实际效果。通过在本地私链环境中部署 Geth 节点, 并引入真实区块链交易数据集, 系统性测试排序优化前后的性能表现。

3.1 实验环境与工具配置

为确保交易排序优化方法的可实施性与实验复现性, 本文在本地搭建以太坊私有链测试平台, 并集成相关工具与数据资源进行实验验证。整体环境配置如下:

(1) 区块链节点平台

本研究采用 Geth 客户端构建以太坊本地私链, 模拟

真实网络中的交易打包与排序过程。系统设定出块间隔为 15 s, 区块 Gas 上限为 12 500 000, 支持模拟高负载交易场景。

(2) 交易数据来源

①Flashbots MEV-Explore 数据集: 提供历史 MEV 攻击样本 (如三明治攻击、前置交易等), 用于训练中构造 MEV 风险评分标签;

②Ethereum Mempool 数据集: 包含实时交易记录, 涵盖 Gas Price、时间戳、合约类型等关键信息, 用于构建强化学习状态空间。

(3) 智能体训练平台

排序策略采用深度 Q 网络建模, 集成经验回放、目标网络等模块以提升训练稳定性和泛化能力。训练过程使用 Python 编写, 主要依赖工具包括:

- ①PyTorch: 实现 DQN 模型的网络结构与参数更新;
- ②Web3.py: 与 Geth 节点交互, 实现交易状态监听与打包提交;
- ③Pandas /Numpy: 用于交易数据处理与特征提取。

通过上述实验环境配置, 本文得以模拟完整的以太坊交易流程, 并实现 DQN 强化学习模型在排序任务中的实际部署, 为后续结果分析提供稳定基础。

3.2 数据集介绍与预处理流程

本研究采用了 Flashbots MEV-Explore 数据集与 Ethereum Mempool 数据集, 用于评估基于 DQN 的交易排序优化方法。

Flashbots MEV-Explore 数据集记录了大量以太坊网络中的 MEV 交易案例, 包括交易类型、金额、Gas 费用及区块位置等关键信息, 有助于深入理解 MEV 的产生机制及其影响。

Ethereum Mempool 数据集则实时捕获了以太坊待处理交易的详细信息, 包括交易发送时间、交易量、Gas 价格和发送账户地址等, 提供了丰富的实时交易特征。

在数据预处理环节, 首先对原始数据进行清洗, 去除异常值与无效数据; 接着通过特征工程方法提取重要特征, 如交易频率、Gas 价格波动性及交易账户活跃度等; 随后采用归一化方法对连续型特征进行处理, 以提高模型训练的稳定性和准确性; 最后, 通过划分训练集、验证集与测试集, 为后续的 DQN 模型训练与性能评估奠定基础。

3.3 模型参数设定

模型参数设定如表 1 所示。

整体配置设计兼顾了交易排序的实时性、网络环境的复杂性以及 DQN 模型的收敛稳定性, 为后续实验中的策略验证与性能评估提供了坚实的基础。

表 1 DQN 训练实验参数设置

参数名称	设定值或说明
强化学习框架	Deep Q-Network
训练轮数	2 000
学习率 α	0.001, 采用 Adam 优化器
批次大小	64, 每轮训练样本数
折扣因子 γ	0.95, 用于平衡长期收益
经验回放池容量	10^4 条历史交易记录
探索策略	ϵ -greedy 策略, 初始 $\epsilon = 1.0$, 线性衰减至 0.01
目标网络更新频率	每 100 步与主网络参数同步一次
奖励函数设计	MEV 交易减少 $\rightarrow$ 给予正向奖励; 低 Gas 交易成功率提升 $\rightarrow$ 给予正向奖励; 平均确认延迟增加 $\rightarrow$ 给予负向奖励

3.4 评估指标说明

为了全面评估基于深度强化学习的以太坊交易排序优化方法的有效性, 本文展示了三个关键评估指标: MEV 提取率、交易公平性和 Gas 费用。

(1) MEV 提取率

MEV 指的是矿工、验证者或排序代理在打包区块时, 通过操控交易执行顺序所获得的额外经济收益。该现象在高频 DeFi 交易和资产价格波动剧烈的场景中尤为常见。为了评估排序策略对 MEV 的控制能力, 本文采用 MEV 提取率作为指标, 其计算公式如下:

$$\text{MEV 提取率} = \frac{N_{\text{MEV}}}{N_{\text{total}}} \times 100\% \quad (9)$$

其中, N_{MEV} 是当前区块中被识别为 MEV 类型的交易数量, N_{total} 是区块中所有交易的总数。

(2) 交易公平性指数

交易公平性衡量的是低 Gas 费用交易在排序优化后能否获得公平的执行机会。传统的 Gas 竞价排序中, 高 Gas 费用交易通常优先执行, 导致低费用交易被延迟或拒绝, 从而影响公平性。本文定义交易公平性指数为:

$$\text{公平性指数} = \frac{N_{\text{low}}^{\text{exec}}}{N_{\text{low}}^{\text{total}}} \times 100\% \quad (10)$$

其中 $N_{\text{low}}^{\text{exec}}$ 是成功打包进区块的低 Gas 费用交易数量, $N_{\text{low}}^{\text{total}}$ 是所有低 Gas 费用交易的总数。

(3) 平均 Gas 费用

平均 Gas 费用衡量的是排序优化后, 整体用户在交易中实际承担的 Gas 成本。其计算方式为:

$$\text{平均 Gas 费用} = \frac{1}{N} \sum_{i=1}^N G_i \quad (11)$$

其中, N 是当前区块中成功执行的交易数量, G_i 是第 i 笔交易的实际支付 Gas 费用。

通过减少交易中的平均 Gas 费用, 优化排序策略能有效降低用户的经济负担, 同时提升系统的资源分配效率。

3.5 实验结果分析

本文从 MEV 提取率、交易公平性指数与平均 Gas 费用三个维度对 FIFO、Gas-Based 以及所提出的 DQN 排序策略进行了实验对比。

(1) MEV 提取率

从图 2 中可以看出, 三种排序策略在 MEV 提取率上的表现存在显著差异。传统的 FIFO 策略和 Gas-Based 的排序策略分别达到 48.5% 和 53.2% 的 MEV 提取率, 表明这两种方法未能有效抑制 MEV 行为, 甚至在 Gas-Based 策略下由于交易优先级与支付费用直接相关, 反而更易受到 MEV 攻击者的利用。而本文提出的 DQN-Based 策略将 MEV 提取率显著降低至 12.7%, 有效压制了前置攻击、三明治交易等 MEV 操作。这一结果表明, 通过深度强化学习自动学习交易排序策略, 能够提升系统的抗 MEV 能力, 从而提高交易的公平性和整体网络稳定性。

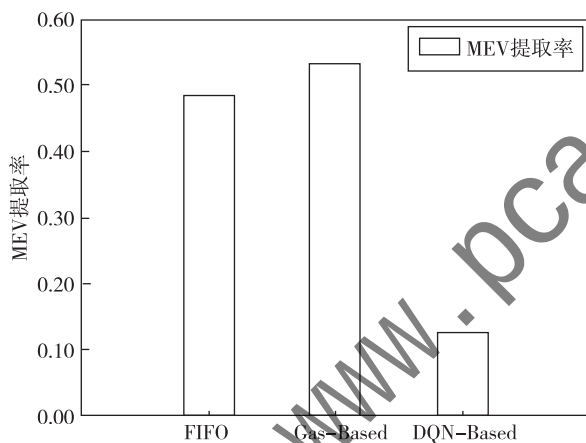


图2 MEV提取率柱状图

(2) DQN 训练过程中的 MEV 提取率变化率

图 3 展示了 DQN 排序策略在不同训练轮次下对 MEV 提取率的优化效果。随着训练轮次的增加, 策略模型逐步学习到更优的交易排序方法, MEV 提取率呈现持续下降趋势。初始阶段 MEV 提取率约为 50%, 此时代理尚未掌握有效的排序策略。随着训练深入, 至第 1 000 轮次 MEV 提取率已降至 18%。在 1 600 轮次之后, 模型趋于收敛, 提取率稳定在约 12.7% 左右。

该结果表明, DQN 强化学习策略在充分训练后能够显著抑制 MEV 行为, 降低区块内潜在的价值被恶意提取的风险, 从而提高交易执行的公平性与系统稳定性。同时, 也验证了通过持续学习与策略迭代, 模型能够有效

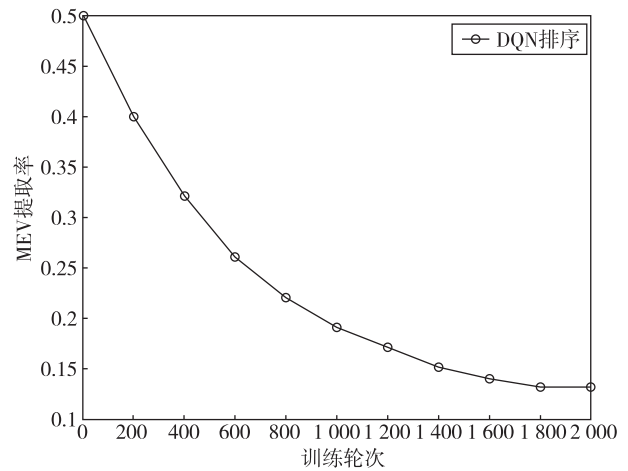


图3 DQN模型MEV提取率收敛趋势图

适应复杂的交易环境, 具备较强的泛化能力和鲁棒性。

(3) 公平性与平均 Gas 费用对比

图 4 展示了三种策略在交易公平性和平均 Gas 成本两个关键指标上的表现。DQN 排序策略在公平性指标上取得最高得分 (0.78), 显著高于 FIFO (0.32) 和 Gas-Based (0.24), 说明其在提升普通用户交易执行机会方面具有明显优势。此外, 在平均 Gas 成本方面, DQN 策略保持在 97 Gwei 左右, 远低于 Gas-Based 策略的 145 Gwei, 也低于 FIFO 策略的 132 Gwei 的水平。这表明 DQN 能在保障公平性的同时控制交易成本, 提升了系统整体经济效率。

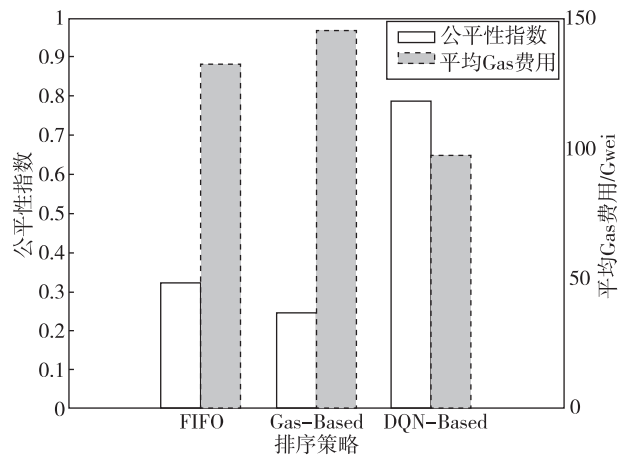


图4 不同排序策略下公平性与平均 Gas 费用对比

4 结论

本文围绕以太坊网络中的交易排序优化问题, 提出了一种基于 DQN 的动态排序方法, 旨在降低 MEV 的产生, 提升交易公平性, 并减少用户交易成本。通过引入多维交易特征构建状态空间, 结合经验回放与 ϵ -greedy 策略优化排序策略, 模型在多个指标上均优于传统 FIFO

和 Gas-Based 方法。实验结果表明,该方法可有效抑制 MEV 提取行为,显著提高低 Gas 交易的执行机会,并在不显著延长确认时间的前提下降低整体交易成本,展现出良好的实用价值与推广前景。

尽管该方法在模拟环境下取得了较好效果,但仍存在两个值得进一步研究的方向:一是如何提升模型在主网复杂交易环境下的适应能力;二是如何平衡模型训练时延与链上实时性需求。未来研究可考虑引入图神经网络、联邦学习等机制,进一步提升模型泛化能力和部署效率,以应对更加动态和安全敏感的区块链交易场景。

参考文献

- [1] BUTERIN V. Ethereum whitepaper [EB/OL]. (2025-02-12). [2025-04-18]. <https://ethereum.org/en/whitepaper/>.
 - [2] ZHANG H, WANG C, XU Z, et al. F3B: a low-overhead blockchain architecture with per-transaction front-running protection [J]. arXiv preprint arXiv: 2205.08529, 2022.
 - [3] NARAYANAN A, BONNEAU J, FELTEN E. Bitcoin and cryptocurrency technologies [M]. Princeton University Press, 2016.
 - [4] CHAN T, ZOHAR A. FIFO ordering and its limits in decentralized systems [J]. ACM Transactions on Economics and Computation, 2019, 7 (4): 32-50.
 - [5] ZHU B Z, WAN X, MOALLEMI C C, et al. Quantifying the value of revert protection [J]. arXiv preprint arXiv: 2410.19106, 2024.
 - [6] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning [J]. Nature, 2015, 518: 529-533.
 - [7] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of go with deep neural networks and tree search [J]. Nature, 2016, 529: 484-489.
 - [8] ZOHAR A, DAIAN P, JUELS A, et al. Flashbots: a transparent solution for MEV in Ethereum [EB/OL]. (2021-xx-xx) [2025-04-18]. <https://flashbots.org>.
 - [9] XU W, ZHANG Y, LIU M, et al. Improving fairness in decentralized exchanges by introducing time-weighted trading priority [J]. Journal of Blockchain Technology, 2021, 14 (2): 85-98.
 - [10] HEIMBACH L, SCHERTENLEIB E, WATTENHOFFER R. The potential of selfregulation for frontrunning prevention on DEXes [J]. arXiv preprint arXiv: 2306.05756, 2023.
 - [11] SARKAR D. FairFlow protocol: equitable maximal extractable value (MEV) mitigation in Ethereum [J]. arXiv preprint arXiv: 2312.12654, 2023.
- (收稿日期: 2025-04-28)
-
- 作者简介:**
- 严彦胜 (1997-), 男, 硕士, 主要研究方向: 区块链。
- 李京 (1966-), 男, 教授, 主要研究方向: 组合软件技术、大型网络系统和分布式算法。
-
- (上接第 19 页)
- [9] 胡枫杰. 基于 LightGBM 网络入侵检测系统的研究 [D]. 西安: 西安电子科技大学, 2021.
 - [10] 孙思佳. 基于模糊 SVD 和极限梯度提升树异常检测方法 [D]. 哈尔滨: 哈尔滨工程大学, 2021.
 - [11] KUMAR P, KUSHWAHA C, SETHI D, et al. Investigating the performance of multivariate LSTM models to predict the occurrence of Distributed Denial of Service (DDoS) attack [J]. PLOS ONE, 2025, 20 (1).
 - [12] 赵玉程, 李英建, 沈世民, 等. 基于网格搜索和投票分类模型的喷油器故障诊断研究 [J]. 机床与液压, 2024, 52 (5): 213-220.
 - [13] 刘俊泽, 汤艳君, 薛秋爽. 基于贝叶斯优化 LightGBM 的物联网入侵检测模型 [J]. 警察技术, 2022 (5): 73-77.
 - [14] RAJKUMAR K, SHALINIE S M. Semi-supervised deep-ELM for DDoS attack detection and mitigation using the OptimalLink model in IoT networks [J]. Computers & Security, 2025, 152 (C).
 - [15] SINHA M, BERA P, SATPATHY M, et al. A hybrid light-weight defense system against address spoofing based DDoS attacks in SDN [J]. Security & Privacy, 2025, 8 (2).
- (收稿日期: 2025-05-28)
-
- 作者简介:**
- 胡宏伟 (2000-), 男, 硕士研究生, 主要研究方向: 计算机网络与信息安全。
- 孙皓月 (1980-), 男, 本科, 副教授, 主要研究方向: 计算机网络与信息安全。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com