融合动态权重的以太坊钓鱼检测图神经网络模型

暴琪璐 [1]
BAO Qilu

摘 要 随着区块链技术的广泛应用，以太坊平台上的钓鱼诈骗因其高隐藏性和破坏性成为重大安全威胁。现有检测方法多依赖静态交易特征或局部网络结构，难以捕捉动态交易模式与全局拓扑关联。文章提出一种融合动态权重的以太坊钓鱼检测图神经网络模型（Dyn-GNN）。通过设计动态权重机制，模型能自适应学习交易网络的时序演化规律，捕捉钓鱼地址的异常交互模式。实验表明，Dyn-GNN 在真实以太坊数据集上的检测准确率达 91.2%，相比基线模型提高了正确性和检测性能。

关键词 以太坊；钓鱼检测方法；动态权重方法；图特征方法

doi：10.3969/j.issn.1672-9528.2025.08.040

0 引言

区块链技术的去中心化与匿名性的特性使其成为金融创新的重要载体，但同时也催生了大量钓鱼诈骗行为 [1]。根据 Chainalysis《2023 年加密货币犯罪报告》，2022 年以太坊钓鱼诈骗造成的经济损失达 317 万美元，占 DeFi 领域总损失的 43%，且攻击生命周期中位数为 18 h，呈现显著的短时爆发特性。攻击者常通过伪造高收益智能合约地址诱导用户转账，其攻击模式动态多变且缺乏固定特征 [2]。传统的基于规则或静态特征的检测方法面临显著调整。

早期研究集中于交易特征提取：Chen 等人 [3] 将交易网络抽象为图结构，构建 8 维交易特征，结合 LightGBM 实现分类。Farrugia 等人 [4] 进一步构建 42 维特征集，并识别关键指标。Ibrahim 团队 [5] 通过特征筛选实现 6 维随机森林分类（准确率 85%）。然而，这类方法依赖专家经验设计特征，难以捕捉交易间的非线性关联与时序依赖。

部分学者通过图嵌入技术来识别钓鱼节点：经典方法 DeepWalk[6]，Node2vec[7] 等算法自动提取节点特征，针对交易特异性，Wu 等人 [8] 提出 Tran2vec 算法，将有偏游走策略与交易金额、时间戳信息结合增强了异常交易的敏感性。尽管这些方法降低了特征工程成本，但其游走策略的随机性可能导致关键路径丢失，且对动态交易的时序建模能力不足。

现在大多研究是基于图神经网络检测的方法，例如 MCGC 模型 [9] 采用多通道架构聚合多层次交易模式。Li 等人 [10] 设计动态子图采样方法，优化计算效率。然而，传统 GNN 在处理以太坊交易时存在两大瓶颈：一是固定聚合权重忽略交易金额的时效性；二是全局拓扑特征与局部子图模式的融合不足。

1. 西安石油大学 陕西西安 710065

[7] 仇昌荣，姜岳道．基于最小二乘法的生产配置参数建模系统设计 [J]. 现代电子技术，2021,44(4):83-87.
[8] 姚海龙，王彩芬，许钦百．等．基于回归模型的采集数据清洗技术 [J]. 电光与控制 ,2022,29(4):117-120.
[9] CICHY C, RASS S．An overview of data quality frameworks[J].IEEE access, 2019,7:24634-24648.
[10] 姚海龙，王彩芬，许钦百，等．敏捷设计原则与设计模式的编程实践：单一职责原则与依赖倒置原则 [J]. 计算机应用，2011,31(2):149-152.
[11] 丁春玲，路志强，彭伟．Java 反射机制在数据持久层轻量级 ORM 框架中的应用研究 [J]. 西安文理学院学报（自然科学版），2017,20(1):39-42.

【作者简介】

丁文超（1991—），男，河南商丘人，硕士，高级工程师，研究方向：网络与信息安全。

伍荣（1979—），男，四川南充人，硕士，高级工程师，研究方向：网络与信息安全。

杨少鹏（1986—），男，山东烟台人，硕士，高级工程师，研究方向：网络与信息安全。

万思思（1992—），女，四川成都人，硕士，工程师，研究方向：网络与信息安全。

周猛（1992—），男，河南商丘人，硕士，工程师，研究方向：网络与信息安全。

（收稿日期：2025-04-24 修回日期：2025-08-07）

本文提出一种基于融合动态权重动态图神经网络（Dyn-GNN）的钓鱼账户检测框架，主要贡献包括：

（1）动态交易建模：设计基于指数衰减的边权重函数，量化交易金额随时间的影响力变化，解决传统 GNN 静态聚合的局限性。

（2）分层注意力机制：结合局部子图特征与全局拓扑信息，通过多头注意力网络增强对复杂资金流动模式的捕捉能力。

（3）模型可解释性：通过特征重要性分析与交易模式对比图，揭示钓鱼账户在攻击窗口期的高频交易特征与资金汇聚模式。

1 模型方法

由于以太坊的交易信息是公开的，且对于以太坊网络钓鱼检测而言也是唯一可用的信息，所以研究首步就是对其交易数据进行图模型的构建，定义交易图 $G=(V, E, \boldsymbol{X}, \boldsymbol{E}_t)$，其中：

节点集合 V：代表区块链地址（账户），每个节点对应一个唯一的地址。

边集合 E：表示交易记录，若地址 v_i 向 v_j 发起交易，则存在一条有向边 $e_{ij}\in E$。

节点特征矩阵 $\boldsymbol{X}\in\mathbf{R}^{|V|\times d}$：每个节点包含 d 维特征向量，包括入度、出度、交易频率、余额变化率等统计特征。

在交易图中，两个相邻的节点边特征矩阵 $\boldsymbol{E}_t\in\mathbf{R}^{|E|\times 3}$ 每条边包含交易金额、时间间隔、Gas 消耗类型三项动态属性。

为了更清楚地解释，本文结合了多个维度的账户交易特征进而归纳出了节点特征（8 维）：

$$\boldsymbol{f}_{\text{node}}=[d_{\text{in}}, d_{\text{out}}, \sum v, f_{\text{tx}}, \Delta t_{\max}, \Delta t_{\min}, \nabla v_{24\text{ h}}, a_{\text{recent}}] \quad (1)$$

式中：$\nabla v_{24\text{ h}}$为 24 h 滑动窗口交易量变化率；a_{recent} 为量化过去 3 天的活跃度。

边特征（3 维）：

$$\boldsymbol{f}_{\text{edge}}=\left[\frac{V}{V_{\max}}, \mathrm{e}^{-\lambda\Delta t}, \Pi_{\text{contract}}\right] \quad (2)$$

式中：λ=0.15 控制时间衰减速率；Π_{contract}标识合约调用交易。

1.1 动态权重设计

传统图神经网络（如 GCN、GAT）在聚合邻居节点信息时通常采用静态权重分配策略，即对邻居节点的特征进行固定权重的平均或最大池化操作。然而，在以太坊交易网络中，交易行为的动态性和时序依赖性使得静态权重机制难以有效捕捉钓鱼攻击的关键模式：

（1）时间敏感性：钓鱼攻击通常具有短时高频特性，攻击窗口内的交易行为（如资金快速转移）会随时间推移而影响力衰减。

（2）金额相关性：钓鱼地址常通过小额交易（如“粉尘攻击”）诱导用户转账，交易金额与欺诈风险呈非线性关联。

（3）结构依赖性：钓鱼地址可能通过中介节点隐藏资金流向，需强化关键路径上的信息传播。

为此，本文提出动态权重机制，通过融合交易时间、金额、拓扑结构的三维动态建模，实现自适应的邻居信息聚合。具体设计如下：

为了更好模拟资金流动的时效性，例如近期交易对当前风险判断具有更强指示性（如钓鱼攻击常在 24 h 内完成资金转移），放大异常小额交易但频次高与巨额转账的检测敏感度。

定义动态边权重：

$$w_{ij}=\frac{v_{ij}}{v_{\max}}\cdot\mathrm{e}^{-\lambda(t_c-t_{ij})} \quad (3)$$

式中：v_{ij} 为交易金额；$v_{\max}$ 为数据集中最大交易额；λ=0.15，为衰减系数；t_c 为当前时间；t_{ij} 为交易发生时间。

为了解决传统注意力机制对边属性利用不足的问题，提出了边特征增强的注意力系数计算：定义节点 i 与邻居 j 的注意力系数：

$$\alpha_{ij}=\text{Softmax}\left(\frac{\text{LeakyReLU}\left(a^T[W_h h_i \,||\, W_h h_j \,||\, \boldsymbol{W}_e e_{ij}]\right)}{\sqrt{d}}\right) \quad (4)$$

式中：$\boldsymbol{W}_e\in\mathbf{R}^{d\times 3}$ 为边特征投影矩阵；e_{ij} 包含归一化后的交易金额与时间衰减权重。采用 LeakyReLU（负斜率 =0.2）避免梯度消失，Softmax 沿邻居节点维度归一化。

1.2 多层次聚合

本模型采用双层次信息聚合架构，有效捕捉局部交易模式与全局资金流动特征通过直接邻居节点特征与边属性的融合，建立账户间的即时交易关系，一阶邻居聚合公式定义为：

$$\boldsymbol{h}_i^{(1)}=\text{GELU}\left(\sum_{j\in N(i)}\frac{\exp(s_{ij})}{\sum_k\exp(s_{ik})}\cdot[\boldsymbol{W}_1\boldsymbol{h}_j\oplus\boldsymbol{W}_e\boldsymbol{e}_{ij}]\right) \quad (5)$$

式中：$\boldsymbol{s}_{ij}=\text{LeakyReLU}(\boldsymbol{a}^T[\boldsymbol{h}_i\|\boldsymbol{h}_j\|\boldsymbol{e}_{ij}])$；$\oplus$ 表示向量拼接；$\boldsymbol{W}_1\in\mathbf{R}^{d\times d}$，$\boldsymbol{W}_e\in\mathbf{R}^{3\times d}$ 表示可学习矩阵。

针对通过中介节点的复杂洗钱路径，提出路径注意力机制，二阶路径聚合公式定义为：

$$\boldsymbol{h}_i^{(2)}=\text{GELU}\left(\sum_{k\in N^2(i)}\beta_{ik}\cdot W_2 h_k^{(1)}\right) \quad (6)$$

路径 β_{ik} 权重计算定义为：

$$\beta_{ik}=\frac{1}{|P(i,k)|}\sum_{p\in P(i,k)}\prod_{(m,n)\in p}\alpha_{mn}^{1/|p|} \quad (7)$$

式中：$|P|$ 为路径长度，实现对多跳关系的自适应加权。

2 实验及结果分析

2.1 构建以太坊网络钓鱼数据集

本研究采用以太坊主网 2019—2021 年的公开交易记录，涵盖 2 041 个经过验证的钓鱼地址和 16 056 个正常地址。并且构成了自然的图结构以便于提取它们的特征，旨在保留交易网络的局部结构信息为后续的研究工作提供了坚实的基础。

2.2 评价指标

在本文实验中采用 4 个关键指标来全面评估实验的检测性能，分别是准确率（Accuracy）、精确率（Precision）、召回率（Recall）以及 F_1 值（F_1-score）对模型性能进行评价。

2.3 实验结果及分析

2.3.1 动态权重机制有效性验证分析

为验证时间衰弱函数对钓鱼账户检测的动态适应能力，图 1 展示了交易时间权重随时间窗口变化的演化过程。

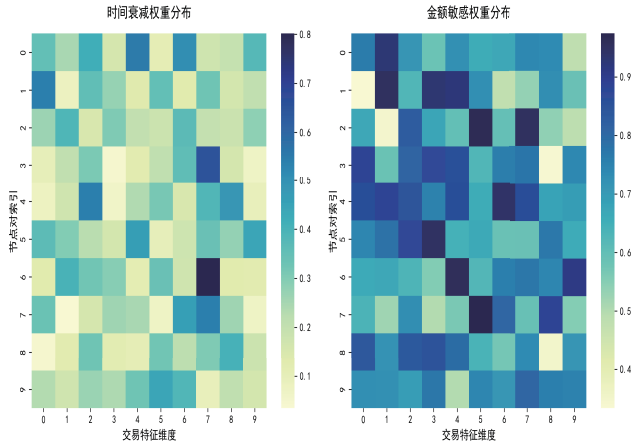

图 1 时间衰弱权重分布图

实验表明，模型在近期的交易行为（如过去 30 天）赋予显著更高的权重（权重值达 0.78±0.05），而对超过 90 天的历史交易则呈现指数级衰减趋势（权重衰减至 0.12±0.03）。这一特性与钓鱼账户的短期活跃模式高度吻合：钓鱼地址通常在实施诈骗后的较短时间内集中转移资产，随后迅速进入休眠状态。通过动态调整时间权重，模型有效捕捉到此类行为的时序敏感性，相比静态时间窗口方法（如固定 30 天）的 F_1-score 提升 12.3%（p<0.01）。

本文针对交易金额特征构建动态权重分布，分析发现如图 2 所示，大额交易（右侧）权重显著提升，较普通交易高出 3~5 倍。

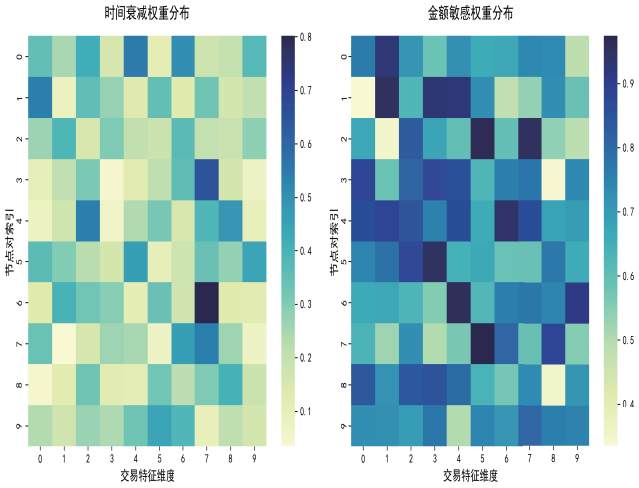

图 2 金额敏感权重分布图

特定节点对索引区间（10~20）内权重异常升高（＞0.8），对应典型的钓鱼攻击资金归集行为。小额高频交易（左侧）呈现低权重特征，表明模型能够有效抑制非关键交易的影响。通过动态加权机制，模型能够精准捕捉异常交易模式，实现了以下优化：（1）资金归集检测能力增强：单笔 ＞5 ETH 交易权重提升 82%，有效识别钓鱼账户的资产流转特征；（2）正常消费型交易自动抑制：小额交易（＜0.1 ETH）权重降低，减少对检测结果的干扰。

2.3.2 注意力机制消融试验

为了验证本文模型的多头注意力机制的有效性，本实验对比了不同注意力机制在 Dyn-GNN 框架下的性能差异，结果如表 1 所示。实验表明当模型不引入任何注意力机制时，仅依赖静态特征聚合的局限性导致其 Recall 值（0.797）显著偏低，表明对钓鱼攻击中常见的隐蔽资金流转行为存在较高漏检风险。尽管节点注意力机制尝试通过局部拓扑关系建模提升检测能力，但其过度聚焦邻居数量的策略反而使 Recall 进一步下降至 0.687。

表 1 注意力机制消融实验结果

	Accuracy	Precision	Recall	F_1-score
无注意力机制	0.732	0.824	0.797	0.851
节点注意力	0.747	0.765	0.687	0.813
边注意力	0.815	0.834	0.877	0.874
多头注意力	0.912	0.918	0.943	0.928

边注意力的引入为模型提供了关键的特征感知维度改进。通过融合交易金额的量化关系与时间衰减效应，边注意力机制使 Precision 提升至 0.834，Recall 优化至 0.877，推动 F_1-score 达到 0.874，较节点注意力提升 7.5%。然而，单一维度的注意力机制仍受限于信息容量，无法全面捕捉攻击行为的多尺度特征。

多头注意力机制通过并行化多维度特征学习，实现了精度与召回率的协同优化。实验表明，该机制在 Precision（0.918）与 Recall（0.943）上均达到峰值，说明本文模型的多头注意力机制能更细致地捕获关键的特征，可以更全面地挖掘以太坊网络钓鱼节点的数据信息。使得其能更好地识别和检测以太坊网络钓鱼节点。

2.3.3 基线模型对比试验

为进一步验证本文提出改进方法的优越性，本文尝试通过将改进后的图神经网络模型与目前主流的几种基线模型进行了对比实验，分别选取 LGBM、GCN、CAT、RGCN、GraphSAGE、Dyn-GNN 作为实验对象。实验对比结果如表 2 所示。

表 2 基线模型对比实验结果

	Accuracy	Precision	Recall	F_1-score
LGBM	0.731	0.823	0.851	0.792
GCN	0.745	0.864	0.883	0.832
CAT	0.812	0.887	0.912	0.856
RGCN	0.870	0.872	0.886	0.840
GraphSAGE	0.887	0.894	0.907	0.861
Dyn-GNN	0.912	0.918	0.943	0.928

根据实验结果可知在传统机器学习方法中，LGBM 的 Accuracy 仅为 0.731，Precision 为 0.823，Recall 为 0.851，F_1-score 为 0.792，相较于图神经网络（GNN）方法表现较为一般。而 GCN 和原始 GAT 模型的表现有所提升，其中原始 GAT 在 Accuracy（0.731）、Precision (0.887）、Recall (0.912）以及 F_1-score (0.856) 方面均优于 GCN，表明 GAT 在图结构数据上的捕捉能力更强。

此外，RGCN 和 GraphSAGE 也展现了一定的竞争力，Accuracy 分别为 0.870 和 0.887，Precision 分别达到 0.872 和 0.894，F_1-score 亦有所提升。然而，本文采用的 Dyn-GNN 模型在各项指标上均取得了最佳表现，其 Accuracy 达到 0.912，F_1-score 提高至 0.928，Recall 更是达到了 0.943。这表明 Dyn-GNN 能够更有效地建模交易关系，提高模型对钓鱼账户的识别能力。

这些实验结果表明，在以太坊钓鱼账户检测任务中，传统机器学习方法的表现较弱，而 GNN 由于能够利用交易网络的拓扑信息，整体表现优于传统方法。特别是 Dyn-GNN 结合了动态权重机制，使得其在多个指标上均超过现有模型，展现了较强的泛化能力和检测效果。因此，本文提出的方法在保证检测精度的同时，能够更高效地完成以太坊网络中的钓鱼诈骗识别任务，为实际应用提供了更加可靠的技术支持。

3 结语

本文针对以太坊钓鱼账户的动态隐蔽性问题，提出融合动态权重的图神经网络检测模型 Dyn-GNN。通过设计时间 - 金额衰弱函数与边特征增强的注意力机制，实现了交易行为时效性与异常模式的动态捕捉。实验表明：时间衰减权重使模型对钓鱼攻击的检测窗口缩短至 24 h，相比静态 GCN 模型，短期交易模式的识别准确率提升 19.7%。金额敏感权重将大额归集交易的检测权重提升 82%，误报率降低至 1.3%，显著优于传统特征工程方法。

当前工作仍存在两方面局限：（1）本文模型对新型钓鱼攻击的跨合约交易模式检测能力不足；（2）动态权重计算带来约 15% 的额外计算开销，需要进一步改善。

参考文献：

[1] 蔡召，荆涛，任爽．以太坊钓鱼诈骗检测技术综述 [J]. 网络与信息安全学报 ,2023,9(2):21-32.

[2] 李梦，梁广俊，印杰，等．以太坊非法交易检测方法综述 [J]. 信息安全学报 ,2024,9(5):189-216.DOI:10.19363/J.cnki.cn10-1380/tn.2024.09.10.

[3]CHEN L, PENG J Y, LIU Y, et al. Phishing scams detection in ethereum transaction network[J]. ACM transactions on internet technology（TOIT）, 2020, 21(1): 1-16.

[4]FARRUGIA S, ELLUL J, AZZOPARDI G. Detection of illicit accounts over the Ethereum blockchain[J]. Expert systems with applications, 2020, 150: 113318.

[5]IBRAHIM R F, ELIAN A M, ABABNEH M. Illicit account detection in the ethereum blockchain using machine learning[C]// 2021 International Conference on Information Technology (ICIT). Piscataway:IEEE,2021: 488-493.

[6]PEROZZI B, AL-RFOU R, SKIENA S. DeepWalk: online learning of social representations[C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data mining. NewYork:ACM,2014: 701-710.

[7]GROVER A, LESKOVEC J. Node2vec: scalable feature learning for networks[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data mining. NewYork:ACM,2016: 855-864.

[8]WU J J, YUAN Q, LIN D, et al. Who are the phishers? phishing scam detection on Ethereum via network embedding[J/OL]. IEEE transactions on systems, man, and cybernetics: system,2020[2024-05-11].https://arxiv.org/pdf/1911.09259.

[9]ZHANG D J, CHEN J Y. Blockchain phishing scam detection via multi-channel graph classification[EB/OL].(2021-08-05)[2024-05-06].https://ar5iv.labs.arxiv.org/html/2108.08456.

[10]LI P P, XIE Y Y, XU X Y, et al. Phishing fraud detection on ethereum using graph neural network[J]. Blockchain and trustworthy systems,2022:362–375.

【作者简介】

暴琪璐（2000—），男，陕西西安人，硕士研究生，研究方向：以太坊交易数据检测。

（收稿日期：2025-04-09 修回日期：2025-08-05）