北京大学
PEKING UNIVERSITY

区块链网络异常检测溯源

李锦凯

2025.2.6

Benchmark研究

北京大学 PEKING UNIVERSITY

Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics

Mark Weber*
MIT-IBM Watson AI Lab
mrweber@mit.edu

Giacomo Domeniconi*
IBM Research
Giacomo.Domeniconi1@ibm.com

Jie Chen
MIT-IBM Watson AI Lab
chenjie@us.ibm.com

Daniel Karl I. Weidele
IBM Research AI
daniel.karl@ibm.com

Claudio Bellei
Elliptic
claudio@elliptic.co

Tom Robinson
Elliptic
tom@elliptic.co

Charles E. Leiserson
MIT CSAIL
cel@mit.edu

MIT-IBM Lab联合工作，该领域被引次数最多

探讨如何利用图卷积网络（Graph Convolutional Networks, GCN）和其他机器学习方法来检测比特币交易中的洗钱行为，以支持金融反洗钱（Anti-Money Laundering, AML）工作。

Weber, Mark, et al. "Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics." arXiv preprint arXiv:1908.02591 (2019).

背景知识

北京大学 PEKING UNIVERSITY

- **金融包容性与AML的矛盾**：AML法规对保护金融系统至关重要，但增加了金融机构的成本，并导致社会边缘群体的金融排斥。约有17亿成年人无法获得银行服务。
- **加密货币的兴起**：比特币等加密货币的出现为支付处理带来了技术革新，但其匿名性也使得犯罪分子能够隐藏在众目睽睽之下。
- **AML的挑战**：AML需要在保护金融系统安全和促进金融包容性之间找到平衡。

数据集

北京大学 PEKING UNIVERSITY

首个大规模公开数据集

•Elliptic数据集：包含203,769个节点（比特币交易）和234,355条有向边（支付流向），其中2%被标记为非法交易，21%为合法交易。每个节点有166个特征，包括本地信息和聚合信息。

1.Weber, Mark, et al. "Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics." arXiv preprint arXiv:1908.02591 (2019).

2.Pareja, Aldo, et al. "EvolveGCN: Evolving graph convolutional networks for dynamic graphs." Proceedings of the AAAI Conference on Artificial Intelligence.

3.https://github.com/Rufaim/EvolveGCN

机器学习baseline

北京大学 PEKING UNIVERSITY

•**机器学习方法：**使用逻辑回归（LR）、随机森林（RF）、多层感知机（MLP）

•**数据划分**：采用70:30的时间划分，前34个时间步用于训练，后15个用于测试。
•**评估指标**：主要关注非法交易的精确率、召回率和F1分数。

Method	Illicit			MicroAVG
	Precision	Recall	F_1	F_1
Logistic Regr^{AF}	0.404	0.593	0.481	0.931
Logistic Regr^{AF+NE}	0.537	0.528	0.533	0.945
Logistic Regr^{LF}	0.348	0.668	0.457	0.920
Logistic Regr^{LF+NE}	0.518	0.571	0.543	0.945
RandomForest^{AF}	0.956	0.670	0.788	0.977
RandomForest^{AF+NE}	0.971	0.675	0.796	0.978
RandomForest^{LF}	0.803	0.611	0.694	0.966
RandomForest^{LF+NE}	0.878	0.668	0.759	0.973

Graph Convolutional Networks (GCN)

北京大学 PEKING UNIVERSITY

GCN的核心思想

通过图卷积操作，将节点的特征与其邻域信息结合起来，从而生成更丰富的节点表示。具体来说，GCN通过以下步骤实现：

- **图卷积操作**：每一层的GCN会聚合节点的特征以及其邻域节点的特征，通过一个可学习的权重矩阵进行变换，然后通过非线性激活函数（如ReLU）引入非线性。
- **邻域聚合**：通过邻接矩阵和归一化操作，将节点的邻域信息聚合到节点的特征表示中。

Graph Convolutional Networks (GCN)

北京大学 PEKING UNIVERSITY

模型构建

•输入特征：使用节点的166个特征作为输入。
•图卷积层：构建一个两层的GCN模型。第一层的输出作为第二层的输入，最终输出层使用softmax函数进行分类。

训练过程

•损失函数：由于数据集中的类别不平衡（非法交易较少），使用加权交叉熵损失函数，为非法样本分配更高的权重。
•优化器：使用Adam优化器，学习率为0.001。
•训练时间：训练1000个epoch。

改进GCN

北京大学 PEKING UNIVERSITY

时序建模的关键点

•金融数据具有时间戳特性，系统动态变化隐含其中。捕捉这些动态变化的模型能更好地进行预测，并在不同时间段上泛化。

GCN与时序扩展：

•传统GCN：图卷积网络（GCN）是一种强大的工具，用于处理图结构数据，但它通常是静态的，无法直接处理随时间变化的数据。

•**EvolveGCN：**为了解决这一问题，提出了EvolveGCN，它通过在每个时间步计算一个独立的GCN模型，并通过循环神经网络（RNN）连接这些GCN模型来捕捉系统的动态变化。

GCN v.s. EvolveGCN

北京大学 PEKING UNIVERSITY

EvolveGCN的核心思想

•独立GCN模型：在每个时间步计算一个独立的GCN模型，以处理该时间点的图结构和节点特征。

•RNN连接GCN模型：

•使用循环神经网络（如GRU）将这些独立的GCN模型连接起来，捕捉从过去到未来的系统动态变化。

•系统状态更新：

•GCN的权重被视为系统的状态，在每次输入新的图信息时通过RNN更新。当前时间步的GCN模型从过去的模型演化而来，反映系统动态。

•图信息表示：

•当前时间步的图信息由最具影响力的前k个节点的嵌入表示。

Table 2: GCN v.s. EvolveGCN

	GCN			EvolveGCN		
	Precis.	Recall	F_1	Precis.	Recall	F_1
Illicit	0.812	0.623	0.705	0.850	0.624	0.720
MicroAVG	0.966	0.966	0.966	0.968	0.968	0.968

关键结论

Table 1: Illicit classification results. Top part of the table shows results without the leverage of the graph information, for each model are shown results with different input: *AF* refers to all features, *LF* refers to the local features, i.e. the first 94, and *NE* refers to the node embeddings computed by GCN. Bottom part of the table shows results with GCN.

Method	Illicit			MicroAVG
	Precision	Recall	F_1	F_1
Logistic Regr^{AF}	0.404	0.593	0.481	0.931
Logistic Regr^{AF+NE}	0.537	0.528	0.533	0.945
Logistic Regr^{LF}	0.348	0.668	0.457	0.920
Logistic Regr^{LF+NE}	0.518	0.571	0.543	0.945
RandomForest^{AF}	0.956	0.670	0.788	0.977
RandomForest^{AF+NE}	0.971	0.675	0.796	0.978
RandomForest^{LF}	0.803	0.611	0.694	0.966
RandomForest^{LF+NE}	0.878	0.668	0.759	0.973
MLP^{AF}	0.694	0.617	0.653	0.962
MLP^{AF+NE}	0.780	0.617	0.689	0.967
MLP^{LF}	0.637	0.662	0.649	0.958
MLP^{LF+NE}	0.6819	0.5782	0.6258	0.986
GCN	0.812	0.512	0.628	0.961
Skip-GCN	0.812	0.623	0.705	0.966

- **RF的优越性**：RF在基准测试中表现最佳，但GCN和Skip-GCN在利用图结构信息方面具有潜力。
- **特征增强的有效性**：将GCN的节点嵌入与原始特征结合可以提升模型性能。
- **时间建模的重要性**：EvolveGCN通过捕捉系统动态，提升了模型对非法交易的检测能力。

TMAS：A transaction misbehavior analysis scheme for blockchain

ELSEVIER

Blockchain: Research and Applications

Volume 5, Issue 3, September 2024, 100197

Research Article

TMAS: A transaction misbehavior analysis scheme for blockchain

Shiyong Huang [a], Xiaohan Hao [b], Yani Sun [c], Chenhuang Wu [d], Huimin Li [d], Wei Ren [a c d] ✉, Kim-Kwang Raymond Choo [e]

基于多种特征模式匹配的研究方案

Huang, S., Hao, X., Sun, Y., Wu, C., Li, H., Ren, W., & Choo, K.-K. R. (2024). TMAS: A transaction misbehavior analysis scheme for blockchain. Blockchain: Research and Applications, 5(2024), 100197. https://doi.org/10.1016/j.bcra.2024.100197

背景

北京大学 PEKING UNIVERSITY

- 区块链技术因其去中心化和匿名性而被广泛应用于加密货币（如比特币），但这些特性也使其容易被用于非法活动，如洗钱。
- 为了应对这一问题，许多国家的政府已经开始探索和实施对加密货币的监管措施。
- 区块链的不可篡改和不可变特性使得交易数据可以被公开验证和分析，为检测非法活动提供了可能。

研究方法

TMAS（A transaction misbehavior analysis scheme ）

方案：该方案包括十个交易图特征、两个启发式洗钱模型和一种账户关联分析方法，用于识别由同一实体控制的不同账户。

- **交易图特征**：包括节点的入度和出度、频繁交易节点、间接交易节点、循环网络、自交易节点、大额交易节点、频繁交易节点、多账户拥有者、双伞模型等。
- **洗钱模型**：
  - **双伞模型**：基于赌场洗钱的现实场景，模拟非法资金通过多个账户分散和聚集的过程。
  - **内部转账模型**：当一个用户拥有多个账户时，通过内部转账分散资金，将非法资金转化为合法收入。

系统设计

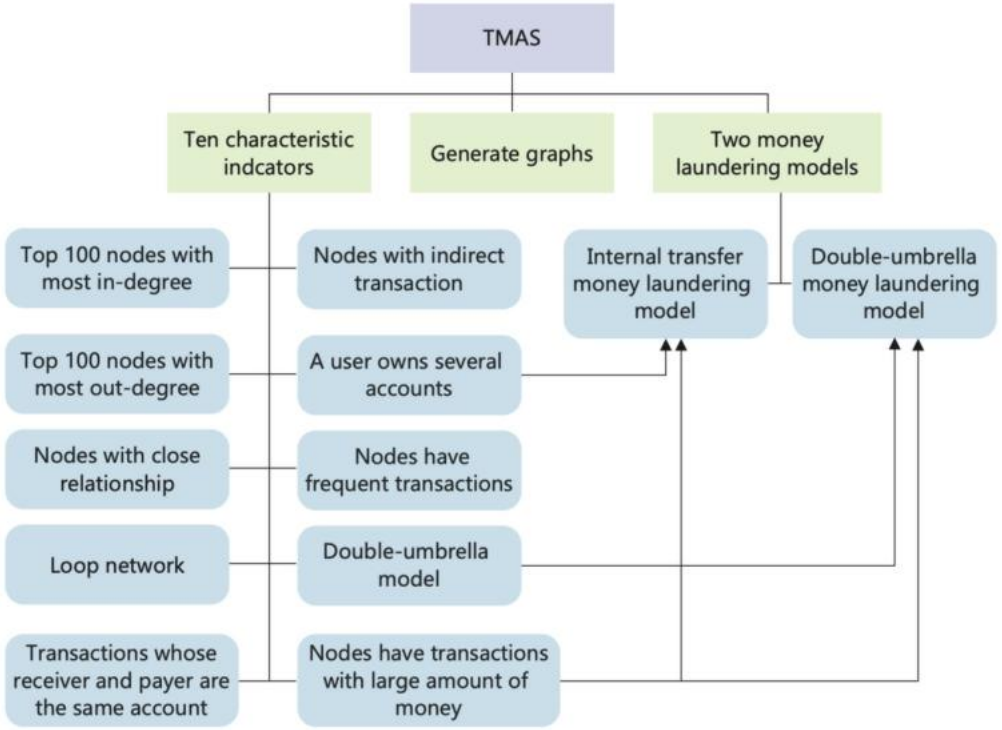

Fig. 1. Overall design of the system.

- **频繁交易节点：**识别出交易频繁的节点，可能代表交易所或矿工。
- **大额交易：**识别出交易金额超过500 BTC的交易。
- **账户间关系紧密度：**识别出交易频繁的账户对。
- **多账户用户：**识别出同一用户拥有的多个账户。
- **间接交易和循环网络：**识别出通过多个中间账户完成的交易路径。
- **双伞模型：**识别出符合双伞洗钱模型的交易模式。
- **内部转账洗钱模型：**识别出同一用户账户之间的大额交易。

Double-umbrella money laundering model

双伞洗钱模型（Double-umbrella money laundering model）是文章中提出的一种用于检测区块链加密货币交易中洗钱行为的模型。该模型的设计灵感来源于现实中的赌场洗钱场景，模拟了非法资金通过复杂的交易路径被分散和重新聚集的过程。

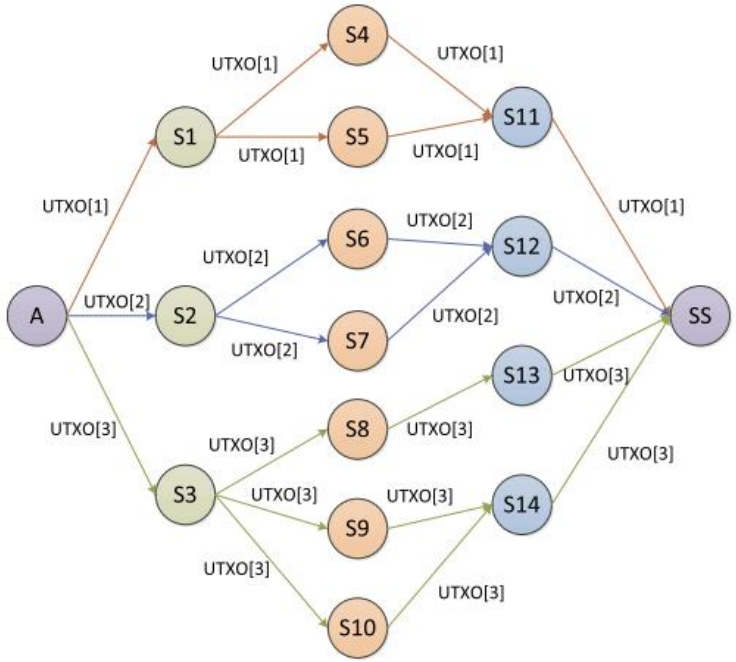

Fig. 4. The case of double-umbrella model 1.

为了检测双伞洗钱模型，文章提出了以下步骤：

1.选择关键账户：
◦选择交易频繁、交易金额大或在异常时间段内进行交易的账户作为起点主账户。
◦通过分析这些账户的交易路径，寻找可能的终点主账户。

2.跟踪交易路径：
◦通过追踪未花费交易输出（UTXO）来确定资金的流动路径。
◦记录每个交易路径的中间账户和交易金额。

3.识别双伞结构：
◦通过分析交易路径，寻找符合双伞模型的结构。
◦检查是否存在从起点主账户到多个中间账户，再从这些中间账户到终点主账户的资金流动。

4.评估模型：
◦根据交易路径的长度、中间账户的数量和交易金额等特征，评估交易是否符合双伞洗钱模型。
◦如果交易路径符合双伞模型且涉及大额资金，则标记为可疑交易。

实验结果

数据集：使用以太坊区块链平台的1亿笔交易数据进行分析。
工具：使用Gephi软件生成交易关系图，展示账户之间的交易路径和关系。

•节点的入度和出度：
•识别出交易频繁的节点，这些节点在交易图中具有高入度和出度。
•示例：节点1的入度和出度总和为65,960，表明该账户非常活跃，可能是交易所或矿工账户。
•在1亿笔交易中，发现了一个符合双伞模型的交易案例，该案例属于多账户模型。
•该模型展示了资金从多个起点账户通过多个中间账户转移到多个终点账户的过程。
•识别出资金从一个账户出发，经过多个中间账户后返回到同一个账户的交易路径。

优劣对比

北京大学 PEKING UNIVERSITY

TMAS方案的优势

1.规则明确、可解释性强：TMAS方案通过定义明确的特征和模型来检测非法交易，如双伞洗钱模型和内部转账洗钱模型。这些规则基于已知的洗钱行为模式，易于理解和解释，监管机构和审计人员可以清楚地看到检测逻辑。

2.特征针对性强：TMAS方案提取了多个与非法交易相关的特征，如节点的入度和出度、交易金额、账户间关系紧密度等，这些特征直接针对洗钱等非法行为的典型特征，能够有效识别异常交易。

3.无需大量数据和复杂训练：与AI模型相比，TMAS方案不需要大量的历史数据进行训练，也不需要复杂的机器学习算法。这使得方案在数据有限或计算资源受限的情况下仍然可以有效运行。

优劣对比

北京大学 PEKING UNIVERSITY

TMAS方案的劣势

1.灵活性不足：TMAS方案依赖于预定义的规则和特征，难以适应不断变化的非法交易模式。一旦出现新的洗钱手法，就需要手动更新规则和模型。

2.难以处理复杂模式：对于一些复杂的、非线性的交易模式，TMAS方案可能无法有效识别。例如，当非法交易通过复杂的网络结构和多层次的交易路径进行时，TMAS方案可能无法准确检测。

3.误报率可能较高：由于TMAS方案基于规则，可能会将一些正常交易误判为非法交易，尤其是在交易模式与规则部分匹配的情况下。

创新点1：强化利用时间特征

北京大学 PEKING UNIVERSITY

利用时空卷积网络（STCNs）

核心概念

•空间特征：在区块链中，空间特征主要指交易网络的结构特征，例如交易之间的连接关系、节点（地址）之间的关系等。
•时间特征：时间特征指交易发生的时间顺序和时间间隔，例如交易的时间戳、交易频率等。
•时空特征：时空特征结合了空间和时间信息，能够更全面地描述交易行为。例如，一个交易不仅与它直接连接的交易有关，还可能受到过去一段时间内其他交易的影响。

创新点1：强化利用时间特征

网络结构：

STCNs通常由以下几部分组成：

•时空卷积层（Spatio-Temporal Convolutional Layer）：通过卷积操作同时处理空间和时间特征。例如，可以使用二维卷积核在交易网络的时间序列上滑动，提取时空特征。

•池化层（Pooling Layer）：用于降采样，减少特征维度，提高计算效率。

•全连接层（Fully Connected Layer）：将提取的时空特征映射到输出空间，例如分类结果（正常交易或异常交易）。

•激活函数（Activation Function）：如ReLU，用于引入非线性，增强模型的表达能力。

创新点2：结合隐私保护

北京大学 PEKING UNIVERSITY

同态加密

背景：同态加密允许在加密数据上直接进行计算，而无需解密数据。这在区块链异常检测中尤为重要，因为交易数据通常包含敏感信息。

应用：在STCNs中，可以使用同态加密技术对输入数据进行加密，确保在模型训练和检测过程中数据的隐私性。例如，使用加法同态加密（如Paillier加密）对交易金额进行加密，然后在加密域内进行卷积操作。

零知识证明

背景：零知识证明允许一方在不泄露任何有用信息的情况下证明某个陈述的真实性。

应用：在STCNs中，可以使用零知识证明技术验证交易的合法性，而无需透露交易的具体内容。例如，通过零知识证明验证交易的签名和金额，确保交易的合规性。

创新点3：TMAS规则匹配结合AI模型

将AI模型算法与TMAS方案中的特征模型匹配相结合，可以充分发挥AI的强大学习能力和TMAS方案中特征提取的针对性，从而更有效地检测区块链中的非法交易。

1. 特征提取与增强

•**利用TMAS特征作为输入**：TMAS方案已经提取了十个关键特征，如节点的入度和出度、交易金额、账户间关系紧密度等。这些特征可以直接作为AI模型的输入，为模型提供丰富的先验知识。

•**AI模型的特征增强**：使用深度学习模型（如ResNet）进一步提取和融合这些特征，以获得更高级的抽象特征。例如，通过残差网络结构ResNet-32，可以自动挖掘特征间的复杂关联关系，学习到包含丰富语义信息的高层抽象特征。

创新点3：TMAS规则匹配结合AI模型

北京大学 PEKING UNIVERSITY

2. 模型融合与集成

•**多视图特征融合**：结合TMAS方案中的特征和AI模型提取的特征，形成多视图特征表示。

•**Transformer模型**：基于Transformer的模型可以学习多视图特征的全局结构和语义特征，捕获视图特征中Token的远程依赖关系。通过对比学习的多视图特征半监督训练方法，可以更准确地捕获底层语义信息，从而更准确地预测非法交易。

3. 可解释性与规则集成

•**可解释AI技术**：将可解释人工智能（XAI）技术集成到AI模型中，以提高模型预测的可解释性。

•**规则与模型的协同**：将TMAS方案中的规则（如双伞洗钱模型和内部转账洗钱模型）与AI模型的预测结果相结合。例如，当AI模型检测到潜在的异常交易时，可以进一步使用TMAS方案中的规则进行验证，从而提高检测的准确性和可信度。

欢迎指导

李锦凯
jinkaili@stu.pku.edu.cn