

基于智能合约的可信数据资产动态定价模型

徐茵¹, 夏琦¹, 高建彬¹, 夏虎¹, 吴超辉², 尹国庆³, 公茂玉⁴

1. 电子科技大学, 四川 成都 611731;
2. 66015 部队, 北京 100192;
3. 32178 部队, 北京 100012;
4. 信息工程大学, 河南 郑州 450001

摘要

数据资产定价作为数据要素市场价值释放的核心环节, 其机制的有效性对市场流通效率具有决定性作用。针对现有定价方案存在的执行过程透明度不足与对非平稳环境适应性差的问题, 提出了一种基于智能合约的可信数据资产动态定价模型 SC-TDP。具体而言, SC-TDP 构建基于马尔可夫决策过程的链上定价闭环, 利用智能合约的不可篡改特性, 确立定价全流程的可信度; 引入基于概念漂移指数的去中心化治理机制, 捕获市场分布的统计变化, 用于保障非平稳环境下的策略鲁棒性; 设计成本感知的自适应批处理策略, 平衡链上更新频次与 Gas 开销, 用于提升系统的工程可行性。在真实数据集上, 将 SC-TDP 与主流定价算法进行对比实验, 实验结果表明 SC-TDP 的综合性能优于对比算法。

关键词

区块链; 智能合约; 强化学习; 数据资产; 动态定价

中图分类号: TP311.1

文献标志码: A

doi: 10.11959/j.issn.2096-0271.

A Smart-Contract-Based Model for Trusted Dynamic Pricing of Data Assets

XU Han¹, XIA Qi¹, GAO Jianbin¹, XIA Hu¹, WU Chaohui², YIN Guoqing³, Gong Maoyu⁴

1. University of Electronic Science and Technology of China, Chengdu 611731, China;
2. The PLA Unit 66015, Beijing 100192, China;
3. The PLA Unit 32178, Beijing 100012, China;
4. Information, Engineering University, Zhengzhou 450001

Abstract

As a core link in unlocking the value of data factor markets, data asset pricing mechanisms have a decisive impact on market circulation efficiency. To address the lack of transparency in the execution process and the poor adaptability to non-stationary environments in existing pricing schemes, this paper proposes SC-TDP, a trusted smart-contract-based dynamic pricing model for data assets. Specifically, SC-TDP builds an on-chain pricing loop based on a Markov decision process (MDP) and leverages the immutability of smart contracts to establish trustworthiness across the entire pricing workflow; introduces a decentralized governance mechanism driven by a concept drift index to capture statistical changes in market distributions and ensure policy robustness under non-stationary environments; and designs a cost-

aware adaptive batching strategy that balances the frequency of on-chain updates against gas overhead to improve the engineering practicality of the system. Experiments on real-world datasets compare SC-TDP with mainstream pricing algorithms, and the results show that SC-TDP achieves superior overall performance.

Key words

Blockchain, Smart Contract, Reinforcement Learning, Data Assets, Dynamic Pricing

0 引言

随着互联网的日益发展,全球数据规模日益增大。根据中国信息通信研究院发布的《大数据白皮书(2020年)》^[1]的预测,2030年全世界数据产生量将达到612ZB,2035年将达到2142ZB。由此产生的海量数据在数据驱动型应用市场的蓬勃发展中含蕴巨大价值,数据资产应运而生。主流的中心化数据交易平台虽在一定程度上疏通了数据流通渠道,但黑箱式的定价与撮合机制,又引发了市场对价格操纵和利润分配失衡的担忧^[2]。

当前,学术界和工业界对数据资产的定价进行了广泛探索,主要形成了三类定价范式:一是价值评估驱动范式,旨在定价前期对数据资产的内在价值提供可量化的锚点,并以此为依据给出标价或价格区间。例如,提出数据估值的统一基准的OpenDataVal^[3]。二是机制主导范式,在已有基础价值的前提下,由市场结构与机制设计来决定最终成交价格,侧重于规则和参与者行为对价格的塑形作用。典型机制包括订阅制、双边议价及各种拍卖形式。例如,基于Stackelberg博弈的动态定价模型^[4]。三是历史学习与迁用范式,利用过往交易数据进行离线模型训练,随后以固定或低频更新的形式部署到线上。例如,在广告竞价场景中以历史轨迹做离线RL

训练,使离线学得策略在生产中实现可上线、可复用、收益可观的改进^[5]。

上述三类范式分别从定价的全生命周期不同环节取得了重要进展,但在真实平台中,仍存在端到端贯通不足的结构性问题,使数据资产定价陷入价值难评估、信任难建立、策略难优化的困境。因此,迫切需要寻求一种方法,实现数据资产定价的智能化、透明化与可信化,促进数据要素市场安全有序、平稳持续发展。

在此背景下,随着区块链技术在金融科技、供应链管理、数字身份等关键领域的应用不断深化^[6],区块链技术则以其不可篡改、可追溯的特性为定价价格形成提供了不可篡改与第三方可验证的审计底座^[7],确保了定价过程的公平性与合规性。而强化学习作为一种强大的动态决策方法,通过与环境的持续交互来学习最优策略,让定价策略根据市场反馈自我优化,从而适应瞬息万变的市场环境。鉴于此,本文从保障价值评估的动态性、定价机制的透明性以及策略优化的自适应性等角度出发,本文提出并实现了一个基于智能合约的可信数据资产动态定价模型,命名为SC-TDP(Smart-Contract-based Trusted Dynamic Pricing)。

SC-TDP的核心技术路线是将复杂的定价问题建模为马尔可夫决策过程(MDP),并在此基础上构建了一个链上-链下协同的定价引擎,通过智能合约库在链上实现了强化学习的闭环逻辑。这种分

层设计将高可信机制锚定在链上，将复杂计算与市场监测留于链下。为回应策略鲁棒性与市场非平稳性的双重挑战，本文设计了一套基于概念漂移的去中心化定价参数治理机制。该机制引入了概念漂移指数（CDI）与阈值触发规则，使其能够在旁路智能体中周期性地评估和校准现行定价策略。当 CDI 超出预设阈值时，系统将自动启动链上治理流程，以安全且可审计的方式演进策略版本，从而确保定价机制在不牺牲去中心化信任模型的前提下，维持对非平稳市场环境的自适应能力。同时，为解决证据留存与链上成本的尖锐冲突，本文设计并实现了 Q 值计算合约（QCC）的成本感知批量更新策略，通过聚合多个经验样本进行一次性更新，在保证模型学习效率 and 收敛速度的前提下，有效摊销了 Gas 成本，从而在策略同步延迟与链上资源成本之间取得了动态的最优平衡。本文的主要贡献如下：

- 本文针对现有定价方案面临的动态适应性不足与信任缺失问题，设计了一种基于智能合约的可信数据资产动态定价模型 SC-TDP。该模型同时利用强化学习的自适应决策能力与区块链的不可篡改特性，构建链上定价引擎，分别从策略演进的智能化与定价执行的透明化两个维度解决定价难题。
- 本文提出基于 CDI 的去中心化参数治理机制。利用 CDI 监测市场分布变化并触发链上治理，以保障定价策略在非平稳环境下演进的鲁棒性；提出成本感知的自适应批处理策略，利用数据驱动与时间驱动的双重触发机制聚合更新请求，以在策略同步延迟与链上 Gas 成本之间取得最优平衡。
- 本文在模拟交易环境与真实数据集上与主流定价算法进行了性能对比。结果

表明，本文提出的 SC-TDP 模型在长期累积收益与抗干扰能力上优于目前的主流方案，并且显著降低了链上执行成本。

1 研究现状

数据资产定价是数据交易系统中的核心环节，数据价值通常用交易时价格来体现。当前，不同定价方法依据关注点不同可分为三类：关注数据资产内在价值的高低来定价属于基于价值的数据定价；确定基础价值的前提下，考虑市场与参与人为影响属于基于经济学的定价；利用过往交易数据进行离线模型训练，随后以固定或低频更新的形式部署到线上属于基于模型的数据定价。三类方法可互补使用以提升市场灵活性与定价公平性。

在基于数据质量与内在价值的评估方面，现有研究致力于建立标准化的量化体系。Wang 等^[8]构建了包含 22 个维度的质量评估框架，后续综述^[9]进一步扩展了其线性加权体系。在此基础上，Zhang 等^[10]针对精确度、完整度等核心维度提出了 Floating 定价法，通过计算实例质量与基准的偏差来确定最终价格。Jiang 提出的 OpenDataVal^[3]则通过统一基准解决了不同估值方法的对齐难题。针对结构化数据的查询场景，Koutris 等^[12]提出的基于查询的定价（QBP）遵循无套利公理，将价格锚定为生成查询视图的最小成本，后续研究^[13]针对 KNN 等特定算法优化了其计算复杂度。然而，在面临概念漂移或跨任务迁移时，这种静态的价值锚点容易失真，无法灵敏地反映实时供需关系和市场中的策略互动^[11]。

基于经济学的定价机制侧重于利用市场结构解决分配问题。经典的供需模型^[14]

试图通过供给与需求函数实现市场出清。近年来出现了许多使用博弈论中方法研究数据定价^[15-21]的文献。博弈论主要研究决策主体的行为发生直接相互作用时的决策以及这种决策的均衡问题。非合作博弈假设交易各方独立决策以追求自身利益最大化，并在此基础上求解纳什均衡作为最终定价。Stackelberg 博弈则引入了领导者与跟随者的角色，强调先行者策略对后续响应者决策的引导作用。讨价还价模型则专注于在信息不完全的条件下，通过多轮协商达成价格协议。拍卖也是最流行的数据定价机制之一。拍卖是通过市场驱动参与拍卖的双方在规则框架内进行自主竞价，从而对商品进行分配，并赋予对应的价格^[22]。尽管这些机制理论严谨，但往往依赖于理想化的信息对称假设，且在实际市场中，平台或数据提供者往往比消费者掌握更多关于市场供需、数据价值或历史交易的信息。这种信息鸿沟可能导致价格被操纵，从而严重影响市场的公平性与透明度。

随着机器学习的发展，基于模型的智能化定价成为新趋势。一类路径关注交易形态的演进，如 Jia 利用沙普利值（Shapley）值量化数据对模型的边际贡献，Chen 等^{[23][24]}及 Agarwal 构建了出售模型实例或预测 API 而非原始数据的交易框架，实现了所有权与使用权的分离。另一种路径是数据定价作为一种动态调整产品价格的方法，结合了强化学习和贝叶斯推理等技术，为市场需求和竞争情况下的价格调整提供了科学依据。例如，Xu 等人^[25]将动态隐私定价建模为多臂老虎机问题，以平衡探索与利用；Bauer 和 Jannach^[26]则引入贝叶斯推理，试图在稀疏和噪声数据环境中实现最优定价。此外，Tsai 等人^[27]提出了基于机器学习的时变智

能定价框架，强调了时间维度对价格调整的影响。同时，自动评估数据产品价值的研究利用机器学习技术。唐兴叶^[28]利用文本分析技术处理消费者反馈，构建了基于顾客感知价值的定价模型；Sun 和 Gong^[29]则结合大数据思维与经济学理论，利用机器学习预测资产的期权价值。为了进一步提高评估精度，倪渊等人^[30]提出了改进的 AGA-BP 神经网络模型，通过迭代训练实现了对数据资源价值的自动评估。这类方法通常采用离线训练+在线预测的模式，高度依赖于大量历史交易数据。这种基于离线数据集训练得出的模型，本质上是对历史市场状态的拟合。在面对数据要素市场高频波动、需求瞬息万变的极端动态场景时，模型往往因无法实时更新参数而出现概念漂移或响应滞后，难以真正适应实时动态市场的定价需求。

2 模型总体设计与理论建模

从保障定价过程的可信度与定价策略的动态演化能力这两个基本维度出发，本研究聚焦于设计并实现一个部署于区块链网络之上的智能化定价引擎。该引擎旨在从根本上解决传统中心化定价模式固有的黑箱操作弊病、由信息不对称引发的信任赤字，以及静态定价规则无法有效适应瞬息万变的市场环境等多重难题，其总体方案如图 1 所示。

具体来说，本方案首先将复杂且动态的数据资产定价问题，严谨地抽象为一个马尔可夫决策过程（MDP）数学模型。这一建模为后续引入强化学习算法提供了坚实的理论框架，使得系统能够在一个充满不确定性的市场环境中进行序贯决策。在此模型基础上，方案的核心是一个部署在

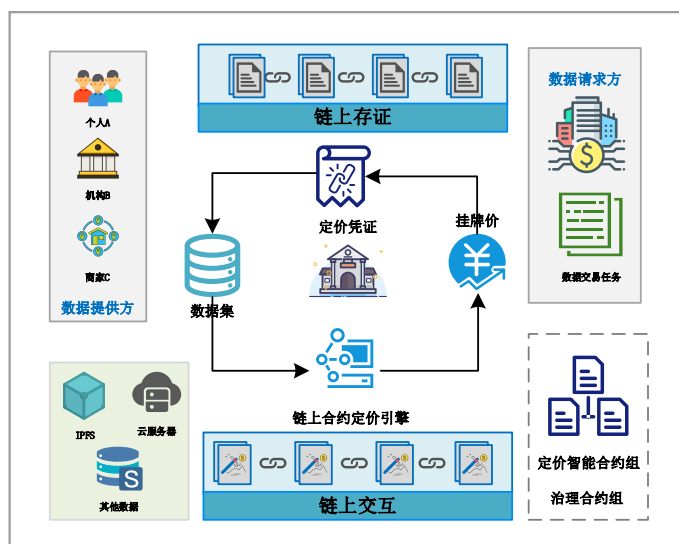


图1 数据资产动态定价方案

区块链上的智能化定价引擎。该引擎由一组智能合约构成，作为MDP框架中的代理（Agent），其行为完全由代码定义，公开透明且自动执行。在每个定价周期开始时，该引擎会主动市场从历史价格与成交结果的滑动窗口中提炼出的关键指标。随后，它会依据通过历史经验学习到的定价策略（Q值函数），计算并输出一个旨在最大化长期期望收益的最优价格决策。该决策被传递至链下数据市场进行实际的交易执行与撮合。交易完成后，为了将分散在链下的、结果不确定的商业行为转化为链上模型可以理解和学习经验，我们设计了一套高效且可信的链下结果反馈机制，负责将每一笔交易的最终结果转化为一个量化的奖励信号并传送回链上。链上定价引擎在接收到奖励信号后，会完成一次完整的强化学习迭代。它将此次的状态-决策-奖励作为一个新的经验样本，用作后续更新内部的定价策略，从而实现策略的自我优化与迭代。

2.1 数据资产的动态定价问题定义

数据资产动态定价问题可以形式化表达为：给定数据资产特征集合 X 和潜在的价格集合 P ，在时间 t ，根据历史交易数据和市场环境信息，确定最优价格 $p_t \in P$ ，使得长期累积收益最大化。以上架或价格刷新作为离散决策时刻 $t=1,2,\dots$ 。给定待定价的数据资产集合 $DP = \{(x_1, p_1), (x_2, p_2), \dots, (x_t, p_t), \dots\}$ ，其中数据资产特征向量 $x_t = [x_{t1}, x_{t2}, \dots, x_{td}]^T \in R^d$ 。表征数据的时效性、完整性等属性， d 为样本的特征维度。市场环境状态 S_t 包含了做出定价决策所需的上下文信息。为了刻画可观测的市场上下文，本方案在区块链维护了两个关键的滑动数据窗口，作为构建 S_t 的核心数据源：

历史价格窗口 $PW[N]$ 是链上存储的、大小为 N 的队列，用于记录最近 N 笔交易的价格历史。该窗口遵循先进先出（FIFO）的滑动机制。当一笔新交易在 t 时刻完成，其价格 p_t 压入 $PW[N]$ 的队尾；若队列长度超过 N ，则最旧的价格 p_{t-N} 将自动从队首移除。

交易结果窗口 $TW[N]$ 是链上存储的、

大小为 N 的队列，用于记录最近 N 笔交易的结果，与 $PW[N]$ 同步更新。当 t 时刻的交易完成，交易结果 $result_t$ 被压入队尾，同时 $result_{t-N}$ 从队首移除。模型可以从 $TW[N]$ 中提取反映市场接受度的关键指标，最典型的是近期交易成功率。

因此，本文所研究的动态定价问题，可以被进一步精确地表述为寻找一个最优定价策略函数 π^* 。该函数 π^* 接收当前资产的内在特征 x_t 以及两个历史信息窗口 $PW[N]$ 和 $TW[N]$ 作为输入，并输出一个最优的价格决策 p_t ，即：

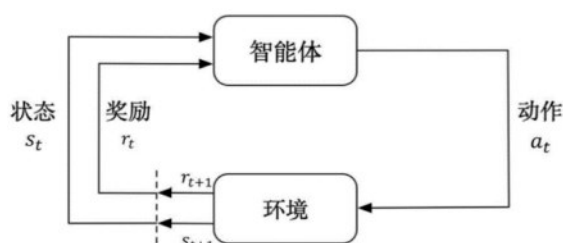
$$p_t = \pi^*(x_t, PW[N], TW[N]) \quad (1)$$

该定义下的核心挑战在于，这个最优策略 π^* 必须能够有效地从 $PW[N]$ 和 $TW[N]$

这种高维、动态变化的序列数据中，提取出反映市场真实状态的关键信息，并应对由于窗口滑动而带来的决策上下文的持续演变。本研究后续将此问题建模为马尔可夫决策过程（MDP），该过程是求解上述最优策略 π^* 的核心方法论。

2.2 动态定价问题的马尔可夫决策过程(MDP)建模

马尔可夫决策过程是一种用于序列化决策的数学框架，适用于动态定价场景建模。假设总决策周期为 T ，智能体在每个周期 $t \in [1, 2, \dots, T]$ 观测环境状态 S^0 ，执行定价动作 a^0 后获得即时奖励 $r^{(t+1)}$ ，并通过状态转移函数 P 更新至下一状态 $S^{(t+1)}$ ，如图2所示。



为解决2.1节所定义的动态定价问题，本文将数据资产动态定价问题建模为马尔可夫决策过程(MDP)，定义为五元组 $M = (S, A, R, P, \gamma)$ 。传统MDP模型部署于中心化服务器，交易参与者无法验证平台提供的市场状态的真实性，缺少奖励函数的计算的公正性验证手段，导致定价策略的执行陷入信任赤字。为从根本上解决这一信任难题，本文的MDP建模创新性地将区块链的透明性、可验证性与自动化特性，深度融合到五元组的定义之中，构建一个可信

MDP框架，其五元组定义为：

(1) **状态空间 S** ：表示系统当前所处的市场状态。本模型的状态 S_t 基于链上公开可验证的数据构建，其核心上下文语境一的历史价格窗口 $PW[N]$ 和交易结果窗口 $TW[N]$ 在区块链上公开维护，确保 Agent 决策所依据的状态是透明、可追溯且无法被恶意篡改的。

(2) **动作空间 A** ：表示可供系统选择的价格调整动作集合，具体包括一系列离散的价格变动幅度，例如“提高价格5%”、

“降低价格10%”或“保持不变”等操作。每一个动作都对应着对当前资产价格的一个特定调整策略。本模型选择和执行该动作的策略 π^* 被编码为链上的自动化业务逻辑（即智能合约），确保决策被刚性执行，杜绝平台方的人为干预。

(3) 奖励函数R：用于度量在某状态下采取某动作所获得的即时反馈。本文将奖励函数定义为资产是否成功交易所带来的实际收益：

$$R(s, a) = price \times 1_{成交} \#(2)$$

其中， $1_{成交}$ 表示交易是否成功的指示函数，若成交则为1，否则为0，从而体现价格调整对交易成功率和最终收益的直接影响。

(4) 状态转移概率P：描述系统在执行动作a后，从当前状态s转移至下一个状态s'的概率分布，即 $P(s'|s, a)$ 。该概率模型体现了市场反馈的随机性和策略执行的非确定性。交易完成后对PW[N]和TW[N]窗口的更新在链上执行并记录的确保整个状态转移轨迹具备完全的可审计性。

(5) 折扣因子 γ ：用于衡量未来收益对当前决策的影响程度，取值范围为 $\gamma \in [0, 1]$ 。较高的折扣因子表明系统更加重视长期收益；而较低的折扣因子则更侧重于短期即时收益的最大化。作为影响长期策略的核心参数， γ 作为链上系统的公共参数进行存储，其设定或修改必须遵循透明的去中心化治理机制，而非由平台单方面决定。

MDP建模的目标是求解一个最优策略 π^* ，使得系统在任意初始状态下，按照该策略采取行动时所获得的期望累计折扣奖励最大化。本节通过将MDP的关键要素锚定于可验证的链上事件，不仅成功地对动态定价问题进行了数学建模，并且在建模的同时系统性地解决了传统中心化方案

固有的信任问题。

3 基于智能合约的可信数据资产动态定价模型框架

本节详细阐述了作为解决方案核心的链上定价引擎的静态架构，包括链上定价层与模型治理层，模型的整体框架如图3所示。该框架在设计上采用了逻辑解耦的两层协同架构来实现可信定价到自适应学习的闭环。这种分层设计的核心思想在于将链上定价层承担的可审计的机制与证据与高算力的数据工作分离，以解决链上信任与链下复杂计算之间的固有矛盾。

其中，链上层是系统的信任根，负责规则的刚性执行、身份校验、以及状态与事件的不可篡改留痕。该层的核心是由一组协同工作的智能合约集群构成，即价值评估合约(VEC)、定价执行合约(PEC)、交易反馈合约(TFC)和Q值计算合约(QCC)。这四个合约组件共同在链上实现了MDP智能体的角色，构成了定价—结果—奖励—批量更新的闭环逻辑，确保每一步均有可审计的链上证据。模型治理层不直接参与实时的MDP学习与在线定价，它的核心职责是订阅链上定价层输出的事件流，离线完成效果评估、市场漂移监测与回放仿真，并以签名工件的形式向链上提交参数与版本的治理提案。这两层架构的安全协同是通过严格定义的窄接口实现的。模型治理层通过上行接口以只读方式订阅链上定价层广播的可重放事件流，进行离线分析；而该层产出的治理提案则通过下行接口提交回链上。为保障链上安全与控制权，这些工件并非自动生效，而是必须经过合约的合法性校验，并遵照既定的治理流程被显式激活，从而确保所有参数变更的可控性、可审计性与可回滚性。

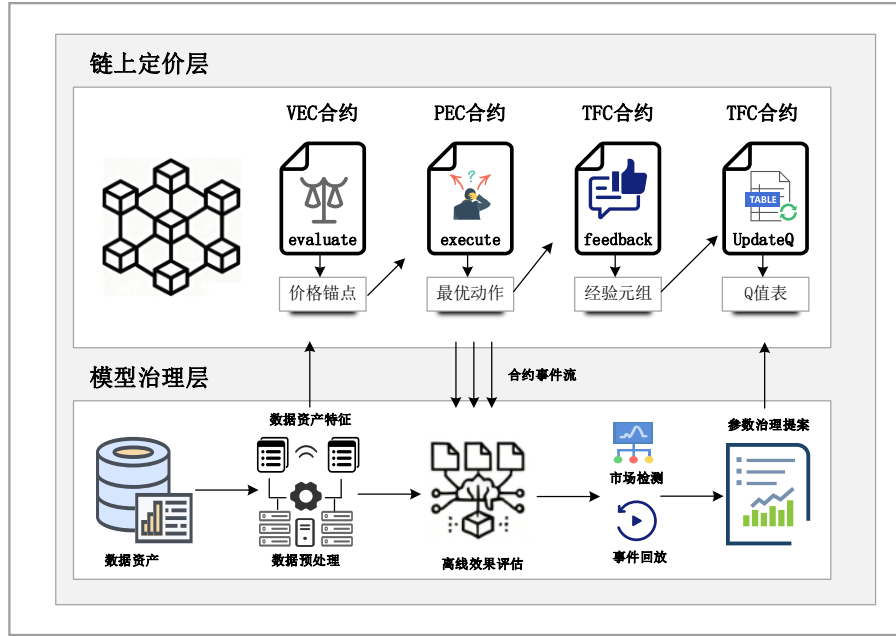


图3 基于区块链与强化学习的定价模型框架示意图

3.1 链上定价层

链上定价层以四个智能合约构成一个可审计的MDP闭环，VEC提供估值锚点，PEC做价格决策，TFC将成交结果规范化为奖励并形成下一状态，QCC在批量触发下更新价值函数并发布版本。为保证可追溯性，层内所有关键中间量以最小摘要在链上留痕，状态以哈希承诺 $h_t = H(s_t)$ 记录，决策与更新以事件统一对外。状态构造遵循资产特征结合近期上下文的原则，资产特征为 x_t ，近期上下文由大小为 N 的历史价格窗口 $PW[L]$ 和交易结果窗口 $TW[L]$ 的统计量 $\psi(PW_L)$ 、 $\psi(TW_L)$ 给出，则本层使用的可计算状态为：

$$s_t = \phi(x_t, \psi_{PW[L]}, \psi_{TW[L]}, meta_t) \# (3)$$

3.1.1 价值评估合约

价值评估合约（VEC）作为定价闭环中的价值锚点，核心功能是一个部署在链上定价层的轻量级、可治理的估值模型

f_{VEC} 。合约的主要函数接收来自数据所有者提交的数据资产特征向量 $x_t = [x_{t1}, x_{t2}, \dots, x_{td}]^T \in R^d$ ，该向量表征了数据的 d 维属性，基础估值 $v_{0,t}$ 的计算过程如下：

$$v_{0,t} = f_{VEC}(x_t, W_{VEC}) = \sum_{i=1}^d w_i \cdot x_{t,i} + b \# (4)$$

其中， $W_{VEC} = [w_1, w_2, \dots, w_d, b]^T$ 是可治理的模型参数， w_i 代表第 i 个资产特征的权重， b 是模型的偏置项。 W_{VEC} 是由模型治理层在离线分析后，通过下行接口以签名工件的形式提交至链上。VEC 计算完成后，将 $v_{0,t}$ 作为价格锚点输出给 PEC 合约。

3.1.2 定价执行合约

定价执行合约（PEC）是链上定价层的策略执行器，负责在每个决策时刻将当前的状态 h_t 和 Q 值函数 $Q(h_t, a)$ ，选择一个最优动作 $a_t \in A$ ，并将其映射为最终的链上价格 p_t 。PEC 采用 ϵ -greedy 策略去实现强化学习中最优策略求解的关键平衡。

PEC 合约的核心决策逻辑如算法 1 所示:

算法 1 定价执行合约 (PEC) 的核心决策逻辑

输入: 当前状态承诺 h_t ; VEC 估值 $v_hat > 0$; 动作集合 A (离散价带/倍率); Q 值表 $Q[h, a]$; 上次成功交易价格 $last_price$ (可空)

输出: 本轮价格 p_t ; 事件 `ListingDecided`($h_t, a_t, p_t, meta$)

```

1: require  $v\_hat > 0$  ▷ 估值必须为正
2: if not HASENOUGHSAMPLES( $L_{min}$ ) then
3:    $\alpha \leftarrow \text{Uniform}(1 - \nu\_cs, 1 + \nu\_cs)$  ▷ 冷启动: 样本不足时, 在估值附近无偏探索
4:    $p^{raw} \leftarrow \alpha \cdot v\_hat$ 
5:    $p_t \leftarrow \text{clamp}(p^{raw}, p_{min}, p_{max})$  ▷ 边界裁剪, 落在  $[p_{min}, p_{max}]$ 
6:   EMIT ListingDecided( $h_t, \text{"COLD\_START"}, p_t, \{\alpha\}$ ) ▷ 留痕: 记录冷启动挂牌
7:   return  $p_t$ 
8: end if
9:  $u \leftarrow \text{Uniform}(0, 1)$ 
10: if  $u < \epsilon$  then
11:    $a_t \leftarrow \text{UNIFORMPICK}(A)$  ▷ 探索: 以  $\epsilon$  概率随机选价带
12: else
13:    $a_t \leftarrow \arg\max_{a \in A} Q[h_t, a]$  ▷ 利用: 选择当前状态下 Q 值最大的动作
14: end if
15:  $p^{raw} \leftarrow a_t \cdot v\_hat$ 
16:  $p_t \leftarrow \text{clamp}(p^{raw}, p_{min}, p_{max})$  ▷ 再次裁剪, 防止越界挂牌
17: if  $last\_price \neq$  and ISSMALLCHANGE( $p_t, last\_price, \Delta_{min}$ ) then
18:    $p_t \leftarrow last\_price$  ▷ 抖动抑制: 变动低于阈值则复用上次价格
19: end if
20: EMIT ListingDecided( $h_t, a_t, p_t, \{\}$ ) ▷ 留痕: 状态承诺/动作/最终挂牌价
21: return  $p_t$ 

```

动作空间 A 被定义为一个离散且有界的系数集合 $A = \{\mu_1, \dots, \mu_K\}$, ϵ -greedy 策略以 $1 - \epsilon$ 的概率利用当前的最优动作, 以 ϵ 的概率在 A 上探索一个随机动作:

$$a_t = \begin{cases} \arg\max_{a \in A} Q(h_t, a), & \text{以概率 } 1 - \epsilon \\ \text{Unif}(A), & \text{以概率 } \epsilon \end{cases} \quad \#(5)$$

在确定 a_t 后, PEC 结合基础估值 $v_{0,t}$, 应用价格边界 $[p_{min}, p_{max}]$ 裁剪, 生成最终价格 p_t :

$$p_t = \text{clip}(\mu_{a_t} \cdot v_{0,t}, [p_{min}, p_{max}]) \quad \#(6)$$

PEC 合约内置了对冷启动场景来保障策略的稳定与鲁棒性。当 TFC 的历史窗口 PW_L, TW_L 未达最小样本门槛 L_{min} 时, 合约将绕过 ϵ -greedy 决策, 转而在 $v_{0,t}$ 锚点附近的一个小范围邻域内进行无偏探索以生成首价, 随窗口数据的填充再自动切换回常态策略。此外, 为抑制价格无效抖动, 合约会校验抖动阈值 θ (若 $|p_t - p_{t-1}|/p_{t-1} < \theta$ 则跳过价格刷新)。合约在执行完毕后, 会将此次决策的完整凭据, 包含状态承诺 h_t 、所选动作 a_t 、最终价格 p_t 固化为一个标准化的 `ListingDecided` 事件, 将其在链上公

开留痕。

3.1.3 交易反馈合约

TFC 合约是连接链上定价层与外部市场的可信数据桥梁，负责接收链下市场的最终交易结果，规范化为 QCC 学习所需的标准数据格式，TFC 合约的反馈处理算法如算法 2 所示：

当一笔交易完成或超时，外部市场调用 TFC 的 reportOutcome 函数上报交易结果 o_t 。 o_t 包含三个关键信息：是否成交 $y_t \in \{0, 1\}$ 、实际收益 payoff_t 、售出时延 τ_t ，其中，售出时延 τ_t 用于度量从资产上架到成功售出的时间成本；若交易超时失败，则 τ_t 取其最大值 - 订单有效期。TFC 首先

根据链上固化的、可治理的奖励函数计算即时奖励 r_t ：

$$r_t = \alpha \cdot \text{payoff}_t - \beta \cdot \tau_t - \gamma_f \cdot (1 - y_t) \quad (7)$$

计算完成后，合约以先进先出方式滚动更新 PW_L 和 TW_L 两个状态窗口。基于更新后的窗口统计量 $\psi(\cdot)$ ，合约必须使用与公式(3)相同的 ϕ 函数及哈希 H 来重建下一状态 s_{t+1} 的摘要及其承诺 $h_{t+1} = H(s_{t+1})$ 。至此，TFC 已集齐了构建一个完整经验元组 (h_t, a_t, r_t, h_{t+1}) 所需的全部信息。该元组是强化学习智能体用于学习的基础数据单元，它封装了一个完整的决策周期信息，包括决策前的状态 h_t 、执行的动作 a_t 、获得的奖励 r_t 以及转移到的下一状态 h_{t+1} 。

算法 2 交易反馈合约 (TFC) 反馈处理算法

输入： 交易标识 tx_uid ；成交标记 $\text{sold} \in \{0, 1\}$ ；成交收益 $\text{revenue} \geq 0$ ；耗时 $\text{delay} \in [0, \tau_{\max}]$ ；滑动窗口 W_price, W_result (容量 N)；上轮留痕 (h_t, a_t, p_t) ；奖励函数 $R(\cdot)$ ；承诺函数 $\text{STATECOMMIT}(\cdot)$

输出： 即时奖励 r_t ；下一状态承诺 h_{t+1} ；事件 $\text{TradeSettled}/\text{TradeFailed}$ 、 SampleBuffered

```

1: require  $tx\_uid$  已绑定  $(h_t, a_t, p_t)$                                 ▷ 一致性校验
2:  $r_t \leftarrow R(\text{sold}, \text{revenue}, \text{delay})$                                 ▷ 将结果规范化为奖励：收益 - 时间成本
3: if  $\text{sold} = 1$  then
4:    $\text{EMIT TradeSettled}(tx\_uid, \text{revenue}, \text{delay}, r_t)$                 ▷ 留痕：成交
5: else
6:    $\text{EMIT TradeFailed}(tx\_uid, \text{delay}, r_t)$                             ▷ 留痕：失败/超时
7: end if
8:  $W\_price.PUSH(p_t)$ ;
9: if  $|W\_price| > N$  then  $W\_price.POPHEAD()$ 
10: end if                                                                ▷ 价格窗口 FIFO 滚动
11:  $W\_result.PUSH(\text{sold})$ ;
12: if  $|W\_result| > N$  then  $W\_result.POPHEAD()$ 
13: end if                                                                ▷ 结果窗口 FIFO 滚动
14:  $\phi_p \leftarrow \text{STATS}(W\_price)$ ;  $\phi_y \leftarrow \text{STATS}(W\_result)$  ▷ 提取窗口统计 (均值/EWMA/波动度/成功率)
15:  $x_t \leftarrow \text{GETASSETFEATURES}(tx\_uid)$                                 ▷ 读取资产内生特征 (只读)
16:  $s_{t+1} \leftarrow \text{CONCAT}(x_t, \phi_p, \phi_y)$ 
17:  $h_{t+1} \leftarrow \text{STATECOMMIT}(s_{t+1})$                                 ▷ 序列化并哈希形成状态承诺
18:  $\text{BUFFER.APPEND}(h_t, a_t, r_t, h_{t+1})$                                 ▷ 写入样本缓冲池，供 QCC 学习
19:  $\text{EMIT SampleBuffered}(tx\_uid, h_t, a_t, h_{t+1})$                     ▷ 留痕：样本入池

```

TFC 合约在获取到一个新的完整经验元组后，会触发两个关键的链上事件。首先根据交易结果 y_t ，公开记录该笔交易的最终商业结果；并将该经验元组写入 Q 值学习的样本缓冲池，供后续的批量学习使用，从而保证结果—奖励—状态滚动的证据链完整性。

3.1.4 Q 值计算合约

QCC 合约是链上定价层链学习执行体，是实现链上强化学习的核心，负责 Q-learning 算法的链上工程实现，如算法 3 所示：

算法 3 Q 值计算合约 (QCC) 的批量更新策略

输入： 样本缓冲池 BUFFER (元素 (h_t, a_t, r_t, h_{t+1}))；动作集合 A ；Q 表 $Q[h, a]$ ；学习率 η ；折扣因子 γ ；触发参数 K_batch, B_max ； $lastUpdateBlock$

输出： 事件 QUpdated(batch_size, gas_used, digest)

```

1: need_by_size  $\leftarrow (|BUFFER| \geq K\_batch)$  ▷ 数据触发：累计到最小批量
2: need_by_time  $\leftarrow (CURRENTBLOCK() - lastUpdateBlock \geq B\_max)$  ▷ 时间触发：超过心跳阈值
3: if not (need_by_size or need_by_time) then
4:   return "SKIP"
5: end if
6: B  $\leftarrow \text{POP\_BATCH}(BUFFER, K\_batch)$  ▷ 按入队顺序取出（可重放审计）
7: gas  $\leftarrow 0$  ▷ 初始化“本批 Gas 用量计数器”
8: for all  $(h_t, a_t, r_t, h_{t+1}) \in B$  do
9:   q_max  $\leftarrow \max_{a' \in A} Q[h_{t+1}, a']$  ▷ 目标中的“下一状态最优 Q”
10:  q_old  $\leftarrow Q.GET(h_t, a_t, 0)$ 
11:  y  $\leftarrow r_t + \gamma \cdot q\_max$ ;  $\delta \leftarrow y - q\_old$  ▷ TD 目标与 TD 误差
12:  Q.SET( $h_t, a_t, q\_old + \eta \cdot \delta$ ) ▷ TD(0) 增量更新
13:  gas  $\leftarrow gas + \text{GAS\_ESTIMATE\_PER\_UPDATE}()$  ▷ 累计 Gas 估算值
14: end for
15: lastUpdateBlock  $\leftarrow CURRENTBLOCK()$  ▷ 更新心跳高度
16: digest  $\leftarrow H(\text{SERIALIZE}(B))$  ▷ 对本批更新序列做承诺摘要
17: EMIT QUpdated( $|B|, gas, digest$ ) ▷ 留痕：发布批次大小/Gas/摘要
18: return "UPDATED"
```

Q-learning 学习一个最优的状态—动作价值函数，这个函数用于评估智能体在特定状态 h 下执行特定动作 a 后所能获得的未来累计折扣奖励。Q-learning 通过时序差分 (TD) 的方式进行学习，无需等到全局结束，仅通过比较当前的估计和下一步的估计之间的差异就能更新当前估计值。QCC 合约利用 TFC 合约提供的经验元组

(h_t, a_t, r_t, h_{t+1}) 来不断迭代优化 Q 值，在链上持久化地存储 Q 值学习产生 Q 值表 $Q(h, a)$ ，如表 1 所示。QCC 合约的 Q 值更新所执行的标准时序差分更新如下：

$$Q(h_t, a_t) \leftarrow (1 - \eta)Q(h_t, a_t) + \eta[r_t + \gamma \max_{a' \in A} Q(h_{t+1}, a')] \quad (8)$$

其中 $\eta \in (0, 1]$ 为学习率、 $\gamma \in [0, 1)$ 为折扣因

子、 A 为离散价带集合。该公式是一个增量更新过程，合约并非用目标价值直接覆盖旧值 $Q(h_t, a_t)$ ，而是以学习率 η 执行一次加权平均，确保 Q 值表不会因单次、可能带有噪声的经验而剧烈波动，从而能够平稳地向最优解收敛。

表1 Q值表示例

状态 承诺 h	0.90	0.95	1.00	1.05	1.10
h1	0.213	0.276	0.318	0.301	0.244
h2	0.045	0.102	0.154	0.165	0.121
h3	0.332	0.348	0.341	0.320	0.271

在链上实现公式(8)的核心工程挑战在于目标价值 $[r_t + \gamma \max_{a' \in A} Q(h_{t+1}, a')]$ 的计算。合约需遍历整个离散动作空间 A ，逐一读取 h_{t+1} 状态行中每个动作 a' 对应的 Q 值，比较找出其中的最大值。这个 $O(|A|)$ 复杂度的循环操作将导致 Gas 费用高昂且不可预测。为在收敛速度与链上成本之间取得平衡，QCC 合约通过一个双重触发机制来执行批量更新：

- 数据驱动触发器 ($\geq K$)：当“缓冲池样本数 $\geq K$ ”时触发。这保证在市场活跃时，合约能以足够多的数据量进行高效更新。

- 时间驱动触发器 ($\geq N$)：当“自上次更新已过去 N 个区块”时触发。这是一个心跳式的兜底机制，用于确保在市场交易稀疏 K 长时间无法攒满时，模型也能被强制更新，从而防止策略因数据饿死而变得陈旧过时。

批量更新完成后，缓冲池将被清空，QCC 合约发布 QUpdatd 事件留痕。在模型下一次定价决策时会读取与 h_t 状态对应的 Q 值集合 $Q(h_t, \cdot)$ ，从而自动感知策略的

演化。

3.2 模型治理层

模型治理层作为本定价模型架构中的旁路智能体，不直接参与在线的 MDP 学习与定价执行，是应对强化学习驱动的链上定价策略在非平稳市场中概念漂移问题而建立。强化学习策略一旦部署，如果缺乏有效的外部治理，模型极易在无法适应市场变化与过拟合/震荡的风险间摇摆。因此，模型治理层是在可验证、可审计、可回滚的约束下，独立于实时报价路径周期性地评估并升级对应的强化学习策略。

我们将模型治理层的设计目标明确为四点：(G1) 稳健性，确保对市场漂移有灵敏但不失控的响应；(G2) 可审计性，所有策略或参数更新可被第三方独立复核；(G3) 可回滚性，在策略效果不佳时能快速安全退回至稳定版本；(G4) 链上开销最小化，大对象在链下保存，链上仅存承诺与元数据。

模型治理层通过链上治理合约和注册表合约实现与链上定价层的协同控制。本层以链上定价层广播的只读事件流作为输入源，这些事件按区块高度和事件序号排序，在链下实时重建可重放的决策与结果时间线。输出是结构化的治理提案 (proposal)，包含变更说明、参数/模型承诺、评测摘要及回滚指针，最终体现为链上的策略激活或回退事件。

模型治理层核心在于维护一个周期性的监测—评估—决策的链下控制回路。在监测阶段，模型治理层持续计算情境漂移指数 (CDI)，以近期窗口与基准窗口的成交率、单位时间收益或价格波动特征的统计差异为核心信号，来量化市场环境的动态变化。在重放的事件时间线上，模型治

理层对候选参数/策略进行离线评估，输出包括总收益、成交率、稳定性、后悔度及 Gas 影响估算等关键指标，并生成数据快照哈希和代码版本哈希的评测摘要。模型治理层的决策是基于 CDI 阈值与评估结果完成的。若 CDI 低于阈值 τ_1 ，则保持现状；介于 τ_1 与 τ_2 时进入观察期并积累更多证据；若超过 τ_2 ，则触发治理提案的生成流程；显著超过 τ_3 时，可能触发应急闸并准备快速回退方案。这一闭环确保了策略的更新节奏由统计证据驱动，而非主观触发。

生成的治理提案最终被打包为签名工件通过下行接口提交至链上，并需遵循一套严格的链上治理流程，由链上治理合约执行。提案首先被提交至链上治理合约。合约对其进行严格校验，包括发起者签名有效性、内容哈希与承诺的匹配性，以及关键参数是否符合预设的参数域校验规则。校验通过的提案将进入公示期，并开放给社区（DAO）进行链上投票。此阶段通过发布 ProposalSubmitted 事件告知社区提案详情，确保透明度（满足 G2）。投票通过的提案并非立即生效，而是被置于一个预设的时间锁延迟期中。此阶段通过发布 ProposalQueued 事件留痕。该延迟期是系统的一道重要安全护栏，为所有利益相关者提供了对即将生效的变更进行最终审计、风险评估和潜在否决的窗口。时间锁期结束后，有权限的实体（如链下 Relayer）可调用执行函数，指令链上治理合约执行提案，原子性地更新注册表合约中记录的策略版本承诺，并进一步更新目标合约中的参数。若治理投票失败或在时间锁期间被否决，系统可发布 ProposalRolledBack 事件（满足 G3）。

模型治理层的设计严格遵循治理与执行分离原则，确保控制权与问责边界绝对清晰，即便模型治理层本身出现问题，链

上定价层也能在不被直接篡改的情况下维持运行，从而在不牺牲去中心化信任模型的前提下，实现定价策略的安全、透明与自适应演进。

4 实验结果与分析

本节将对实验环节进行阐述。首先对实验准备工作进行说明，包括实验环境、实验数据集和评价指标；其次，分别对本文中进行的实验内容进行阐述，并对实验结果进行分析。

4.1 实验设置

本实验构建了链上-链下协同的仿真交易环境。链上部分基于以太坊 Hardhat 测试网搭建，智能合约通过 Solidity(v0.8.21) 编写；链下交互与市场模拟模块基于 Python3.10 与 Ethers.js 实现。本文从 Kaggle 平台选取两个在任务类型与数据结构上具有代表性的公开数据集作为待售数据资产，将其预处理定义为特征包，通过统一量化数据资产的内在价值，生成一个高维的价格-购买率概率表，该表作为链下模拟器的核心输入，为强化学习智能体提供了一个基于价格的市场需求函数，为价格博弈过程提供驱动依据。

为全面评估 SC-TDP 在策略优化与去中心化特性上的优势，本文选取了三种在决策模式与信任模型上具有代表性的算法作为对比基线，如表 2 所示：所有方法共用同一市场模拟器、相同随机种子集与相同价带集合 $A = \{0.90, 0.95, 1.00, 1.05, 1.10\} \times \hat{v}_0$ 。除特别说明外，RL 模型统一使用 ϵ -greedy 策略，学习率 $\eta = 0.10$ ，折扣因子 $\gamma = 0.95$ 。

表2 对比模型特性

模型	规则	策略	决策位置	信任模型
静态固定定价	价格恒等于VEC估值锚	静态	链上	去中心化
短期调整定价	根据近期成交结果对价格进行小幅调整	静态	链上	中心化
C-RL	在中心服务器上离线更新Q表	动态	链下	中心化
MDP动态学习(SC-TDP)	在智能合约中离线更新Q表	动态	链上	去中心化

4.2 SC-TDP的优化效果验证

本节旨在通过消融实验验证引入强化学习策略及链上治理机制对系统盈利能力与市场适应性的增益。实验总轮次设定为T=1000，并在t=500时刻引入显著的市

场漂移（模拟需求曲线结构性下移），以此评估各模型在非平稳环境下的动态响应能力。图6展示了静态基线、短期调整策略与SC-TDP在累计净收益（Rev）与成交成功率（CR）上的对比结果。

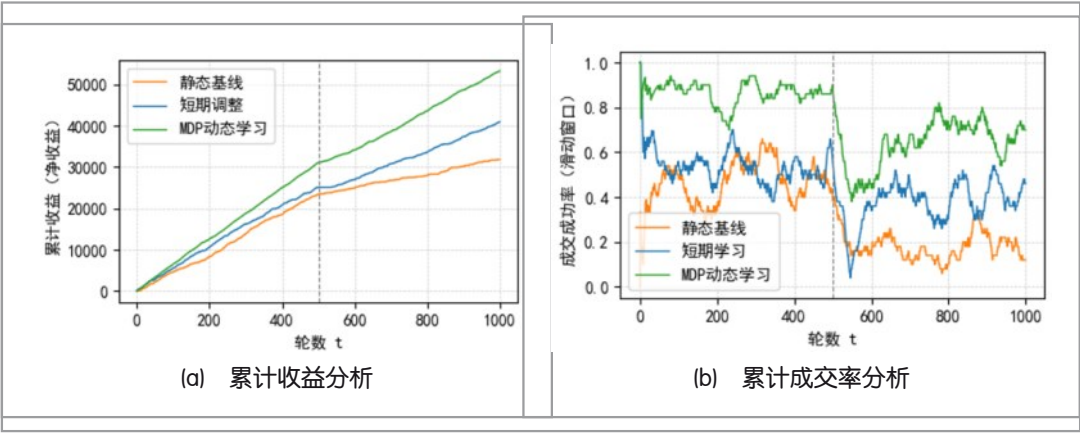


图4 SC-TDP的优化效果测试

如图4(a)所示，MDP策略引入了显著且持续的收益优势。实验初期（冷启动阶段），各策略因处于尝试期，累计收益曲线几乎重合。随轮次增加，MDP与短期调整策略迅速超越静态基线，且MDP与基线的差距呈单调扩大趋势。在t=500发生市场漂移后，静态基线收益曲线迅速趋平，短期调整策略虽维持一定增长，但MDP策略保持了更陡峭的斜率。这表明在治理层监控下，系统通过捕捉CDI越限信号主动触发宏观校准，使模型迅速摆脱旧环境束缚，

从而在非平稳市场中获得更高的长期收益。

图4(b)显示，冷启动阶段各策略CR剧烈波动。静态策略在t≈100时达到峰值后，因缺乏适应力逐渐下降并稳定在低位。短期调整策略因贪婪规则陷入成交即提价、失败即降价的短视循环，长期震荡于0.5左右。相比之下，MDP策略凭借全局搜索跳出局部陷阱，将成功率稳定在0.8以上。在t=500引入漂移瞬间，各策略CR断崖式下跌。静态基线停滞于0.2，短期策略徘徊在中等水平。MDP策略虽遭遇最深回

撤，但随即展现强劲的 V 型反转并回升至 0.8。该现象并非模型失效，而是治理层应对漂移支付的必要探索代价。通过主动重置参数，系统强制模型跳出局部最优，以短期波动换取了对新环境规律的快速掌握，实现了更优的自愈与价值捕获。

4.3 模型治理下的去中心化信任保障验证

本节从系统的鲁棒性和成本优势两组核心实验来验证 SC-TDP 的链上定价引擎在复杂、对抗和资源受限场景下的可靠性与工程可行性。

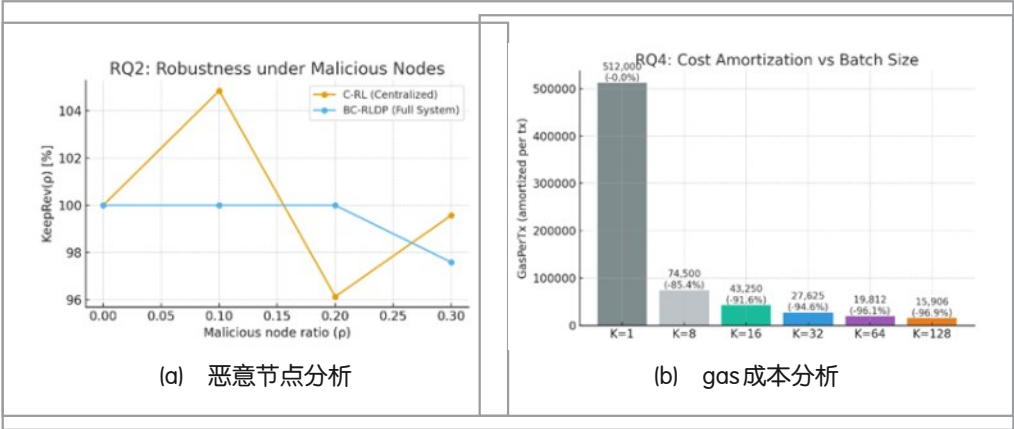


图5 SC-TDP的优化效果测试

抵御恶意节点攻击的鲁棒性验证通过引入不同比例 {0%，10%，20%，30%} 的恶意节点 p 模拟外部攻击，验证 SC-TDP 的去中心化信任机制在抵御训练样本污染方面的有效性。图 5(a) 显示，当 $p \leq 0.20$ 时，SC-TDP 利用共识机制保持了 100% 的收益保持率。相比之下，C-RL 在 $p=0.20$ 时保持率降至 96.1%。值得注意的是，C-RL 在 $p=0.10$ 时的反常收益提升源于虚假高反馈引发的过度乐观偏差，导致模型误选高风险策略并在攻击加剧后迅速崩溃。这一结果充分验证了 SC-TDP 模型的去中心化架构是保障定价策略在对抗环境中安全演进和维持收益稳定的关键。

图 5(b) 展示了不同批量大小 K 下的 Gas 成本摊销效果。相对于实时更新，批量更新显著降低了人均成本：K=8 时成本降低

85.4%，K=32 时降低 94.6%，K=128 时节约率达 96.9%。结果表明，QCC 的双阈值批量更新策略成功将固定开销摊销至多个样本，在保障持续学习的同时实现了链上系统长期运营的经济性与工程可行性。

5 结束语

本文针对传统数据资产定价机制中存在的动态适应性不足与节点可信性低的双重挑战，提出并实现了一种基于区块链和强化学习的动态定价模型（SC-TDP）。该模型通过将马尔可夫决策过程（MDP）与区块链技术相结合，构建了一个去中心化、透明且自动化的数据资产定价闭环。

本文的主要贡献可以总结为：

(1) 提出了一个完整的链上定价引擎架构。该架构由价值评估 (VEC)、定价执行 (PEC)、交易反馈 (TFC) 和 Q 值计算 (QCC) 四个核心智能合约组成, 在链上实现了强化学习从状态感知、决策、反馈到策略更新的完整闭环逻辑。(2) 设计了应对市场非平稳性的去中心化治理机制。(3) 通过引入概念漂移指数 (CDI) 与阈值触发规则, 使系统能够在不牺牲去中心化信任的前提下, 安全、可审计地演进策略版本, 保持对市场变化的自适应能力。(4) 实现了成本感知的自适应批量更新策略。针对链上高昂的 Gas 成本, 本模型为 QCC 设计了数据驱动和时间驱动的双重触发机制, 有效摊销了链上成本, 在策略同步延迟与工程可行性之间取得了动态平衡。实验结果表明, 本模型在模拟交易环境和真实数据集上的验证中, 均表现出较高的适应性、精确度和策略稳健性, 验证了该方案的有效性与工程可行性。

今后的工作可聚焦于提升模型的泛化能力与安全性, 例如将该框架扩展至数据捆绑或 API 服务等复杂形态, 并研究融合零知识证明 (ZKP) 等隐私技术以保护链上策略的机密性。此外, 深度学习算法的扩展性, 引入深度强化学习 (DRL) 以克服当前 Q-learning 在面对高维状态空间时的扩展性瓶颈, 也将是提升模型决策精度的关键方向。

参考文献:

[1] 中国信息通信研究院. 大数据白皮书(2020年). 2020. http://www.caict.ac.cn/kxyj/qwfb/bps/202012/t20201228_367162.htm.

[2] Nguyen T. L., Nguyen L., Hoang T., Bandara D., Wang Q., Lu Q., ... & Chen,

S. (2025). Blockchain-empowered trustworthy data sharing: Fundamentals, applications, and challenges[J]. ACM Computing Surveys, 2025, 57(8): 1-36.

[3] Jiang K., Liang W., Zou J. Y., & Kwon Y. (2023). Opendataval: a unified benchmark for data valuation[C]//Advances in Neural Information Processing Systems. 36: 28624-28647.

[4] Xu C. Z., Zhu K., Yi C. Y., Wang R. (2020). Data pricing for Blockchain-based car sharing: A Stackelberg game approach[C]//Proc. of IEEE Global Communications Conference (GLOBECOM 2020). Taipei: IEEE, 2020: 1-5.

[5] Korenkevych D., Cheng F., Balakir A., Nikulkov A., Gao L., Cen Z., ... & Zhu, Z. (2024). Offline reinforcement learning for optimizing production bidding policies [C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2024). Barcelona, Spain: ACM, 2024: 5251-5259.

[6] Kodadi S. Integrating blockchain with database management systems for secure accounting in the financial and banking sectors[J]. Journal of Science and Technology, 2023, 8(9): 9-27.

[7] Sober M., Scaffino R., et al. (2023). A Blockchain-Based IoT Data Marketplace [J]. Cluster Computing, 2023, 26(6): 3523-3545.

[8] WANG R Y, STRONG D M. Beyond accuracy: What data quality means to data consumers [J]. J. Manage. Inf. Syst., 1996:12 (4):5-33.

[9] WANG J, LIU Y, LI P, et al. Overview of data quality: Examining the dimensions, antecedents, and impacts of data quality [J]. J. Knowl. Econ., 2024:15 (1):1159-1178.

- [10] Zhang D, Wang HZ, Ding XO, Zhang YC, Li JZ, Gao H. On the fairness of quality-based data markets. arXiv: 1808.01624, 2018.
- [11] WANG J, LIU Y, LI P, et al. Overview of data quality: Examining the dimensions, antecedents, and impacts of data quality [J]. *J. Knowl. Econ.*, 2024; 15(1): 1159–1178.
- [12] KOUTRIS P, UPADHYAYA P, BALAZINSKA M, et al. Query-based data pricing[J]. *J. ACM*, 2015;62(5):1–44.
- [13] Koutris P, Upadhyaya P, Balazinska M, Howe B, Suciu D. Query-based data pricing. In: *Proc. of the 31st ACM SIGMOD-SIGACT SIGAI Symp. on Principles of Database Systems*. Scottsdale: ACM, 2012. 167 – 178.
- [14] Jia R., et al. Efficient Task-Specific Data Valuation for Nearest Neighbor Algorithms. *PVLDB*, 12(11), 2019: 1610 – 1623.
- [15] Mei LJ, Li W, Nie K. Pricing decision analysis for information services of the Internet of Things based on Stackelberg game. In: Zhang ZJ, Zhang RT, Zhang JL, eds. *LISS 2012*. Berlin, Heidelberg: Springer, 2013. 1097 – 1104.
- [16] Liu K, Qiu XY, Chen WH, Chen X, Zheng ZB. Optimal pricing mechanism for data market in Blockchain-enhanced internet of things. *IEEE Internet of Things Journal*, 2019, 6(6): 9748 – 9761.
- [17] Xu CZ, Zhu K, Yi CY, Wang R. Data pricing for Blockchain based car sharing: A Stackelberg game approach. In: *Proc. of 2020 IEEE Global Communications Conf. Taipei: IEEE*, 2020. 1 – 5.
- [18] Kang X, Zhang R, Motani M. Price-based resource allocation for spectrum-sharing femtocell networks: A Stackelberg game approach. *IEEE Journal on Selected Areas in Communications*, 2012, 30(3): 538 – 549.
- [19] Yao HP, Mai TL, Wang JJ, Ji Z, Jiang CX, Qian Y. Resource trading in Blockchain-based industrial Internet of Things. *IEEE Transactions on Industrial Informatics*, 2019, 15(6): 3602 – 3609.
- [20] Rawat DB, Shetty S, Xin CS. Stackelberg-game-based dynamic spectrum access in heterogeneous wireless systems. *IEEE Systems Journal*, 2016, 10(4): 1494 – 1504.
- [21] Jung K, Park S. Privacy bargaining with fairness: Privacy-price negotiation system for applying differential privacy in data market environments. In: *Proc. of the 2019 IEEE Int'l Conf. on Big Data*. Los Angeles: IEEE, 2019. 1389 – 1394.
- [22] McAfee RP. A dominant strategy double auction. *Journal of Economic Theory*, 1992, 56(2): 434 – 450.
- [23] Chen LJ, Koutris P, Kumar A. Towards model-based pricing for machine learning in a data marketplace. In: *Proc. of the 2019 Int'l Conf. on Management of Data*. Amsterdam: ACM, 2019. 1535 – 1552.
- [24] Tang RM, Wu HY, Bao ZF, Bressan S, Valduriez P. The price is right-models and algorithms for pricing data. In: *Proc. of the 24th Int'l Conf. on Database and Expert Systems Applications*. Prague: Springer, 2013. 380 – 394.
- [25] XU L,JIANG C X,QIAN Y,et al.Dynamic privacy pricing: a multi-armed bandit approach with time variant rewards[J]. *IEEE Transactions on Information Forensics and Security*, 2017, 12(2): 271–285.
- [26] BAUER J,JANNACH D.Optimal pricing in e-commerce based on sparse and noisy data [J]. *Decision Support Systems*, 2018,106:53–63.
- [27] Tsai Y C,Cheng Y D,Wu C W,et al.Time-

dependent smart data pricing based on machine learning[C]//Canadian Conference on Artificial Intelligence. Springer. Cham, 2017:103-108.

[28] 唐兴叶. 基于文本分析与机器学习的顾客感知价值定价模型[D]. 厦门:厦门大学,2019.

[29] Sun C Y,Gong L T.Research on interest rate pricing under the big data thinking: an empirical analysis based on machine learning[J]. Fin. Theory Pract, 2017, 18(7): 1-5.

[30] 倪渊,李子峰,张健. 基于 AGA-BP 神经网络的网络平台交易环境下数据资源价值评估研究[J]. 情报理论与实践,2020,43(01):135-142.

收稿日期: XXXX-XX-XX

通信作者: xiaqi@uestc.edu.cn

基金项目: 国家自然科学基金资助项目(No. U22B2029)

Foundation Items: The National Natural Science Foundation of China (No. U22B2029)