

УДК 004.85

Методы классификации сервисов в сети биткоин на основе статистических признаков и векторных представлений

Кузовчиков Максим Евгеньевич mkuzovchikov@gmail.com

Вишняков Игорь Эдуардович vishnyakov@bmstu.ru
SPIN-код: 5500-1834

Белова Наталья n.virtseva@bmstu.ru
SPIN-код: 3582-3061

МГТУ им. Н.Э. Баумана, Москва, Россия

Аннотация. Проведено исследование методов классификации биткоин-адресов, связанных с различными сервисами, на основе статистических признаков транзакционной активности и на основе векторных представлений эго-сетей адресов. Результаты показали, что применение эмбедингов и нейронных сетей обеспечивает значительно более высокие показатели классификации. Работа подтверждает эффективность графовых моделей и машинного обучения для анализа структуры транзакций в сети биткоин.

Ключевые слова: биткоин, анализ транзакций, сервисы-миксеры, машинное обучение, классификация адресов, эго-сети, graph2vec, эмбединги

Введение. В настоящее время растет интерес к анализу транзакций в сети биткоин, а особое внимание уделяется сервисам-миксерам [1], которые повышают анонимность путем перемешивания средств различных пользователей. В работе рассмотрены методы классификации биткоин-адресов сервисов на основе различных признаков их транзакционной активности.

Методы и материалы; результаты. Для классификации были собраны данные из открытого источника WalletExplorer [2], включающие адреса миксеров, биржевых сервисов и игровых платформ. Рассматривались два подхода к классификации: на основе статистических признаков и на основе векторных представлений эго-сетей адресов.

В первом случае для каждого адреса формировался набор из 8 признаков: общее количество транзакций адреса в качестве отправителя или получателя, сумма средств, поступивших и выведенных с адреса, среднее значение входящих и исходящих переводов, время между первой и последней транзакцией адреса. Результаты классификации представлены в табл. 1.

Второй подход основывался на построении эго-сетей адресов. Эго-сеть адреса представляет собой локальный подграф, строящийся вокруг выбранного узла и включающий его соседей. Эго-сети адресов использовались для извлечения векторных представлений с помощью алгоритма graph2vec [3]. Полученные эмбединги использовались для классификации адресов с помощью моделей машинного обучения. Результаты представлены в табл. 2.

Таблица 1

Результаты классификации по статистическим признакам

Модель	Доля верных классификаций	Точность	Полнота	F_1 -мера
Логистическая регрессия	0,431	0,517	0,497	0,408
Градиентный бустинг	0,720	0,690	0,534	0,556
Случайный лес	0,629	0,508	0,503	0,504

Таблица 2

Результаты классификации на основе векторных представлений

Модель	Доля верных классификаций	Точность	Полнота	F_1 -мера
Логистическая регрессия	0,854	0,755	0,818	0,775
Градиентный бустинг	0,959	0,966	0,846	0,874
Случайный лес	0,940	0,904	0,830	0,857
Нейронная сеть	0,979	0,951	0,952	0,951

Использование векторных представлений позволило значительно улучшить показатели. Нейронная сеть, обученная на эмбедингах, достигла точности до 0,98 и показала устойчивость к дисбалансу данных. Это подтверждает эффективность применения графовых моделей и эмбедингов для анализа структуры транзакций.

Заключение. Данная работа показала, что методы машинного обучения позволяют эффективно классифицировать биткоин-адреса сервисов. Подход на основе статистических признаков обеспечивает базовый уровень точности. Применение эмбедингов на основе алгоритма graph2vec дает возможность достичь высокой точности и устойчивости к шуму и дисбалансу данных.

Список источников

- [1] Shojaeinasab A., Motamed A., Bahrak B. Mixing detection on bitcoin transactions using statistical patterns. IET Blockchain, 2023, vol. 3, pp. 136–148. <https://doi.org/10.48550/arXiv.2204.02019>
- [2] Smart Bitcoin block explorer. URL: <https://www.walletexplorer.com> (accessed 20.05.2025).
- [3] Narayanan A., Chandramohan M., Venkatesan R., Chen L. Graph2vec: learning distributed representations of graphs. 13th International Workshop on Mining and Learning with Graphs, 2017, pp. 1–8. <https://doi.org/10.48550/arXiv.1707.05005>