面向加密货币反洗钱的智能监管研究：模型、方法与应用

张昳斐[1]， 袁勇[1]， 杨东[2]， 王飞跃[3]
（1. 中国人民大学数学学院，北京 100872；
2. 中国人民大学法学院，北京 100872；
3. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室，北京 100190）

摘 要：近年来，基于区块链技术的加密货币市场正在快速发展，新兴商业模式和应用场景不断涌现，形成了数万亿美元规模的新经济体系。加密货币具有去中心化、匿名性和易于跨境流通等特性，极大地丰富和发展了以法定货币为核心的主流金融体系，同时也被广泛应用于洗钱等非法金融活动。因此，面向加密货币反洗钱的监管研究成为新兴热点领域，而机器学习和人工智能则成为加密货币交易行为监管、交易网络分析、安全风险评估等领域的主要研究方法。首先，系统性地梳理了近年来加密货币反洗钱的智能监管研究进展，从传统机器学习、深度学习、集成学习、图分析和启发式方法等维度归纳总结了反洗钱监管研究的主要模型、方法和应用模式。同时，整理了加密货币反洗钱研究文献中的常用数据集，并基于Elliptic数据集给出现有模型和算法的性能对比。最后，讨论了加密货币反洗钱的研究挑战与未来发展方向，以期为加密货币和去中心化金融产业的繁荣与健康发展提供有益的参考和借鉴。

关键词：加密货币；反洗钱；区块链；机器学习；人工智能
中图分类号：F831
文献标志码：A
doi: 10.11959/j.issn.2096−6652.202522

Intelligent regulation for cryptocurrency anti-money laundering: models, methods, and applications

ZHANG Yifei[1], YUAN Yong[1], YANG Dong[2], WANG Fei-Yue[3]
1. School of Mathematics, Renmin University of China, Beijing 100872, China
2. School of Law, Renmin University of China, Beijing 100872, China
3. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Abstract: In recent years, blockchain-enabled cryptocurrency market has been rapidly evolving to a trillion-dollar economic system with various kinds of emerging business models and application scenarios. Cryptocurrencies, characterized by decentralization, anonymity, and ease of cross-border circulation, have significantly enriched and evolved the mainstream financial system centered around fiat currencies. However, they are also extensively used in illegal financial activities including money laundering. As such, research on anti-money laundering (AML) regulation for cryptocurrencies has become a burgeoning field of interest, with machine learning and artificial intelligence (AI) emerging as primary methodologies in areas such as cryptocurrency transaction monitoring, transaction network analysis, and security risk assessment. First, the recent research progress on intelligent AML regulation for cryptocurrencies was systematically reviewed. The

收稿日期：2025−03−04；**修回日期：**2025−05−06
通信作者：袁勇，yong.yuan@ruc.edu.cn
基金项目：国家重点研发计划项目（No.2022YFB2720401）；国家自然科学基金项目（No.72171230）；北京市未来区块链与隐私计算高精尖中心项目
Foundation Items: The National Key Research and Development Program of China (No.2022YFB2720401), The National Natural Science Foundation of China (No.72171230), Project from Beijing Advanced Innovation Center for Future Blockchain and Privacy Computing

main models, methods, and application patterns in AML regulatory research were summarized from multiple dimensions, including traditional machine learning, deep learning, ensemble learning, graph analysis, and heuristic approaches. Meanwhile, commonly used datasets in cryptocurrency AML research were compiled, and a performance comparison of existing models and algorithms was provided based on the Elliptic dataset. Finally, the research challenges and future directions of cryptocurrency AML were discussed, aiming to offer valuable references and insights for the prosperity and healthy development of the cryptocurrency and decentralized finance industry.

Key words: cryptocurrency, anti-money laundering, blockchain, machine learning, artificial intelligence

0 引言

加密货币市场近年来快速发展，投资者数量呈爆炸式增长，新兴商业模式和应用场景不断涌现，已成为全球经济和金融系统的重要组成部分。2024 年 12 月，全球加密货币市场总市值超过 3.91 万亿美元，加密货币币种已超过 1 万种，其中比特币和以太坊（Ethereum）市值占比分别为 53.6% 和 11.8%。加密货币市场的繁荣得益于其去中心化、匿名性、易于跨境流通等特性，也极大地丰富和发展了以法定货币为核心的主流金融体系[1-2]。

现阶段，加密货币的市场监管体系尚不完善，其去中心化、共识信任、强匿名性、分布式存储、低交易成本、易于跨境流通等特点，使其成为洗钱等非法金融活动的重要选择。统计数据显示，2016 年至 2023 年，全球加密货币洗钱规模总计 1 477 亿美元，且 2020 年以来洗钱金额以每年 67% 的速度快速增长。从丝绸之路的黑市交易到跨境网络赌博，从欺诈性的“杀猪”骗局到毒品交易等，加密货币洗钱的方法和手段也逐渐多样化。因此，国际上相关监管机构正在不断加强对加密货币交易的监管，以应对这一严峻挑战。

加密货币洗钱的主要监管机构一般包括各国的金融监管机构、反洗钱机构以及国际组织。例如，中国监管机构主要包括中国人民银行、国家互联网信息办公室、工业和信息化部、公安部等，这些部门协同合作，共同打击与加密货币洗钱相关的非法金融活动。美国的证券交易委员会、商品期货交易委员会和金融犯罪执法网络分别负责加密货币的证券属性监管、商品属性监管和反洗钱监管。欧盟则通过欧洲银行管理局和欧洲证券及市场监督管理局等机构实施统一监管框架。国际层面，金融行动特别工作组制定了全球反洗钱标准，要求加密货币服务提供商遵循严格的客户尽职调查和反洗钱措施。

为应对洗钱等非法金融活动带来的风险，各国均加强了对加密货币的监管力度。2024 年 12 月，欧盟发布《加密资产市场监督法案》，成为全球首个实施统一加密货币监管框架的地区。英国计划于 2026 年制定全面的加密监管制度。2025 年 1 月，美国总统特朗普签署行政命令，计划成立数字资产市场工作组，以制定管理数字资产的联邦监管框架，管理美国数字资产的发行和运营等。日本则通过《支付服务法》和《金融工具和交易法》对加密货币交易所进行严格监管，要求其获得运营许可并满足资金和信息安全要求。我国则一直对加密货币持谨慎态度，全面禁止私人加密货币交易和“挖矿”活动，明确加密货币相关业务活动属于非法金融活动。同时，我国还修订了《中华人民共和国反洗钱法》，进一步强化了对加密货币洗钱行为的打击。

然而，尽管各国和国际组织通过多维度的监管政策不断加强对加密货币洗钱的打击力度，但加密货币的特性仍带来了诸多难以克服的技术挑战，使传统金融领域的洗钱监管方法难以有效应对。加密货币洗钱的技术挑战主要体现在其匿名性、快速流动性、跨境交易的复杂性以及数据的海量与类别不平衡等方面。具体来说，加密货币的匿名性使交易双方身份难以确定，混币技术则进一步增加了资金流向的追踪难度。加密货币的去中心化、快速流动性与无国界性使洗钱资金能够在短时间内跨境转移，不同国家的监管政策差异与技术手段错配极大地增加了洗钱监管的难度。此外，加密货币市场的海量、高频数据和高度类别不平衡（合法交易与洗钱交易的比例悬殊）使传统金融的洗钱监管方法失效。这些技术挑战促使研究者转向利用人工智能和机器学习等方法，通过实时监测、模式识别和风险评估，提高对洗钱行为的检测精度和监管力度。

总体来说，加密货币反洗钱（anti-money laundering，AML）监管研究正在经历从传统方法到先

进人工智能方法的演变。最初，研究者们借鉴传统金融领域的反洗钱方法，主要采用启发式和基于规则的方法来识别可疑交易。这些方法依赖于预设的规则和专家经验，虽然在一定程度上有效，但随着加密货币交易的复杂性和规模的增加，其局限性逐渐显现。随后，研究者们发现加密货币交易天然具有图结构，这促使越来越多的研究转向图分析方法。通过构建交易图，研究者们能够更直观地分析资金流动和交易模式，从而更有效地识别洗钱活动。随着数据集的不断丰富，机器学习方法开始被广泛应用于加密货币反洗钱研究。在标签数据有限的情况下，无监督学习方法被用来识别潜在的洗钱行为。随着更多标记数据的积累，监督学习方法越来越多地被用于训练模型以提高检测的准确性。近年来，深度学习方法在加密货币反洗钱研究中也得到了广泛的应用。深度神经网络能够自动学习复杂的交易特征，从而更准确地识别洗钱活动。此外，主动学习方法和强化学习方法也被引入，以减少对大量标记数据的依赖，提高模型的适应性和准确性。目前，集成学习方法被认为是加密货币反洗钱研究中较为先进的技术。通过结合多种机器学习模型，集成学习方法能够充分利用不同模型的优势，提高洗钱检测的准确性和鲁棒性。

本文系统地阐述了加密货币反洗钱监管领域的研究进展，总结了现有技术挑战与未来研究方向，致力于为后续研究提供模型、技术和方法等维度的文献基础。需要说明的是，现有研究已针对加密货币的安全与监管等宏观领域进行了综述，例如，文献[3]系统地梳理了加密货币交易中的潜在攻击类型、检测技术与防御措施，为理解加密货币交易的整体安全性提供了新视角；文献[4]关注加密货币生态中的“挖矿”监管、智能合约安全分析、用户身份识别、交易节点追溯等方向。本文则从微观视角出发，聚焦于加密货币反洗钱这一具体领域，为加密货币生态的安全和监管研究提供了新的参考与借鉴。

本文采用系统性研究与综述方法，聚焦加密货币反洗钱领域的技术进展和应用现状，通过制定文献检索策略来收集和分析近年来的相关文献，梳理最新研究成果和技术发展趋势。具体来说，本文以“cryptocurrency/加密货币”和“anti-money laundering/反洗钱”为检索关键词，检索范围为 Web of Science、EI、CNKI 等数据库，限定文献发表时间为 2019 年至 2024 年。经过人工筛选，剔除与研究问题无关或重复的文献，共梳理出 45 篇核心文献作为综述工作的基础文献，以确保文献数据的时效性、多样性和权威性。

1　加密货币反洗钱监管的研究热点

本节首先给出加密货币反洗钱研究任务的问题定义，接下来概述加密货币反洗钱研究的主要热点。

1.1　问题定义

加密货币的交易数据本质上是一种图结构数据，通过将加密货币交易系统中的实体（用户、地址、交易、账户等）表示为节点，并将资金的流向表示为节点之间的有向边，可以构建出图结构的加密货币交易网络。因此，研究者们大多从节点、边和图 3 个维度，即基于节点（用户/地址/交易/账户）、边（资金流向）和图（整个交易网络）来分析加密货币交易，主要研究方法则包括基于规则或启发式的方法、图分析方法、传统机器学习方法、深度学习以及集成学习方法等。

形式上，加密货币反洗钱任务的基本模型是通过分析交易图 $G=(V, E)$ 来识别洗钱行为。其中，节点集 V 表示交易实体（如账户、地址等）的集合，每个节点 $v_i \in V$ 代表一个交易实体，具有特征向量 $\boldsymbol{X}_i$；边的集合 E 表示资金流向，每条边 $e_{ij} \in E$ 代表实体 v_i 和 v_j 之间的资金流动，具有特征向量 $\boldsymbol{A}_{ij}$；交易网络则表示为图 G，其包含所有节点和边的信息。

1.2　研究热点

表 1 详细总结了各个维度和研究方法的特点、典型算法、研究热点与代表性文献。需要说明的是，由于加密货币反洗钱研究领域各种新颖的分析框架层出不穷，为把握关键问题，本文列举的算法均为该研究方法中的主流算法。这些算法通常被作为后续研究中对比实验的基线方法，而代表性文献则为算法的典型实现举例，并非首个应用该算法的文献。

总体来说，现阶段的加密货币反洗钱监管研究主要围绕以下 5 个方向展开，即实体识别与去匿名性、异常行为检测、资金流动追踪与路径分析、交易图构建与特征工程，以及非法活动判定与分类。下面将逐一介绍这些研究方向的核心内容、研究方法及面临的挑战。

表1 研究热点

研究方法		节点（实体/账户/地址/交易）	边（资金流向）	图（交易网络）
基于规则或启发式	特点	依赖于专家知识定义的规则，方法简单但灵活性差，难以应对复杂洗钱行为	通过分析边的特征来识别异常资金流动。这种方法易于实现，但规则固定，难以适应复杂的洗钱模式	通过分析图的整体结构特征来识别异常交易网络。这种方法易于理解和实现，但规则固定，难以适应复杂的洗钱模式
	典型算法	Dempster Shafer理论[5]、规则引擎[6]、阈值检测[6]、风险等级匹配[6]、实体图分析[7]、地址图分析[7]、多输入交易分析[8]、交易网络分析[8]、用户网络分析[8]、截断污染策略[9]、污染分析[9]		
	研究热点	1. 去匿名性[8]； 2. 地址聚类[10]	基于污染分析的洗钱行为追踪[9]	1. 分析交易图、特征提取[7]； 2. 交易数据集构建[11]
图分析/拓扑分析	特点	通过节点的连接关系和图的拓扑结构识别异常，能够捕捉复杂关系，但对图结构要求高	通过分析边的特征，识别异常的资金流动路径。这种方法能够捕捉资金流动的复杂关系，但计算复杂度较高，对图的结构要求严格	通过社区检测、路径分析等技术，识别出异常的交易群体和资金流动路径。这种方法能够捕捉交易网络的复杂结构关系，但计算复杂度较高，对图的结构要求严格
	典型算法	联合图分析（union-find graph algorithm）[7]、交易网络分析[8]、时间循环检测（temporal cycle detection）[12]、LOF[13]、深度自编码器[13]、Hybrid Motifs[14]、DeepWalk[15]、Node2Vec[15]、GCN[16]、K-means聚类[16]		
	研究热点	1. 检测混币服务[13-14]； 2. 交易节点分类[15-16]	1. 识别交易流行为[7]； 2. 追踪资金流动[8]； 3. 分析资金流情况[12]	构建交易图并提取图特征[15]
无监督学习	特点	通过聚类等技术，无须标记数据即可发现未知异常模式，但结果不稳定		将交易网络中的节点和边映射到低维空间，从而识别出异常的图结构，能够发现未知的异常模式，但对初始值和参数设置敏感，结果可能不稳定
	典型算法	LOF[13,17-18]、K-means聚类[13,19-21]、iForest[17]、基于聚类的异常因子[17]、主成分分析[17]、GMM[17]、基于马氏距离的异常检测[18]、期望最大化[18]		
	研究热点	1. 混币服务检测[13]； 2. 异常交易检测[17,22]； 3. 异常用户检测[19,21]； 4. 异常钱包检测[23]		识别异常交易簇[20]
传统的监督学习	特点	利用标记数据训练模型，识别精度较高，但依赖大量标记数据且泛化能力有限		
	典型算法	DT[24-26]、LR[11,17,24,26-27]、SVM[18,24,26-27]、kNN[18,26,28]		
	研究热点	1. 用户合法性分类[24]； 2. 地址分类[25,27]； 3. 交易合法性分类[11,18,26,28]		
深度学习	特点	自动学习节点特征，适用于大规模复杂数据，但需大量数据和计算资源，解释性较差		能够自动学习图的特征表示，捕捉交易网络的复杂结构和模式，尤其适用于大规模复杂数据，但需要大量数据和计算资源，模型解释性较差
	典型算法	EvolveGCN[11]、GCN[11,29-32]、GAT[29,32]、GNN[29-31]、HGNN[29]、深度图信息最大化（deep graph infomax，DGI）[30]、图同构网络（graph isomorphism network，GIN）[30]、SRL[31]、LSTM[33]、RecGNN[33]、DGR[34]、对抗性损失图神经网络（adversarial loss based GNN）[35]		
	研究热点	非法交易识别[11,29,31-38]		可疑子图分类[30]
集成学习	特点	结合多个基学习器的预测结果，减少单一模型的偏差和方差，预测准确性较高		
	典型算法	GB[31,39-41]、AdaBoost[39-41]、XGBoost[42-45]、LightGBM[43,45]、CatBoost[43]、Bagging Classifier[39-40,46]、RF Classifier[39-40,47-48]、ET Classifier[39-40]、Stacking[46,48-49]、混合集成学习[40,50]		
	研究热点	1. 非法交易识别[26,40,43,47,49-50]； 2. 实体类型识别[39,41]； 3. 非法账户识别[44]		

1.2.1 实体识别与去匿名性

加密货币反洗钱研究过程中，破解加密货币匿名性、识别实体身份或类型为后续分析提供了结构化基础，尤为重要。由于加密货币的匿名性，交易背后的实体（如个人或组织）往往难以被直接识别，而明确实体身份是理解交易背景和动机的关

键。去匿名化通过分析地址关联性、交易模式等特征，将匿名地址映射到真实实体。目前，该领域的主要研究热点包括以下4个。（1）去匿名性：通过启发式规则或监督学习将匿名地址映射到真实实体；（2）地址聚类：基于交易模式相似性将地址归并为同一实体；（3）实体类型识别：利用监督学习预测实体类型（交易所、暗网市场等）；（4）交易节点分类：结合图特征和嵌入技术对节点分类。这些研究热点均以明确实体身份为核心，且需结合交易模式和图结构分析。目前，去匿名化面临数据稀疏性和隐私保护的双重挑战，且随着混币服务等匿名增强技术的普及，实体识别的难度进一步增加。

1.2.2　异常行为检测

异常行为检测旨在通过无监督学习或图分析发现不符合正常交易模式的异常交易、用户或子图结构。例如，频繁的小额交易、异常的资金流动或混币服务的使用都可能成为洗钱的迹象。目前，该领域的主要研究热点包括以下3个：（1）异常交易/用户/钱包检测，基于统计特征或图特征识别异常点；（2）混币服务检测，通过图拓扑分析或交易模式识别混币服务；（3）异常交易簇识别，利用聚类算法或子图分析发现异常交易群体。

图分析和无监督学习方法因其不依赖标签数据的特性，在异常检测中应用广泛。这些研究热点均以发现偏离正常模式的行为为目标，需处理高维和非结构化数据，但高误报率、动态交易模式适应的挑战和缺乏高质量标签数据限制了监督学习的应用。

1.2.3　资金流追踪与路径分析

资金流追踪与路径分析是揭示洗钱行为的重要手段。通过构建交易图并分析资金流向，可以识别洗钱路径和污染行为。目前，该领域的主要研究热点包括以下3个。（1）基于污染分析的洗钱行为追踪：利用已知非法资金源头，通过启发式规则追踪后续流向；（2）资金流追踪：构建交易图并分析时间循环、交易模式等动态特征；（3）交易流行为识别：通过有向无环图或子图模式匹配识别洗钱特征。这些方法依赖于交易图的方向性和动态结构特性，能够有效追踪复杂资金流动。然而，随着混币服务和隐私币的普及，资金追踪的难度显著增加，且大规模交易图的计算复杂度也限制了算法的实时性。

1.2.4　交易图构建与特征工程

交易图构建与特征工程是加密货币反洗钱研究的基础环节。交易图通过将实体建模为节点、资金流向建模为边，将原始链上数据转化为结构化图数据。目前，该领域的主要研究热点包括以下3个。（1）交易图构建：定义节点和边，将原始链上数据构建为不同的交易图。例如，比特币的未花费交易输出（unspent transaction output，UTXO）模型和以太坊的账户模型分别适用于不同类型的图构建规则。（2）图特征提取：生成拓扑特征（度中心性、聚类系数）和时序特征（交易间隔）。（3）交易数据集构建：标注数据集并设计特征标签。例如，Elliptic数据集就是典型的交易图数据集，并标注了非法/合法交易标签。交易图构建与特征工程是所有上层分析的基础，直接影响模型性能和可解释性。然而，交易图构建面临数据噪声和异构性挑战，且特征工程的设计高度依赖领域知识，限制了模型的泛化能力。

1.2.5　非法活动判定与分类

非法活动判定与分类任务依赖于带标签数据集，通过各种机器学习模型识别非法交易或账户。目前，该领域的主要研究热点包括以下3个：（1）非法地址/交易识别，常用于比特币数据集，基于标签数据训练二分类模型；（2）可疑子图分类，通过子图表示学习识别犯罪模式；（3）非法账户识别，常用于以太坊数据集，结合账户行为特征和交易图嵌入进行分类。然而，非法活动判定面临标签数据稀缺和模型可解释性不足的挑战，且洗钱行为的动态演变要求模型具备较强的适应性。

2　加密货币反洗钱监管的研究方法

本节将结合上述研究热点，给出现有研究的基本模型和方法，并详细论述各类方法的研究进展和现状。

2.1　基于规则和启发式的方法

加密货币反洗钱监管的早期研究主要是借鉴传统金融领域的经验并结合加密货币的特点，采用基于规则和启发式的方法实现反洗钱监管任务。传统技术手段和数据获取能力有限，难以对大量复杂的加密货币交易数据进行深入分析和处理，因此需要借助规则来简化问题，例如，通过设定交易金额、频率等阈值，初步筛选可疑交易。加密货币的匿名性和去中心化特性使得资金流向难以追踪，传统的

反洗钱方法难以直接应用，而启发式方法则可以根据已知的洗钱模式和特征，结合专家经验，设计特定规则来识别潜在的洗钱行为，从而在一定程度上弥补技术和数据的不足。

在基于法定货币的传统金融系统中，反洗钱系统主要通过明确的规则进行推理[51]，识别出符合预定义规则的可疑交易[52]，并提供可解释、高效和符合伦理要求的反洗钱决策支持。例如，Khanuja 等[5]提出了一种创新的可疑交易监测方法，即依据反洗钱规则标记可疑交易，并利用证据理论综合分析交易特征，量化评估交易可疑性。Panigrahi 等[53]同样基于证据理论对信用卡交易进行多层次的证据融合与动态信念调整，以提高欺诈检测的准确性和鲁棒性。

然而，基于规则的方法存在一定的局限性。例如，规则可能过于僵化，难以应对复杂的洗钱手段和不断变化的洗钱模式。因此，后续研究更多地结合了启发式方法和先进的数据分析技术，以进一步提升反洗钱的效果。启发式方法是一种基于经验和直觉的问题解决策略，利用已有的知识和规则指导问题的求解过程。这种方法特别适用于处理复杂、不确定或不完全确定的问题。在反洗钱研究领域，启发式方法依赖于专家经验和历史数据中的规则，并将这些经验和规则结合起来指导模型的构建和监测策略的制定。例如，参考以往洗钱案件的特征和规律，制定出更有效的监测规则，使模型能够识别出与洗钱行为相关的特定模式和异常行为。同时，启发式方法还允许在研究过程中灵活调整和改进策略，以适应洗钱手段的不断变化[6]。

在加密货币反洗钱研究中，基于启发式的方法更为常见。Ron 等[7]通过启发式方法分析比特币交易图，从大量匿名的交易数据中提取出有意义的统计信息，揭示比特币的存储与流通情况、交易规模与频率等关键特征。Reid 等[8]利用启发式方法分析网络属性和比特币流动路径，揭示了用户之间的潜在关系和流动规律，从而有效进行去匿名性研究。Meiklejohn 等[10]应用启发式和聚类方法对交易数据进行分析，构建出资金流动的网络图谱，通过分析图谱中的交易模式，进一步识别出异常或可疑的模式（如“剥离链”），并追踪资金流动路径，锁定可能涉及洗钱行为的实体和账户，为打击加密货币洗钱行为提供有力的线索和依据。

启发式方法也被广泛应用于构造加密货币反洗钱数据集。Elliptic 数据集的构造过程通过启发式推理[11]，利用比特币交易的伪匿名性和公开信息，提取交易特征并进行合法与非法标记，为后续的机器学习和图卷积网络（graph convolutional network，GCN）等方法提供了有力的数据支持。Wu 等[9]提出了一个名为 XBlockFlow 的框架，用于识别和分析以太坊上的洗钱行为，并构建了一个名为 EthereumHeist 的详细数据集。该数据集的构建过程是一种典型的启发式方法，其核心思想是利用已知的安全事件信息和对洗钱行为特征的经验判断，逐步推断和追踪洗钱资金的流动路径。

另外，启发式方法可用于地址聚类，即通过预定义的规则识别由同一用户控制的多个地址，以实现一定程度的去匿名化。例如，Liu 等[54]提出了改进的多条件识别方法，通过输入地址脚本类型与输出地址脚本类型一致性、输出金额的小数部分非零，以及输出金额小于所有输入地址的支付金额等规则来识别一次性找零地址，并将其与输入地址聚类到同一用户实体。这种方法可以更好地反映真实用户的交易行为，发现潜在的非法交易模式，为反洗钱监管提供了更可靠的分析工具。

2.2 基于图分析的方法

加密货币的交易记录天然具有图拓扑结构的特性，每个交易或账户可以被视为一个节点，而资金流向则构成了节点之间的边。这种结构使研究人员能够直观地分析资金在不同地址之间的流动情况，从而更容易发现异常的交易模式和潜在的洗钱行为。因此，基于加密货币交易的许多研究是在交易图的基础上进行的。例如，Reid 等[8]构建了 2 个有向无环图，分别用于追踪比特币在用户之间的流动和分析交易随时间的变化；Ron 等[7]分析了比特币交易图中的子图，识别出交易流的特征行为。通过图的结构特性（如节点的连接关系、边的方向和权重等），研究者能够追踪资金流动路径、识别交易模式和关联不同地址。例如，通过分析交易图中的长连续交易链、分叉-合并模式等，可以揭示资金的转移路径和潜在的洗钱行为。

与比特币的 UTXO 模式不同，以太坊采用账户模式来管理其状态和交易。Lal 等[12]通过构建以太坊区块链的交易图，利用图分析技术来检测和分析交易中的时间循环（即资金从一个账户出发经过一系列交易后又回到该账户的闭环路径），通过分析时间循环中的资金流动情况，揭示了不同类型的

恶意活动（如赌博、网络钓鱼和洗钱）在时间循环上的独特交易模式和行为特征。

图分析在检测比特币混币服务方面也有比较多的应用。去中心化的加密货币的交易记录在区块链上公开可查，为隐藏交易地址的真实身份，比特币混币服务应运而生。混币服务将多个用户的交易混合在一起，使得追踪资金流向变得困难。因此，检测比特币混币服务对于防止洗钱具有重要意义。Nan 等[13]采用深度自编码器进行图嵌入，提取比特币交易图的特征，再使用 K-means 聚类算法和局部异常概率来检测与混币服务相关的异常节点。Wu 等[14]构建了 2 种网络图，提出了属性时序异构模式（ATH Motifs）的概念，用于捕捉交易网络中的复杂动态过程和资金流动模式，通过从网络图中提取多层面的特征，构建了一个基于正负未标记学习的检测模型。该模型能够有效识别出与混币服务相关的地址，并帮助识别洗钱行为。

图分析与机器学习技术相结合是加密货币反洗钱监管领域的重要方法，即先将加密货币交易表示为图结构，再学习节点特征，进而对节点进行分类，再检测比特币网络上的洗钱活动。例如，Hu 等[15]通过构建比特币交易图，提取了包括节点的入度和出度、统计特征以及网络拓扑特征在内的多种图特征，再利用图嵌入技术（如 Node2Vec）将节点映射到低维向量空间，以捕捉节点之间的复杂关系。在此基础上，结合 Node2Vec 基分类器，实现了对洗钱交易和正常交易的有效区分。后续研究中，研究者开始使用 GCN 来处理比特币交易图的图结构数据。GCN 是一种专门针对图结构数据设计的神经网络，能够有效地捕捉节点之间的关系和图的拓扑结构特征。例如，Alarab 等[16]首先利用 GCN 对比特币交易图进行建模，然后将 GCN 输出的节点嵌入与线性层输出的特征进行拼接，再通过多层感知机（multilayer perceptron，MLP）进行进一步的特征学习和分类预测。这种方法不仅充分利用了图结构数据的优势，还结合了机器学习在特征学习和分类预测中的强大能力，从而提高了对非法交易的识别准确率。

2.3　基于机器学习的方法

2.3.1　传统机器学习

目前传统的机器学习方法在加密货币反洗钱监管方面的应用主要分为监督学习算法和无监督学习算法两类。监督学习算法利用已标注的交易数据进行模型训练，能够有效识别洗钱行为。例如，决策树（decision tree，DT）、逻辑回归（logistic regression，LR）、支持向量机（support vector machine，SVM）和 k 最近邻（k-nearest neighbors，kNN）等经典算法被广泛应用于对加密货币交易数据的分类和预测。这些算法通过提取交易特征（如交易金额、频率、时间等）和用户行为模式构建分类模型，以区分正常交易与可疑交易。然而，由于加密货币交易数据的复杂性和非线性特征，传统监督学习算法在处理大规模数据时可能面临局限性，例如，数据不平衡问题导致模型对少数类别的识别能力不足。

无监督学习算法则不依赖于带标签的数据，而是通过异常检测和聚类分析来识别数据中的异常行为。这种方法特别适用于处理那些没有明确标签的大量交易数据。例如，孤立森林（isolation forest，iForest）和局部异常因子（local outlier factor，LOF）算法被用于检测交易中的异常行为，而 K-means 聚类和高斯混合模型（Gaussian mixture model，GMM）则用于将交易数据分群，以识别洗钱行为的模式。通过这些方法的灵活应用以及与其他监督学习方法的巧妙融合，研究者们能够更有效地分析和识别加密货币交易中的洗钱活动，从而为监管机构和金融机构提供强有力的工具，以应对日益复杂的洗钱挑战。基于传统机器学习算法的代表性文献见表 2。

2.3.2　主动学习

非法交易并不总是表现为明显的异常点，而是可能隐藏在合法交易中，导致一些基于比特币交易数据集的无监督异常检测方法在检测非法交易方面表现欠佳。因此，Lorenz 等[17]提出并验证了主动学习方法在标签稀缺情况下的有效性。实验结果表明，使用主动学习方法，仅需 5% 的标签就可以达到与完全监督学习基线相当的性能。这在实际应用中具有重要意义，因为获取大量标注数据通常是昂贵且耗时的。

Wang 等[55]提出了一种基于强化学习的主动学习模型 GraphALM，旨在提高检测区块链交易中洗钱活动的性能。这种方法通过优化采样策略和分类模型，提高了对区块链交易中洗钱活动的识别准确率，尤其是对少数类别的非法交易样本识别的准确率。利用主动学习方法，减少对大量未标记数据进行标注的需求，降低数据标注的成本和时间。

表2 基于传统机器学习算法的代表性文献

类别	方法	具体算法	代表性文献
无监督学习	异常检测	iForest	文献[17]
		LOF	文献[13,17-18]
		基于聚类的异常因子（cluster-based outlier factor，CBLOF）	文献[17]
		基于马氏距离的异常检测（Mahalanobis distance-based approach，MDB）	文献[18]
		主成分分析（principal component analysis，PCA）	文献[17]
	聚类	K-means 聚类	文献[13,19-21]
		GMM	文献[22]
		期望最大化（expectation-maximization）	文献[23]
监督学习	分类	DT	文献[24-26]
		LR	文献[11,17,24,26-27]
		SVM	文献[18,24,26-27]
		kNN	文献[18,26,28]

2.3.3 强化学习

强化学习作为一种自适应学习技术，通过与环境的交互不断优化策略，为解决加密货币交易中的洗钱检测问题提供了新的思路。通过结合主动学习和自适应策略，强化学习模型能够有效应对数据不平衡问题，并在复杂的交易网络中识别潜在的洗钱活动。

现有文献中，强化学习已被用于开发自适应的防御机制，以应对动态威胁环境。例如，Actor-Critic 算法在攻击防御模拟中表现出较高的成功率，优于传统威胁缓解策略[56]。这种自适应能力使得强化学习在检测和防御复杂威胁方面具有显著优势，因此强化学习在加密货币反洗钱领域应用前景广阔。再如，GraphALM 模型[24]通过强化学习解决了不平衡数据集中的未知特征查询问题，并通过联合损失函数实现了鲁棒且公平的分类，有效提高了区块链交易中洗钱活动的检测性能。

2.3.4 深度学习

传统机器学习算法在处理结构化数据和特征工程方面表现出色，但在面对加密货币交易网络的复杂性和动态性时，逐渐显露出局限性。例如，Alotibi 等[23]通过对比多种机器学习算法在 Elliptic 数据集上的表现，发现随机森林（random forest，RF）在处理不平衡数据集方面表现出色，但其性能仍受限于特征选择和数据预处理的复杂性。这表明，尽管传统方法在某些情况下能够取得较好的效果，但在面对加密货币交易网络的高维性和动态性时，其性能提升空间有限。

随着深度学习技术的兴起，研究者们开始探索其在加密货币反洗钱中的应用，发现深度学习算法能够更有效地挖掘交易网络中的隐藏模式和复杂关系，从而显著提升检测性能。尤其是图神经网络（graph neural network，GNN）及其变体，如 GCN 和图注意力网络（graph attention network，GAT）能够充分利用加密货币交易网络的图结构，挖掘节点间的复杂关系。例如，Mohan 等[36]和 Pan[37]提出了 GCN 和 GAT 在比特币交易网络分析中的应用。这些模型利用图结构数据的特性，通过卷积操作提取特征，用于节点分类等任务，显著提高了非法交易检测的准确性。Ferretti 等[29]研究了异构图神经网络（heterogeneous graph neural network，HGNN）在识别非法和恶意行为中的应用，展示了其在处理包含不同种类节点和边的图结构数据方面的优势。Alarab 等[33]提出的循环图神经网络（robust recurrent graph convolutional network，RecGNN）模型结合了时间序列和图拓扑信息，即结合了修改版的长短期记忆（long short-term memory，LSTM）模型和 GNN，以利用比特币数据的时间行为和图结构进行有效预测，显著提高了非法交易检测的准确性。Lo 等[30]提出的 Inspection-L 框架，通过自监督学习生成节点嵌入，结合 RF 算法，进一步提高了检测性能。子图表示学习（subgraph representation learning，SRL）和动态图表示学习（dynamic graph representation learning，DGR）技术的引入，进一步提升了深度学习模型在加密货币反洗钱中的性能。Bellei 等[31]介绍了 SRL 技术在复杂网络中的局部结构分析中的应用，展示了其在识别特定子图模式方面的优势。Qiao 等[34]提出的 DynAEGCN 模型，利用 DGR

来适应不断变化的交易模式，有效识别异常交易。

随着深度学习技术在加密货币反洗钱研究中的应用不断拓展，GNN的改进和特定问题的针对性解决方案也成了研究重点。Weber等[11]、Humranan等[32]均探讨了GCN结合焦点损失（focal loss，FL）在处理类别不平衡问题上的有效性，并通过实验验证了这种方法在提高非法交易检测性能方面的优势。Adloori等[38]则提出了一种新颖的GAT与残差网络结合的架构（GATResNet），在Elliptic数据集上的实验表明，该模型在检测非法交易方面具有较高的准确性。此外，对抗性损失架构和时间去偏技术也被引入深度学习模型，以提高模型的泛化能力和鲁棒性。例如，Singh等[35]提出了一种基于GNN的对抗性损失架构，在欺诈分类和时间分类之间进行对抗性学习，生成不带时间偏差的特征，从而提高模型在未见过的时间步上的性能。

2.3.5　集成学习

加密货币反洗钱研究中，单个分类器因数据复杂性、不平衡性及特征局限性，容易出现泛化能力不足、偏向多数类及模式挖掘不充分等问题。集成学习可以组合多个不同的学习器来提高模型的准确性和鲁棒性，有效应对数据的复杂性和多样性，且能够处理加密货币交易数据中常见的不平衡问题。例如，通过SMOTE技术生成合成的少数类样本，使用集成模型如RF、梯度提升树（gradient boosting decision tree，GBDT）等能够更好地学习和区分正常与可疑交易。此外，集成学习可以综合考虑交易的多种特征和不同模型的优势，如将基于交易特征的模型与基于图结构的模型相结合，充分利用数值特征和网络结构特征的信息，更准确地识别洗钱行为。同时，它还支持模型的自动化更新和迭代，以适应加密货币交易模式和洗钱手段的变化，保持模型的时效性和有效性。

基于集成学习算法的代表性文献见表3。

（1）Boosting

Boosting集成学习方法通过组合多个弱学习器（通常是DT）来构建强大的预测模型，其核心思想是逐步优化模型，每次迭代都尝试减少前一个模型的错误。Boosting算法有多种实现，其中最著名的是梯度提升（gradient boosting，GB）和自适应提升（adaptive boosting，AdaBoost）。

GB可以用于回归和分类问题，常用的损失函数包括平方误差损失（用于回归）和对数损失（用于分类）。

早在1995年，Senator等[57]就认为基于树的模型在检测传统金融交易中的洗钱行为方面很有潜力。然而，由于缺乏标记数据，因此很难采用这些方法。基于树的模型利用一系列if-then规则来生成预测。随后，Harlev等[39]、Jullum等[42]、Savage等[58]、Weber等[11]、Zhang等[59]的研究均表明基于树的分类器可以有效地检测洗钱活动。例如，Savage等[58]提出，RF与网络分析和社区检测相结合，可以有效地判断洗钱交易。Jullum等[42]使用GB算法检测洗钱交易，因为它效率高、可扩展性强，并且可以通过使用GPU减少训练时间，所以性能优于银行常用的基于规则的手动检查系统。RF和XGBoost（extreme gradient boosting）都是DT集成，只是RF为Bagging方法，而XGBoost为Boosting方法。

XGBoost、LightGBM（light gradient boosting machine）、CatBoost（categorical boosting）是GB

表3　基于集成学习算法的代表性文献

方法	具体算法		代表性文献
Boosting	GB		文献[39-41,43]
	GB的高效实现	XGBoost	文献[42-45]
		LightGBM	文献[43,45]
		CatBoost	文献[43]
	AdaBoost		文献[39,41,44]
Bagging	Bagging Classifier		文献[39-40,46]
	Bagging的特殊实现	RF	文献[11,39,44,47]
		ET Classifier	文献[39-40]
Stacking	Stacking Classifier		文献[46,48-49]
混合集成学习	RF、ET Classifier、Bagging Classifier		文献[40]
	Bagged CNN、Bagged LSTM、Boosted CNN、Boosted LSTM		文献[50]

的高效实现，在传统的GB算法基础上进行了多项优化，使其在处理大规模、复杂且不平衡的数据集时表现出色，因而被广泛应用于加密货币反洗钱研究。Vassallo等[43]采用这3种分类器在交易和账户层面检测加密货币洗钱行为，通过组合多个弱学习器，能够更好地处理数据的复杂模式，提高预测的准确度和鲁棒性。在处理不平衡数据方面，这些算法提供了多种技术，如调整样本权重和使用不同的损失函数，有效减少了假阳性和假阴性，提高了模型的泛化能力。同时，这些算法支持并行和分布式计算，能够快速处理大规模数据集，显著减少了训练时间和资源消耗。另外，XGBoost和CatBoost能够自动处理缺失值，CatBoost还能自动处理类别特征从而提高模型训练效率，因此非常适合捕捉加密货币交易数据中的复杂模式和动态变化。

在比特币交易数据中，合法与非法交易的特征复杂且数据集高度不平衡。每次迭代中，自适应增强（adaptive boosting，AdaBoost）能够对错误分类的样本增加权重，以便于后续的弱学习器更加关注这些难以分类的样本，从而逐步提升对少数类（非法交易）的识别能力，有效减少假阴性，这对于精准打击洗钱活动至关重要。此外，AdaBoost的这种逐步优化方式使其能够动态适应数据的变化，不断调整模型以提高预测准确性。虽然其他算法也有处理复杂数据和不平衡数据的能力，但AdaBoost的这种独特的权重调整和逐步优化机制，使其在比特币反洗钱任务中表现尤为出色，能够更有效地识别和防范非法交易。

（2）Bagging

Bagging，即bootstrap aggregating，通过原始数据集上多次有放回抽样，产生多个不同的训练数据集，再在每个数据集上训练一个基学习器（如DT），最后通过投票或平均方法整合这些基学习器的预测结果。这种方法能够显著减小模型方差，提高模型的稳定性和泛化能力。由于加密货币交易数据的复杂性和高度不平衡性，Bagging通过生成多个不同的训练集，让每个基学习器都能从不同的角度学习数据，从而减少过拟合的风险。此外，Bagging对异常值和噪声数据具有较强的鲁棒性，这是因为每个基学习器只在部分数据上训练，不太可能被少数异常值所影响。

Bagging Classifier是Scikit-learn库中实现Bagging方法的分类器，可以使用任何分类器作为基学习器。与上述提到Bagging的原理相同，在加密货币反洗钱任务中，可以通过该分类器减少模型方差，提高模型的稳定性和泛化能力。例如，Li等[46]提出基于区块链的加密货币反洗钱集成学习框架BELFAL（blockchain-based ensemble learning framework for anti-money laundering），使用Bagging算法（包括简单平均硬投票、简单平均软投票、加权平均硬投票和加权平均软投票）集成多个异构模型的预测结果，显著提高了加密货币反洗钱任务的性能。另外，Bagging Classifier也可以用来解决去匿名性问题，即预测尚未识别实体的类型。Harlev等[39]与比特币分析公司Chainalysis合作，使用已知身份和类型的434个实体（约2亿笔交易）作为训练数据集，构建了区分10个类别的分类器。研究发现，RF、Bagging Classifier和GB在有和没有过采样的情况下都表现优异。

RF是Bagging的一种特殊实现，专门用于DT。它在Bagging的基础上，引入了额外的随机性，即在每次分裂节点时随机选择一部分特征（通常为总特征数的平方根）进行分裂。这种方法进一步增加了模型的多样性，减少了过拟合的风险。在加密货币反洗钱任务中，RF被广泛应用于检测异常交易和洗钱行为[11,40,47]，以及预测实体类型[39]。比特币交易数据中可能存在大量的噪声和异常交易，RF能够有效地过滤这些干扰，准确识别出非法交易。此外，RF可以评估特征的重要性，帮助研究人员理解哪些特征对预测结果影响最大，这对于解释模型的决策过程和进一步优化特征选择非常有帮助。特别地，Weber等[11]在使用LR、RF、MLP和GCN等方法进行分类任务预测非法交易时，发现RF表现最佳，甚至优于包含了图结构信息的GCN。

与RF类似，极端随机树（extremely randomized trees classifier，ET Classifier）也是基于DT的集成学习方法，但其在构建DT时采用了不同的策略。ET Classifier在选择分裂节点时，不仅随机选择特征子集，还在每个特征上随机选择分裂点，而不是像RF那样选择最优分裂点。这种随机性进一步增加了模型的多样性，使得ET Classifier在处理高维数据时更加高效，且能够更好地捕捉数据中的复杂模式。在比特币交易数据中，特征数量众多且存在高度的非线性关系，ET Classifier的这种高随机性使其能够更有效地探索特征空间，提高模型的泛化能力。此外，ET Classifier在训练过程中不需

要进行 Bootstrap 采样，这使得其训练速度通常比 RF 更快，尤其适合处理大规模数据集。因此，ET Classifier 也被广泛应用于检测异常交易和洗钱行为，以及预测实体类型的研究[39-40]。

（3）Stacking

Stacking 结合多个基学习器的结果，利用元学习器进行最终决策，从而提高模型的预测性能。这种方法在处理复杂和多样化的加密货币交易数据时，具有模型多样性与互补性，提高预测性能、灵活性、自适应性以及可解释性等显著优势。目前，Stacking 在加密货币反洗钱领域的应用还相对较少，但已有研究表明其在提高检测性能方面具有一定的潜力。例如，Md 等[49]和 Kumar 等[48]通过结合不同的基学习器（如 kNN、LR、GBDT、RF 等）和元学习器（如 LR），在加密货币交易数据上取得了显著的性能提升，展示了 Stacking 在处理复杂和动态变化的交易数据，特别是在提高模型的准确性和鲁棒性方面的优势。Md 等[49]发现以多项式朴素贝叶斯和 RF 为基础学习器、以 LR 为元学习器的 Stacking 分类器取得了最高的准确率（97.18%）。另外，Li 等[46]提出的 BELFAL 框架中也使用了 Stacking 方法，使用多种不同的机器学习模型作为基学习器（包括 LR、SVM、kNN、DT 和 MLP），采用 RF 作为元学习器将这些基学习器的预测结果进行进一步的训练和整合，实验结果表明该方法在 F1 值和 Macro-F1 值上优于单一模型，为加密货币反洗钱任务提供了一种更准确、鲁棒性更高的检测方法。

（4）混合集成学习

加密货币交易数据的复杂性、高维度以及不平衡性，使得单一类型的集成学习算法难以全面捕捉数据中的所有信息。例如，RF 虽然在处理高维数据和减少过拟合方面表现出色，但对于某些复杂的非线性关系可能效果不佳；XGBoost 虽然在处理大规模数据集时效率高、准确性好，但对数据中的噪声和异常值较为敏感。因此，将这些算法融合起来，可以相互补充，发挥各自的优势。

混合集成学习方法融合了多种异质算法的优势，能够更全面、准确地检测洗钱活动。例如，Alarab 等[40]结合 RF、ET Classifier 和 Bagging Classifier 在不同方面的优势，实现了比单一模型更精准的预测。这种混合集成学习方法通过计算多种不同方法的预测概率均值来生成最终预测，不仅增强了模型对数据的泛化能力，还有效降低了因单一模型偏差导致的预测错误，提高了预测的稳定性和可靠性。此外，该方法在处理比特币交易数据时，能够更好地应对数据的复杂性和不平衡性，减少假阳性和假阴性的发生，这对于反洗钱任务来说至关重要。实验结果也证明了该集成学习方法在准确率和 F1 值上的卓越表现，准确率达到 98.13%，F1 值达到 83.36%，显著优于其他单一监督学习方法，从而证实了其在比特币反洗钱问题中的有效性和实用性。

另外，深度学习模型也可以与集成学习深度结合。例如，Umer 等[50]提出的混合集成学习方法通过结合 2 种深度学习模型和 2 种集成技术，实现了对加密货币欺诈交易的高效预测。具体来说，该研究使用 CNN 提取交易数据中的结构化特征，并使用 LSTM 捕捉时间序列数据中的长期依赖关系，使模型能够更全面地处理复杂的交易数据。同时，使用 Bagging 并行训练多个模型并取平均预测结果来减少方差，提高了模型的稳定性和鲁棒性；使用 Boosting 逐步训练多个模型并加权平均预测结果来减少偏差，进一步提升了模型的准确性。这种混合集成学习方法不仅充分利用了不同模型的优势，还通过集成技术进一步优化了模型的性能，显著提高了欺诈交易预测的准确性和鲁棒性。

3　数据集与模型性能

本节将探讨加密货币反洗钱监管研究中的主流数据集及其在反洗钱研究中的应用，以及多种常见分类算法在 Elliptic 数据集上的性能表现。

目前用于加密货币反洗钱分析的主流数据集可以分为三大类：（1）以 Elliptic 数据集为代表的比特币洗钱交易标签数据集；（2）以 EthereumHeist 数据集为代表的以太坊洗钱交易标签数据集；（3）以 Chainalysis 数据集为代表的比特币实体标签数据集。前两者与加密货币洗钱问题直接相关，而对 Chainalysis 数据集的研究可以通过区分实体类型从而间接帮助反洗钱研究。就数据集规模而言，Elliptic 数据集无论是特征维度还是节点数量都远大于 EthereumHeist 数据集，因此目前绝大多数反洗钱研究是基于 Elliptic 数据集开展的。

目前，最大规模的带标签开源交易数据集是比特币 Elliptic 数据集，由麻省理工学院与 Elliptic 合作创建，包含 20 万笔比特币交易，涉及洗钱交易的总金额达 60 亿美元。这些交易被标记为“合法”或“非法”，每个交易节点包含 166 个特征，涵盖

了交易的时间、金额、历史行为等信息，Elliptic数据集典型节点特征举例见表4。Elliptic数据集庞大的数据规模和丰富的特征为机器学习模型的训练提供了基础，是现有文献中应用最广泛的数据集。基于该数据集，结合多种机器学习方法（如LR、RF、MLP和GCN），研究者能够成功识别非法比特币交易[11]。2024年，该团队进一步创建了Elliptic2数据集，其包含了121 810个标记的比特币集群子图，位于一个包含4 900万个节点集群和1.96亿次边缘交易的背景图中。该数据集专注于通过SRL揭示加密货币中的洗钱活动模式，并准确分类新的犯罪活动[30]。

表4 Elliptic数据集典型节点特征举例

类别	特征	解释
本地特征	时间步	交易发生的时段
	输入/输出数量	交易中的输入或输出的数量
	交易费	进行交易所需支付的费用
	输出量	交易中的输出金额或数量
聚合特征	最大/最小值	例如，邻居交易中的最大和最小交易费
	标准差	描述邻居交易某一特定数据（如交易费）的离散程度
	相关系数	描述邻居交易的某一数据与其他数据之间的关联程度

现有文献中，许多学者致力于在Elliptic数据集的基础上进一步扩展。例如，Elliptic++数据集包含超过82.2万个比特币钱包地址（节点），以及127万条时间交互信息，为更全面深入地分析比特币网络中的交易行为提供了更丰富的数据基础[60]。Song等[61]基于Elliptic数据集重构，通过增加历史交易数据、引入钱包地址节点、提取元路径信息等改进，创建了异构比特币交易行为数据集HBTBD，并将构建HBTBD的代码发布于GitHub。由于Elliptic数据集的特征都是匿名的，缺乏精确描述，且没有缺失值，这与真实世界的需求不符，因此，Wang等[55]在Elliptic数据集的基础上构造了RD（realistic and demand）数据集，增加了更有意义的特征，同时引入了缺失值，以更符合真实世界场景。

除了直接对交易进行分类外，实体识别与分类（即去匿名性）也是关键的分析手段。去匿名性能够将交易背后的实体（如个人或组织）识别出来，从而揭示交易的真实参与者。这一过程对于理解交易的背景和动机至关重要。因为即使交易本身看似合法，其背后的实体也可能涉及非法活动。例如，一个交易可能通过多个看似无关的地址进行，但通过去匿名性可以发现这些地址实际上属于同一非法组织，从而揭示潜在洗钱行为。Chainalysis提供的比特币实体集群数据集就是处理这类任务的数据集[41]。这些数据不仅涵盖交易的详细信息，还通过标签数据和机器学习模型，将链上地址与暗网、诈骗分子、勒索软件、去中心化金融（decentralized finance，DeFi）平台、矿池等合法或非法实体对应起来。这种丰富的数据集为研究人员和执法机构提供了强大的工具，以深度调查任何一笔链上交易及其关联交易，并实时监控交易，以满足反洗钱等合规要求。

上述数据集仅局限于比特币交易。对以太坊交易来说，Wu等[9]提出了首个详细的以太坊洗钱数据集EthereumHeist。该数据集包含交易的时间和金额、服务提供商地址的标签、洗钱地址所在的层级等丰富的信息。现阶段，针对其他加密货币反洗钱任务的公开可用的数据集仍然欠缺。表5展示了加密货币反洗钱数据集的详细情况。

表6给出了各种分类算法在Elliptic数据集上的性能表现，这些分类算法常常被当作基线方法。

4 挑战与展望

目前，加密货币反洗钱监管的相关学术研究和技术应用均取得了较好的进展。现有文献提出的研究方法为加密货币交易行为监管、交易网络分析和安全风险评估提供了强大的技术支撑；高质量数据集则为模型训练和验证奠定了坚实的数据基础，以支持加密货币反洗钱的量化评估和深入分析。

然而，随着加密货币行业的迅猛发展，新币种、新技术和新监管需求的不断涌现也给加密货币反洗钱研究带来诸多亟待解决的新问题和新挑战。首先，加密货币市场的新币种、新模式和新场景快速发展演化。例如，隐私币的兴起和智能合约的广泛应用，进一步加剧了匿名性与隐私保护的挑战，使得传统监管手段难以有效应对；应用场景的日益复杂，尤其是利用DeFi平台和跨链交易洗钱的案例层出不穷，导致资金流向更加隐蔽，数据的复杂性与不平衡性问题愈发突出。其次，新兴技术的快速迭代带来了交易规模的指数级增长和交易模式的快速演变，使模型的泛化能力和实时性面临严峻考验。最后，监管需求的不断强化对跨境监管协调、数据共享机制以及监管资源与技术能力提出了更高

表5　加密货币反洗钱数据集的详细情况

数据集	对象	特征	类别	规模
Chainanalysis	实体	98	交易所、托管钱包、商户服务、矿池、混币服务、赌博、诈骗、暗网市场、勒索软件、被盗比特币、个人钱包、其他	约3.95亿笔交易（已标记数据957个实体，未分类数据153 293个集群）
Elliptic	交易	166	非法、合法、未知	203 769个交易节点
Elliptic2	交易	56	非法、合法、未知	121 810个标记的比特币集群子图
Elliptic++	交易 钱包地址	交易：183 钱包地址：56	非法、合法、未知	交易数据集：203 769个交易节点，234 355条有向边 钱包地址数据集：822 942个钱包地址
HBTBD	交易 钱包地址	交易：166 钱包地址：8	交易：非法、合法、未知 钱包地址：只有输入没有输出、只有输出没有输入、既有输入又有输出	46 045个比特币交易实体 319 311个相关的比特币钱包地址
RD	交易	186	非法、合法、未知	203 769个交易节点
EthereumHeist	账户	11	可疑洗钱账户、提供某些服务但未被Etherscan标记的账户	2018—2022年以太坊代表性的73个盗窃案，洗钱账户16万个以上

表6　Elliptic数据集上各种算法结果

模型	精确率	召回率	F1值
kNN[40]	61.60%	63.99%	62.77%
GB[40]	99.84%	59.37%	74.46%
Adaptive Boosting[40]	96.28%	71.83%	82.28%
Bagging Classifier[40]	96.41%	72.11%	82.51%
RF[40]	97.38%	72.20%	82.92%
ET Classifier[40]	98.70%	70.36%	82.15%
集成学习[40]	99.11%	71.93%	83.36%
XGBoost[43]	92.10%	73.20%	81.50%
LightGBM[43]	93.20%	73.20%	82.00%
CatBoost[43]	93.60%	72.80%	81.90%
DT[45]	50.2%	66.6%	57.2%
SVM[45]	65.7%	63.8%	64.7%
LR[11]	53.7%	52.8%	53.3%
MLP[11]	78.0%	61.7%	68.9%
GCN[11]	81.2%	51.2%	62.8%
跳跃图卷积网络（Skip-GCN）[11]	81.2%	62.3%	70.5%
动态图卷积网络（EvolveGCN）[11]	85.0%	62.4%	72.0%

要求，而现有体系在这方面的不足逐渐显现。未来的研究亟须在这些方面取得突破，开发更精准的检测算法，探索多模态、多源数据融合的新路径，在隐私保护与监管之间寻求动态平衡，并利用大模型技术和基于区块链的监管模式等创新手段，推动加密货币反洗钱监管迈向新的高度，实现行业的健康、有序发展。

本节将详细讨论目前加密货币反洗钱监管领域面临的上述挑战，并展望未来的发展方向。图1展示了由现有技术存在的局限性引发的新挑战与需求，以及从需求出发而引申的未来方向展望之间相互支撑的逻辑关系。

4.1　挑战

加密货币反洗钱研究面临的挑战主要涉及技术层面的匿名性、数据复杂性以及模型泛化能力等问题，也涵盖了监管层面的跨境协调、资源不足以及数据共享机制不完善等现实困境。

4.1.1　技术层面的挑战

（1）匿名性和隐私保护

加密货币的匿名性为用户提供了隐私保护，也为洗钱等非法活动提供了便利。特别地，混币技术将资金与多个其他交易混合，极大地增加了资金流向的追踪难度。例如，去中心化的加密货币混币服务Tornado Cash被广泛用于洗钱活动，尤其是在一些重大加密货币盗窃事件中。美国财政部下属的外国资产控制办公室指出，2022年3月，黑客组织Lazarus Group利用Tornado Cash清洗了超过4.55亿美元的被盗加密货币。隐私币（如门罗币、零币等）的兴起进一步加剧了这一问题，其加密技术使得交易几乎无法被追踪[62]。由于隐私币的市场份额小、高度匿名性和技术的复杂性，现阶段针对隐私币的反洗钱研究少之又少。因此，如何在保护用户隐私的同时有效识别和追踪可疑交易，如何对隐私币反洗钱等尚未开拓的领域进行系统性研究，都是当前亟待解决的技术难题。

（2）数据复杂性与不平衡性

在涉及大量交易的情况下，加密货币市场的高速交易特性使追踪资金流向变得极为困难。同时，加密货币资金能够跨境自由流动，不受传统金融体系的限制，进一步加剧了数据处理和分析的难度[63]。近年来，兴起的DeFi也为加密货币交易提供了新的平台。DeFi平台通过智能合约自动执行

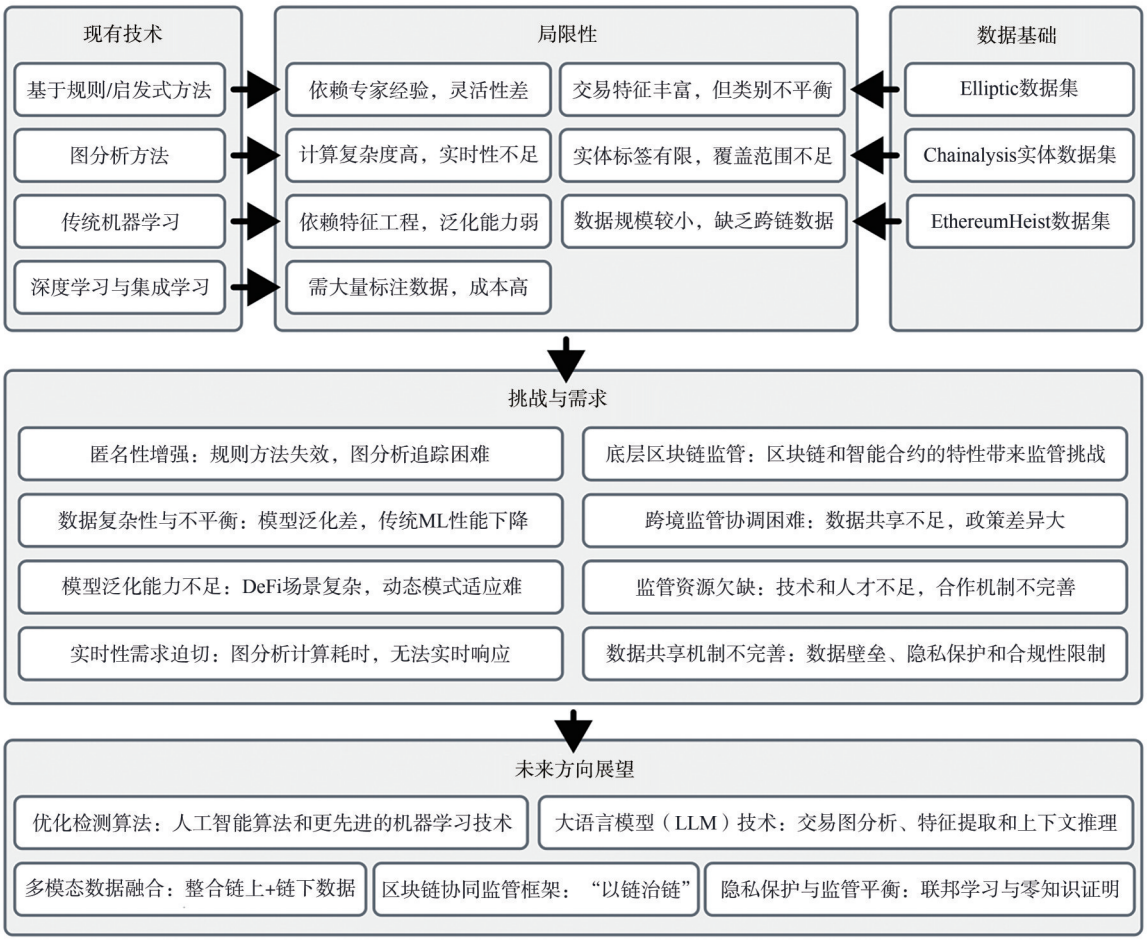

图 1 挑战与展望

金融操作，不需要中心化的中介机构。这种去中心化的交易模式拥有更高的隐私性和安全性，但也使得交易数据更加分散和难以追踪。同时，在绝大多数交易场景和数据集中，合法交易与洗钱交易的样本比例悬殊，导致数据类别不平衡。例如，Elliptic 的研究表明，其模型训练数据中只有不到万分之一的交易被标记为洗钱交易[30]。这种不平衡使得传统的机器学习方法在检测洗钱行为时容易出现误判。

（3）模型的泛化能力和实时性

加密货币市场的动态性强，交易模式和洗钱手段不断演变[64]，现有的机器学习模型在面对新交易模式和数据时，往往需要重新训练和调整。例如，DeFi 的兴起为洗钱提供了新的途径，通过去中心化交易所 DEX 进行跨链洗钱成为可能。如何提高模型的泛化能力，使其能够适应不断变化的市场环境，是当前研究的重要方向。同时，对洗钱行为的检测需要实时性，以便及时采取措施，但目前的模型在实时性方面仍存在不足。

（4）底层区块链监管的挑战

区块链的去中心化和不易篡改性、智能合约的自动化性质等，都为洗钱行为监管带来巨大挑战[4]。首先，区块链将交易记录存储在分布式网络的多个节点上，这些节点共同验证和记录交易。这种去中心化特性使得监管机构难以采用传统金融系统的中心化监管模式。其次，智能合约的自动化性质意味着一旦满足预设条件，合约将去中心化自动执行，不受人工干预。这种自动化特性使得监管者需要更新传统的法律体系，以应对新出现的交易方式和商业模型。智能合约的不可逆性则使得监管机构在发现非法交易时，也无法像在传统金融系统中那样冻结或撤销交易。最后，区块链的不易篡改特性使得监管机构在发现非法交易时难以进行有效的干预和纠正，导致“监易管难”，迫使监管机构更加依赖事前预防和实时监管而非事后干预，这也增加了监管的难度和复杂性。

4.1.2　监管层面的挑战

（1）跨境监管协调困难

加密货币的跨境特性迫切需要各国之间加强协调与合作。然而，不同国家和地区的监管政策差异较大，导致跨境监管协调困难。例如，俄罗斯在2024年将加密货币“挖矿”合法化，而西方国家则持谨慎态度，这种政策差异使得跨境洗钱行为的监管难以有效实施。同时，国际合作机制现阶段尚不完善，各国在信息共享、技术交流和联合执法等方面均存在不足。例如，金融行动特别工作组虽然制定了全球反洗钱标准，但在具体实施过程中，各国的执行力度和效果都存在差异[65]。

（2）监管资源与技术能力不足

加密货币市场的复杂性和强技术依赖性要求监管机构具备强大的技术能力和充足的监管资源。然而，目前许多国家的监管机构在技术手段和专业人才方面存在不足。例如，Qiao 等[34]的研究表明，尽管人工智能模型在检测加密货币洗钱方面表现出色，但误报率仍然较高，需要专业人员进行二次筛选和确认。同时，监管机构与技术研究机构之间的合作机制尚不完善，导致技术成果难以快速转化为实际监管能力[66]。

（3）数据共享机制不完善

加密货币反洗钱行为监管需要大量的交易数据来训练和验证模型，但目前数据共享机制尚不完善。不同机构和企业之间存在数据壁垒，导致数据难以获取和共享。例如，传统的金融机构与加密货币反洗钱机构之间缺乏有效的数据共享平台，使得反洗钱监测分析系统无法实现跨机构的客户尽职调查。此外，数据的隐私保护和合规性要求也限制了数据的共享范围[64]。

4.2　展望

目前面临的种种挑战为未来的研究提供了广阔的探索空间和方向。加密货币反洗钱研究正在技术、监管、数据共享以及各方合作等多个维度展现出新的希望和潜力。接下来将借鉴传统金融领域的反洗钱实践、当前已开展的探索性研究，以及未来可能涉及的研究方法，对未来加密货币反洗钱研究展开进一步探讨。

（1）开发更精准的检测算法

未来的研究需要进一步优化现有的机器学习和人工智能算法[67-68]，开发更精准的洗钱检测模型[69]。通过引入更先进的深度学习技术，如GNN和强化学习，更好地处理加密货币交易的图结构数据，提高模型的检测精度和泛化能力。同时，可以结合主动学习和迁移学习方法，减少对大量标记数据的依赖，提高模型的适应性和实时性[70]。

（2）多模态、多源数据融合

除交易数据，还可以尝试引入其他类型的数据，如用户行为数据、社交媒体数据等，通过分析用户在社交媒体上的言论和行为，可以发现与加密货币交易相关的异常情绪或可疑活动[71]。一些工具（如LunarCrush和Glassnode）已经提供了对加密货币市场情绪的分析，可以作为一项强有力的数据补充。将链上数据（如交易记录、钱包地址等）与社交媒体数据相结合，可以更全面地追踪资金流向和用户行为。例如，通过分析链上数据中的交易模式和社交媒体上的用户行为，可以发现某些用户可能在进行洗钱活动。

（3）隐私保护与监管平衡

监管机构需要在数据共享和隐私保护之间找到平衡，即在有效监测洗钱行为的同时，保护用户数据隐私[72-73]。例如，通过数据脱敏、加密、匿名化等技术手段，在保护个人隐私的同时，实现数据的有效共享和利用。此外，建立完善的数据管理制度和流程，明确数据的收集、存储、使用、共享等环节的责任和义务，也是实现隐私保护与监管平衡的重要措施[74]。在实际应用中，联邦学习技术提供了一种可能的解决方案[75]。其允许在不共享原始数据的情况下进行模型训练，从而保护数据隐私，同时实现多模态数据的有效融合。这种技术可以在保护用户隐私的同时，提高反洗钱监测的准确性和效率。此外，零知识证明和同态加密等隐私保护技术正在逐步应用于加密货币反洗钱领域[76]。例如，Zcash通过零知识证明技术在保护用户隐私的同时，允许监管机构对可疑交易进行有限度的追踪。

（4）大语言模型技术的应用

大语言模型（large language model，LLM）凭借其强大的推理能力和对复杂数据的处理能力，为加密货币交易分析提供了新的研究路径和应用可能性。传统金融机构在反洗钱领域已经积累了丰富的经验和技术，可以将LLM技术借鉴到加密货币反洗钱领域[77]。例如，兴业银行的AML-GPT大模型可以通过分析加密货币交易数据、用户行为数据和社交媒体数据，智能生成可疑交易分析报告。另

外，LLM技术在交易图分析、特征提取和上下文解释等方面均取得了初步的研究进展。Lei等[78]的研究表明，LLM在加密货币交易图分析中具有显著潜力，尤其是在反洗钱领域，结合LLM4TG格式和CETraS算法，LLM能够有效处理复杂的交易图，并提供有价值的分析结果和解释。Nicholls等[79]利用LLM生成交易的上下文叙述，并通过嵌入向量计算相似性，识别出与目标交易相似的非法交易。这种方法不仅提高了非法交易的识别效率，还为分析人员提供了详细的解释和报告，增强了模型的可解释性和可信度。未来的研究可以通过结合大模型与其他先进技术（如图神经网络和可解释人工智能方法），进一步提升加密货币交易分析的效率和准确性。

（5）基于区块链的监管模式

区块链技术的分布式、去中心化特征虽然带来了监管的挑战，但也提供了新的监管思路。研究者已提出基于区块链对加密货币进行监管，形成多链协同治理的“以链治链”监管框架[80]，利用区块链的不易篡改和透明性特点，提高监管的效率和准确性，实现对加密货币交易的实时监测和管理[81-82]。

5 结束语

本文对加密货币反洗钱监管研究进展和现状进行了综述，阐述了该领域在各研究方向和研究方法方面的研究热点和代表性工作，总结了该领域的数据集及其相关研究，梳理和量化评估了主流方法在这些数据集上的性能表现，最后探讨了该领域面临的技术挑战与发展趋势。随着加密货币市场的日益繁荣，基于加密货币的洗钱行为监管的需求日益迫切，亟待研究者采用先进的人工智能方法提供更精准、高效的监管方案，期待本文可为未来的研究工作提供有益的参考与借鉴。

参考文献：

[1] LIU F, FAN H Y, QI J Y. Blockchain technology, cryptocurrency: entropy-based perspective[J]. Entropy, 2022, 24(4): 557.

[2] LIU F, FENG Z F, QI J Y. A blockchain-based digital asset platform with multi-party certification[J]. Applied Sciences, 2022, 12(11): 5342.

[3] 刘峰, 江佳齐, 黄灏. 面向加密货币交易介质及过程的安全综述[J]. 信息网络安全, 2024, 24(3): 330-351.
LIU F, JIANG J Q, HUANG H. Security overview of cryptocurrency trading media and processes[J]. Netinfo Security, 2024, 24(3): 330-351.

[4] 王佳鑫, 颜嘉麒, 毛谦昂. 加密数字货币监管技术研究综述[J]. 计算机应用, 2023, 43(10): 2983-2995.
WANG J X, YAN J Q, MAO Q A. Overview of cryptocurrency regulatory technologies research[J]. Journal of Computer Applications, 2023, 43(10): 2983-2995.

[5] KHANUJA H K, ADANE D S. Forensic analysis for monitoring database transactions[C]//Security in Computing and Communications. Heidelberg: Springer, 2014: 201-210.

[6] SALEHI A, GHAZANFARI M, FATHIAN M. Data mining techniques for anti money laundering[J]. International Journal of Applied Engineering Research, 2017, 12(20): 10084-10094.

[7] RON D, SHAMIR A. Quantitative analysis of the full Bitcoin transaction graph[C]//Financial Cryptography and Data Security. Heidelberg: Springer, 2013: 6-24.

[8] REID F, HARRIGAN M. An analysis of anonymity in the Bitcoin system [C]//Security and Privacy in Social Networks. New York: Springer, 2012: 197-223.

[9] WU J J, LIN D, FU Q S, et al. Toward understanding asset flows in crypto money laundering through the lenses of Ethereum heists[J]. IEEE Transactions on Information Forensics and Security, 2023, 19: 1994-2009.

[10] MEIKLEJOHN S, POMAROLE M, JORDAN G, et al. A fistful of Bitcoins: characterizing payments among men with no names[C]//Proceedings of the 2013 Conference on Internet Measurement Conference. New York: ACM, 2013: 127-140.

[11] WEBER M, DOMENICONI G, CHEN J, et al. Anti-money laundering in Bitcoin: experimenting with graph convolutional networks for financial forensics[J]. arXiv preprint, 2019: arXiv:1908.02591.

[12] LAL B, AGARWAL R, SHUKLA S K. Understanding money trails of suspicious activities in a cryptocurrency-based blockchain[J]. arXiv preprint, 2021: arXiv:2108.11818.

[13] NAN L H, TAO D C. Bitcoin mixing detection using deep autoencoder [C]//Proceedings of the 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC). Piscataway: IEEE Press, 2018: 280-287.

[14] WU J J, LIU J L, CHEN W L, et al. Detecting mixing services *via* mining Bitcoin transaction network with hybrid motifs[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(4): 2237-2249.

[15] HU Y N, SENEVIRATNE S, THILAKARATHNA K, et al. Characterizing and detecting money laundering activities on the Bitcoin network[J]. arXiv preprint, 2019: arXiv:1912.12060.

[16] ALARAB I, PRAKOONWIT S, NACER M I. Competence of graph convolutional networks for anti-money laundering in Bitcoin blockchain[C]//Proceedings of the 2020 5th International Conference on Machine Learning Technologies. New York: ACM, 2020: 23-27.

[17] LORENZ J, SILVA M I, APARÍCIO D, et al. Machine learning methods to detect money laundering in the bitcoin blockchain in the presence of label scarcity[C]//Proceedings of the First ACM International Conference on AI in Finance. New York: ACM, 2021: 1-8.

[18] CHEN B J, WEI F S, GU C X. Bitcoin theft detection based on supervised machine learning algorithms[J]. Security and Communication Networks, 2021(1): 6643763.

[19] CHANG T H, SVETINOVIC D. Improving Bitcoin ownership identification using transaction patterns analysis[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50(1): 9-20.

[20] ZHANG R, ZHANG G F, LIU L, et al. Anomaly detection in Bitcoin information networks with multi-constrained meta path[J]. Journal of Systems Architecture, 2020, 110: 101829.

[21] SHAYEGAN M J, SABOR H R, UDDIN M, et al. A collective anomaly detection technique to detect crypto wallet frauds on Bitcoin network[J].

Symmetry, 2022, 14(2): 328.

[22] YANG L X, DONG X W, XING S Y, et al. An abnormal transaction detection mechanim on Bitcoin[C]//Proceedings of the 2019 International Conference on Networking and Network Applications (NaNA). Piscataway: IEEE Press, 2019: 452-457.

[23] BAEK H, OH J, KIM C Y, et al. A model for detecting cryptocurrency transactions with discernible purpose[C]//Proceedings of the 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN). Piscataway: IEEE Press, 2019: 713-717.

[24] NERURKAR P, BHIRUD S, PATEL D, et al. Supervised learning model for identifying illegal activities in Bitcoin[J]. Applied Intelligence, 2021, 51(6): 3824-3843.

[25] LIANG J Q, LI L J, CHEN W Y, et al. Targeted addresses identification for Bitcoin with network representation learning[C]//Proceedings of the 2019 IEEE International Conference on Intelligence and Security Informatics (ISI). Piscataway: IEEE Press, 2019: 158-160.

[26] CHANG Y L, HSIANG K L, FU C T. A systematic review of detecting illicit bitcoin transactions[J]. Procedia Computer Science, 2022, 207: 3217-3225.

[27] LIN Y J, WU P W, HSU C H, et al. An evaluation of Bitcoin address classification based on transaction history summarization[C]//Proceedings of the 2019 IEEE International Conference on Blockchain and Cryptocurrency (ICBC). Piscataway: IEEE Press, 2019: 302-310.

[28] ALOTIBI J, ALMUTANNI B, ALSUBAIT T, et al. Money laundering detection using machine learning and deep learning[J]. International Journal of Advanced Computer Science and Applications, 2022, 13(10): 732-738.

[29] FERRETTI S, D'ANGELO G, GHINI V. On the use of heterogeneous graph neural networks for detecting malicious activities: a case study with cryptocurrencies[C]//Proceedings of the 4th International Workshop on Open Challenges in Online Social Networks. New York: ACM, 2024: 33-40.

[30] LO W W, KULATILLEKE G K, SARHAN M, et al. Inspection-L: self-supervised GNN node embeddings for money laundering detection in Bitcoin[J]. Applied Intelligence, 2023, 53(16): 19406-19417.

[31] BELLEI C, XU M H, PHILLIPS R, et al. The shape of money laundering: subgraph representation learning on the blockchain with the Elliptic2 dataset[J]. arXiv preprint, 2024: arXiv: 2404.19109.

[32] HUMRANAN P, SUPRATID S. A study on GCN using focal loss on class-imbalanced Bitcoin transaction for anti-money laundering detection[C]//Proceedings of the 2023 International Electrical Engineering Congress (iEECON). Piscataway: IEEE Press, 2023: 101-104.

[33] ALARAB I, PRAKOONWIT S. Robust recurrent graph convolutional network approach based sequential prediction of illicit transactions in cryptocurrencies[J]. Multimedia Tools and Applications, 2024, 83(20): 58449-58464.

[34] QIAO C B, TONG Y Z, XIONG A, et al. Block-chain abnormal transaction detection method based on dynamic graph representation[M]// Game Theory for Networks. Cham: Springer, 2022: 3-15.

[35] SINGH A, GUPTA A, WADHWA H, et al. Temporal debiasing using adversarial loss based GNN architecture for crypto fraud detection[C]// Proceedings of the 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA). Piscataway: IEEE Press, 2021: 391-396.

[36] MOHAN A, KARTHIKA P V, SANKAR P, et al. Improving anti-money laundering in Bitcoin using evolving graph convolutions and deep neural decision forest[J]. Data Technologies and Applications, 2023, 57(3): 313-329.

[37] PAN Y C. Enhancing predictive models for illicit activities in the Bitcoin transaction network using advanced graph analytical techniques[J]. Applied and Computational Engineering, 2024, 48(1): 78-86.

[38] ADLOORI H, DASANAPU V, MERGU A C. Graph network models to detect illicit transactions in block chain[J]. arXiv preprint, 2024: arXiv: arXiv: 2410.07150.

[39] HARLEV M A, SUN YIN H H, LANGENHELDT K C, et al. Breaking bad: de-anonymising entity types on the Bitcoin blockchain using supervised machine learning[C]//Proceedings of the 51st Hawaii International Conference on System Sciences 2018. Honolulu: Hawaii International Conference on System Sciences (HICSS), 2018: 3497-3506.

[40] ALARAB I, PRAKOONWIT S, NACER M I. Comparative analysis using supervised learning methods for anti-money laundering in Bitcoin[C]//Proceedings of the 2020 5th International Conference on Machine Learning Technologies. New York: ACM, 2020: 11-17.

[41] SUN YIN H H, LANGENHELDT K, HARLEV M, et al. Regulating cryptocurrencies: a supervised machine learning approach to de-anonymizing the Bitcoin blockchain[J]. Journal of Management Information Systems, 2019, 36(1): 37-73.

[42] JULLUM M, LØLAND A, HUSEBY R B, et al. Detecting money laundering transactions with machine learning[J]. Journal of Money Laundering Control, 2020, 23(1): 173-186.

[43] VASSALLO D, VELLA V, ELLUL J. Application of gradient boosting algorithms for anti-money laundering in cryptocurrencies[J]. SN Computer Science, 2021, 2(3): 143.

[44] FARRUGIA S, ELLUL J, AZZOPARDI G. Detection of illicit accounts over the ethereum blockchain[J]. Expert Systems with Applications, 2020, 150: 113318.

[45] HEGADI R, TRIPATHI B, NAMRATHA S, et al. Anti-money laundering analytics on the Bitcoin transactions[C]//Information Security, Privacy and Digital Forensics. Singapore: Springer, 2023: 405-418.

[46] LI Z Y, YAO R Z, YANG D, et al. BELFAL: a blockchain-based ensemble learning framework for anti-money laundering in cryptocurrency markets[C]//Proceedings of the 2024 IEEE 30th International Conference on Parallel and Distributed Systems (ICPADS). Piscataway: IEEE Press, 2024: 520-527.

[47] LEE C, MAHARJAN S, KO K, et al. Toward detecting illegal transactions on Bitcoin using machine-learning methods[C]//Blockchain and Trustworthy Systems. Singapore: Springer, 2019: 520-533.

[48] KUMAR N, SINGH A, HANDA A, et al. Detecting malicious accounts on the ethereum blockchain with supervised learning[C]//Cyber Security Cryptography and Machine Learning. Cham: Springer, 2020: 94-109.

[49] MD A Q, SATYA SREE NARAYANAN S M, SABIREEN H, et al. A novel approach to detect fraud in Ethereum transactions using stacking[J]. Expert Systems, 2023, 40(7): e13255.

[50] UMER Q, LI J W, ASHRAF M R, et al. Ensemble deep learning-based prediction of fraudulent cryptocurrency transactions[J]. IEEE Access, 2023, 11: 95213-95224.

[51] BELLOMARINI L, LAURENZA E, SALLINGER E, et al. Rule-based anti-money laundering in financial intelligence units: experience and vision[C]//Proceedings of the 14th International Rule Challenge, 4th Doctoral Consortium, and 6th Industry Track @ RuleML+RR 2020. Berlin: Springer, 2020: 133-144.

[52] RAJPUT Q, KHAN N S, LARIK A, et al. Ontology based expert-system for suspicious transactions detection[J]. Computer and Information Science, 2013, 7(1): 103-114.

[53] PANIGRAHI S, KUNDU A, SURAL S, et al. Credit card fraud detection:

a fusion approach using Dempster-Shafer theory and Bayesian learning[J]. Information Fusion, 2009, 10(4): 354-363.

[54] LIU F, LI Z H, JIA K, et al. Bitcoin address clustering based on change address improvement[J]. IEEE Transactions on Computational Social Systems, 2024, 11(6): 8094-8105.

[55] WANG Q Y, TSAI W T, SHI T Y. GraphALM: active learning for detecting money laundering transactions on blockchain networks[J]. IEEE Network, 2025, 39(2): 294-303.

[56] OLUTIMEHIN A T. The Synergistic role of machine learning, deep learning, and reinforcement learning in strengthening cyber security measures for crypto currency platforms[J]. Asian Journal of Research in Computer Science, 2025,18(3): 190-212.

[57] SENATOR T E, GOLDBERG H G, WOOTON J, et al. The FinCEN artificial intelligence system: identifying potential money laundering from reports of large cash transactions[C]//Proceedings of the Seventh Conference on Innovative Applications of Artificial Intelligence. Palo Alto: AAAI Press, 1995: 156-170.

[58] SAVAGE D, ZHANG X Z, WANG Q M, et al. Detection of money laundering groups: supervised learning on small networks[J]. Proceedings of the AAAI-17 Workshop on AI and Operations Research for Social Good. Palo Alto: AAAI Press, 2017: 43-49.

[59] ZHANG Y, TRUBEY P. Machine learning and sampling scheme: an empirical study of money laundering detection[J]. Computational Economics, 2019, 54(3): 1043-1063.

[60] ELMOUGY Y, LIU L. Demystifying fraudulent transactions and illicit nodes in the Bitcoin network for financial forensics[C]//Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2023: 3979-3990.

[61] SONG J L, GU Y J. HBTBD: a heterogeneous Bitcoin transaction behavior dataset for anti-money laundering[J]. Applied Sciences, 2023, 13(15): 8766.

[62] CARLETTI R, LUO X J, ADELOPO I. Understanding criminogenic features: case studies of cryptocurrencies-based financial crimes[J]. Journal of Financial Crime, 2025, 32(3): 681-705.

[63] 张俊, 袁勇, 王晓, 等. 量子区块链: 融合量子信息技术的区块链能否抵御量子霸权?[J]. 智能科学与技术学报, 2019, 1(4): 409-414.
ZHANG J, YUAN Y, WANG X, et al. Quantum blockchain: can blockchain integrated with quantum information technology resist quantum supremacy?[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(4): 409-414.

[64] VENČKAUSKAS A, GRIGALIŪNAS Š, POCIUS L, et al. Machine learning in money laundering detection over blockchain technology[J]. IEEE Access, 2024, 13: 7555-7573.

[65] 丁文文, 王帅, 李娟娟, 等. 去中心化自治组织: 发展现状、分析框架与未来趋势[J]. 智能科学与技术学报, 2019, 1(2): 202-213.
DING W W, WANG S, LI J J, et al. Decentralized autonomous organizations: the state of the art, analysis framework and future trends[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(2): 202-213.

[66] SHAFIN K M, RENO S. Integrating blockchain and machine learning for enhanced anti-money laundering system[J]. International Journal of Information Technology, 2025, 17(4): 2439-2447.

[67] 王飞跃. DeepSeek 呼唤 DeepThink: 重视 AI 治理与社会范式变革[J]. 智能科学与技术学报, 2025, 7(1): 1-3.
WANG F Y. DeepSeek calls DeepThink: rethinking AI governance and societal paradigm shift[J]. Chinese Journal of Intelligent Science and Technology, 2025, 7(1): 1-3.

[68] 王飞跃. 新 AI 与智能科技前沿: 从 PAPAI++到新产业生态[J]. 智能科学与技术学报, 2024, 6(4): 413-415.
WANG F Y. New AI for new frontier of intelligent science and technology: from PAPAI++ to new enterprise ecology[J]. Chinese Journal of Intelligent Science and Technology, 2024, 6(4): 413-415.

[69] ALNAQBI M, AL-ALI M M, ALREMEITHI M, et al. Different techniques and algorithms to combat the issue of money laundering in Bitcoin[C]//Proceedings of the 2022 International Conference on Electrical and Computing Technologies and Applications (ICECTA). Piscataway: IEEE Press, 2022: 122-126.

[70] TURNER A B, MCCOMBIE S, UHLMANN A J. Analysis techniques for illicit Bitcoin transactions[J]. Frontiers in Computer Science, 2020, 2: 600596.

[71] DOU Y T, LIU Z W, DENG Y T, et al. Enhancing graph neural network-based fraud detectors against camouflaged fraudsters[C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management. New York: ACM, 2020: 315-324.

[72] 王飞跃, 王艳芬, 陈薏竹, 等. 联邦生态: 从联邦数据到联邦智能[J]. 智能科学与技术学报, 2020, 2(4): 305-311.
WANG F Y, WANG Y F, CHEN Y Z, et al. Federated ecology: from federated data to federated intelligence[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(4): 305-311.

[73] 李娟娟, 王戈, 王晓, 等. 加密管理: 一种基于区块链的新型组织管理模式[J]. 智能科学与技术学报, 2022, 4(2): 145-156.
LI J J, WANG G, WANG X, et al. Crypto management: a novel organizational management model based on blockchain[J]. Chinese Journal of Intelligent Science and Technology, 2022, 4(2): 145-156.

[74] 刘峰, 杨杰, 李志斌, 等. 一种基于区块链的泛用型数据隐私保护的安全多方计算协议[J]. 计算机研究与发展, 2021, 58(2): 281-290.
LIU F, YANG J, LI Z B, et al. A secure multi-party computation protocol for universal data privacy protection based on blockchain[J]. Journal of Computer Research and Development, 2021, 58(2): 281-290.

[75] KONEČNÝ J, MCMAHAN H B, YU F X, et al. Federated learning: strategies for improving communication efficiency[J]. arXiv preprint, 2016: arXiv: 1610.05492.

[76] KOVAČEVIĆ M, GRBIĆ T, ČAPKO D, et al. A zero-knowledge proof for the syndrome decoding problem in the lee metric[J]. arXiv preprint, 2025: arXiv: 2502.11641.

[77] CHENG Y, GUO J J, LONG S Q, et al. Advanced financial fraud detection using GNN-CL model[J]. arXiv preprint, 2024: arXiv: 2407.06529.

[78] LEI Y C, XIANG Y X, WANG Q, et al. Large language models for cryptocurrency transaction analysis: a Bitcoin case study[J]. arXiv preprint, 2025: arXiv:2501.18158.

[79] NICHOLLS J, KUPPA A, LE-KHAC N A. Large language model XAI approach for illicit activity Investigation in Bitcoin[J]. Neural Computing and Applications, 2024, 36(17): 18427-18445.

[80] 陈晓丰, 宋兆雄, 郑佩玉, 等. 一种多链协同治理的“以链治链”监管框架[J]. 计算机研究与发展, 2024, 61(9): 2290-2306.
CHEN X F, SONG Z X, ZHENG P Y, et al. A multichain-collaborating governing chain-supervising-chain supervision framework[J]. Journal of Computer Research and Development, 2024, 61(9): 2290-2306.

[81] LIU F, HE S H, LI Z H, et al. An overview of blockchain efficient interaction technologies[J]. Frontiers in Blockchain, 2023, 6: 996070.

[82] 刘峰, 张嘉淏, 周俊杰, 等. 基于改进哈希时间锁的区块链跨链资产交互协议[J]. 计算机科学, 2022, 49(1): 336-344.
LIU F, ZHANG J H, ZHOU J J, et al. Novel Hash-time-lock-contract based cross-chain token swap mechanism of blockchain[J]. Computer Science, 2022, 49(1): 336-344.

[作者简介]

张昳斐（1997– ），女，中国人民大学数学学院硕士生，主要研究方向为区块链与加密货币监管。

杨东（1975– ），男，博士，中国人民大学法学院教授，主要研究方向为金融科技、区块链、数字货币。

袁勇（1980– ），男，博士，中国人民大学数学学院教授，主要研究方向为区块链、计算经济学和分布式人工智能。

王飞跃（1961– ），男，博士，中国科学院自动化研究所复杂系统管理与控制国家重点实验室主任，澳门科技大学特聘教授，主要研究方向为平行系统的方法与应用、社会计算、平行智能、知识自动化。