ПРИКЛАДНЫЕ ОНТОЛОГИИ ПРОЕКТИРОВАНИЯ

УДК 004.89 *Научная статья* DOI: 10.18287/2223-9537-2025-15-3-334-350

Обнаружение аномальных транзакций криптовалюты с помощью нейронных сетей и онтологий

© 2025, И.В. Котенко ✉, Д.С. Левшун, К.Н. Жернова, А.А. Чечулин

Санкт-Петербургский федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН), Санкт-Петербург, Россия

Аннотация

Выполнен поиск эффективного подхода к выявлению аномалий в транзакциях криптовалют с помощью нейронных сетей (свёрточных, глубоких, управляемых рекуррентных блоков) и сравнение с другими подходами, применяемыми к задаче поиска нелегальных транзакций в сети криптовалют. Проведён анализ работ в областях анализа данных транзакций в сети криптовалют, визуализации данных для анализа транзакций и поиска аномальных транзакций с помощью компьютерного зрения. Определена предметная область исследования. Проблема выявления аномалий в транзакциях криптовалют основывается на том, что в наборах данных транзакций нет прямых указаний на личность отправителя и получателя. Актуальность и значимость работы заключается в предлагаемом в ней способе выявления с высокой точностью аномальных транзакций в режиме, приближенном к реальному времени. Выполнены экспериментальные исследования набора данных транзакций криптовалют с помощью нейронных сетей и не нейросетевых классификаторов с последующим сравнением полученных результатов с другими исследованиями. Эксперименты показали, что управляемые рекуррентные блоки позволили справиться с задачей лучше других сравниваемых нейросетевых подходов: аккуратность 0.94, точность 0.95, полнота 0.93 и *F*-мера 0.94, что доказывает высокую эффективность предложенной модели. Однако данная модель уступает традиционным алгоритмам машинного обучения, таким как оптимизированный распределённый градиентный бустинг. Новизной предложенного подхода является то, что он основан на анализе статистической информации о графе транзакций и использовании для этого технологии глубокого обучения и градиентного бустинга. Областью применения предложенного подхода является создание программных средств для поиска нелегальных транзакций криптовалюты в системах информационной безопасности и в задачах цифровой криминалистики.

Ключевые слова: *поиск аномалий, компьютерная безопасность, машинное обучение, нейронные сети, визуальный анализ данных, криптовалюта, онтологии.*

Цитирование: *Котенко И.В., Левшун Д.С., Жернова К.Н., Чечулин А.А.* Обнаружение аномальных транзакций криптовалюты с помощью нейронных сетей и онтологий. *Онтология проектирования*. 2025. Т.15, №3(57). С.334-350. DOI:10.18287/2223-9537-2025-15-3-334-350.

Финансирование: исследование выполнено при поддержке совместного гранта Российского научного фонда и Санкт-Петербургского научного фонда № 24-21-20058, https://rscf.ru/project/24-21-20058/.

Вклад авторов: *Котенко И.В.* – постановка проблемы, общее руководство работой и планирование экспериментов; *Левшун Д.С.* – проведение эксперимента и интерпретация результатов; *Жернова К.Н.* – поиск и описание набора данных для эксперимента, обзор релевантных работ; *Чечулин А.А.* – разработка концепции исследования, подготовка набора данных.

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов.

Введение

Криптовалюта, как удобный способ осуществления платежей без участия посредников, имеет недостатки, в т.ч. в контексте информационной безопасности. Использование криптовалют имеет устойчивый рост, в т.ч. и преступлений, связанных с их использованием. В 2022 году был зафиксирован очередной всплеск киберпреступлений с их участием, а в 2023 и 2024 годах уровень таких инцидентов оставался высоким[1]. Таким образом, данная область требует принятия как дополнительных законодательных мер[2], так и новых решений в области информационной безопасности.

Выявление аномальных транзакций в режиме реального времени остаётся сложной задачей. В связи с этим обнаружение нелегальных цепочек транзакций, как правило, осуществляется постфактум и становится предметом исследования в области цифровой криминалистики. Основные методы выявления подозрительных транзакций в цифровой криминалистике включают: применение алгоритмов искусственного интеллекта (ИИ), мониторинг сетевой активности и анализ графов транзакций или пользователей.

В данной работе предложен подход к обнаружению нелегальных транзакций, основанный на анализе статистической информации о графе транзакций и использовании для этого технологии глубокого обучения и градиентного бустинга над решающими деревьями.

1 Онтологическая схема связей

Для описания предлагаемого подхода необходимо очертить границы предметной области (ПрО) в терминах понятий, их атрибутов и взаимосвязей. Это позволит сформировать концептуальную модель ПрО, которая может быть использована в т.ч. для решения других подобных задач. Схема взаимосвязей сущностей, относящихся к данной ПрО, и их атрибутов приведена на рисунке 1.

Криптовалюта – цифровая валюта, которая создаётся с помощью криптографических алгоритмов и содержится в специальном виртуальном *кошельке* пользователя, посредством которого можно совершать оплату или на который можно принимать средства, используя *приватный ключ* для авторизации. Цифровая валюта создаётся посредством криптографических вычислений на основе *хэш-функции*. Основными атрибутами криптовалюты являются *хэш* (результат вычислений с помощью хэш-функции), а для кошелька атрибутами являются *идентификатор* владельца (отправителя или получателя) и его *приватный ключ*.

Хэш-функция транзакции – специальный алгоритм для вычисления определённой буквенно-цифровой последовательности заданной длины на основе данных о транзакциях в предыдущем блоке. Полученная буквенно-цифровая последовательность называется *хэш*, который создаётся с помощью хэш-функции, и при его вычислении создаётся криптовалюта (отношения 1 и 2 см. рисунок 1).

Сеть криптовалют – сеть адресов кошельков пользователей, по которой могут осуществляться транзакции криптовалюты. С одной стороны, сеть обеспечивает быстроту и безопасность денежных переводов, с другой – позволяет использовать криптовалюту для оплаты нелегальных схем ввиду сложности выявления юридической личности отправителя и получателя.

Транзакция криптовалюты – операция платежа, получения или обмена криптовалюты. Осуществление транзакции криптовалюты происходит следующим образом[3]: начало транзакции, проверка приватного ключа, отправка/получение, подтверждение правильности транзакции, завершение транзакции. Атрибуты транзакции – это *идентификаторы отправителя* и *получателя, интервал* между транзакциями и *количество транзакций*. *Псевдонимность* криптовалютных транзакций достигается тем, что они записываются с использованием криптографических адресов – случайных наборов символов, выступающих в роли цифровых идентификаторов пользователей.

[1] 2024 Crypto Crime Trends: Illicit Activity Down as Scamming and Stolen Funds Fall, But Ransomware and Darknet Markets See Growth. 2024. January 18, 2024. https://www.chainalysis.com/blog/2024-crypto-crime-report-introduction/.

[2] *Лузгин Андрей*. В России готовятся принять регулирование криптовалют. Что нужно знать. 17 июля 2024 г. https://www.rbc.ru/crypto/news/6697b1f79a794796605a1a81.

[3] Что такое криптовалютная транзакция и как она работает. 18 декабря 2024 г. https://cryptomus.com/ru/blog/what-is-a-cryptocurrency-transaction-and-how-does-it-work.

Приватный ключ – код, который требуется для авторизации в сети криптовалют для отправки и получения криптовалюты.

Паттерн транзакций – набор характеристик, присущий цепочке транзакций: интервал между транзакциями и их количество.

Нормальная транзакция – транзакция, характеристики которой соответствуют типичным значениям, присущим законным операциям в сети криптовалют.

Аномальная транзакция – транзакция, характеристики которой отличаются от обычных характеристик транзакций. Аномальные транзакции могут оказаться нелегальными.

Нелегальная транзакция – аномальная транзакция, которая направлена на финансирование незаконной деятельности.

Отмывание денег на основе криптовалюты – совершение транзакций криптовалюты для придания правомерного вида владению, пользованию или распоряжению имуществом, приобретённым преступным путём, либо помощи лицу, совершившему преступление, избежать уголовной ответственности за его совершение.

Блокчейн – это технология выстраивания цепочки *блоков*, в которых совершаются вычисления посредством криптографических алгоритмов, и с помощью хэш-функции шифруется информация о предыдущем блоке. *Децентрализованность* обеспечивается принципами технологии блокчейна: сеть криптовалюты представляет собой распределённый реестр данных, поддерживаемый в актуальном состоянии. Особенностью блокчейна является то, что каждый последующий блок содержит информацию о предыдущем – в частности, хэш-значение, вычисленное на основе данных о транзакциях, зафиксированных в предыдущем блоке.

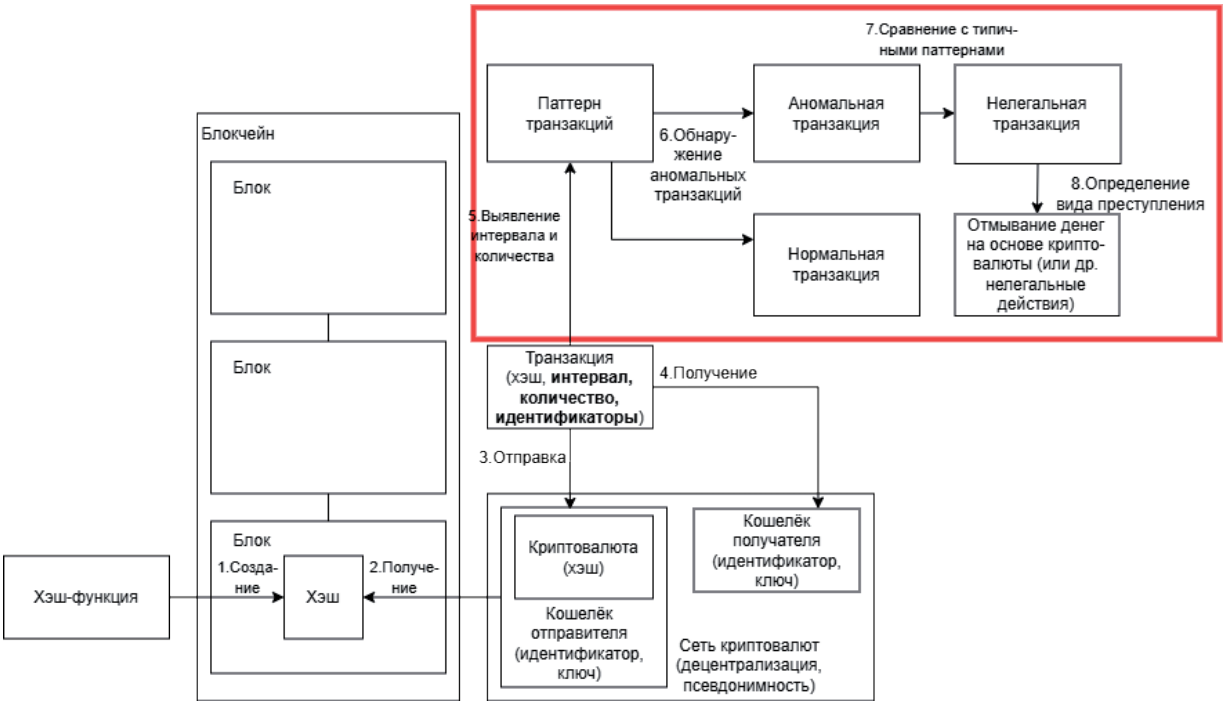


Рисунок 1 – Онтологическая схема связей в области обнаружения аномальных транзакций криптовалюты

На рисунке 1 ключевые понятия представлены в виде прямоугольных блоков; в круглых скобках указаны значимые их атрибуты. Взаимосвязи между понятиями обозначены стрелками и пронумерованы для удобства обращения к ним. Отношения иерархической принадлежности одного понятия другому отражены графически с помощью вложенности блоков (например, хэш принадлежит блоку, блок – блокчейну; криптовалюта – виртуальному кошельку, а кошелёк – сети криптовалют).

В большом прямоугольнике сверху справа показан участок области используемого в данной работе подхода, который основан на отношениях 5–8. Цель подхода заключается в выявлении аномальных транзакций на основе их статистических данных (отношение 6). Под статистическими данными транзакции понимаются данные, которые вычислены на основе таких важных для полученной схемы атрибутов транзакции, как интервал между транзакци-

ями, количество транзакций и идентификаторы отправителя и получателя. Чтобы обнаружить аномалию, необходимы статистические данные о транзакции, поэтому важно отношение 5 – выявление интервала между транзакциями и их количества. При этом не любая аномальная транзакция может оказаться нелегальной. Для того чтобы достоверно определить её нелегальность, нужно сравнить её поведение с паттернами, типичными для нелегальных транзакций (отношение 7). Отношение 8 показывает, что последующий анализ позволит определить конкретный тип операции с криптовалютой.

2 Релевантные работы

Приведён анализ литературы, связанной с отслеживанием незаконных транзакций биткоина с помощью машинного обучения (МО), в т.ч. посредством нейронных сетей (НС). Представлены работы в данной области, в которых изучаются методы обнаружения легализации доходов, полученных преступным путём. Криптовалюта используется в разных преступных схемах, среди которых: финансирование терроризма[1]; шантаж, торговля на чёрном рынке оружием и нелегальными препаратами, финансовые пирамиды [1]; распространение вредоносного программного обеспечения [2] и т.д. Для определения степени легальности транзакции существуют реестры легитимных и нелегитимных адресов [4, 5], однако количество адресов постоянно увеличивается, что ограничивает эффективность подобных списков.

Для обнаружения аномалий в сетях криптовалют используются четыре основных метода: поведенческий анализ на основе построения графа транзакций и применение алгоритмов МО; визуальный анализ графа транзакций; статистический анализ на основе проверки выполнения наборами данных определённого правила; применение методов МО к несвязанным данным о транзакциях. Часто эти методы объединяются.

Большая часть работ, в которых исследовалась возможность обнаружения преступлений, связанных с криптовалютой, посвящена способам их обнаружения с помощью методов ИИ, таких как использование различных классификаторов.

В [1] используются ориентированные графы для отображения графов сети биткоина, основываясь на предположении, что характеристики подграфов этого графа, содержащих адреса, расположенные близко к какому-либо адресу, отражают некоторую информацию об этом адресе. Применяются ансамблевые методы. Так, в [3] применено ансамблевое обучение путём объединения моделей МО, таких как деревья решений, наивный байесовский классификатор, *k* ближайших соседей (*k-NN*) и случайный лес (*Random Forest, RF*) для обнаружения мошенничества в транзакциях биткоина. В целях балансировки данных используется генерация синтетических данных и случайное удаление данных. В [4] предложена модель из ансамбля деревьев решений, которая обучается на отличительных признаках различных групп нелегальных пользователей с целью отличить их от законных.

В [5] использовано построение графа транзакций, по которому определяются пути переводов криптовалюты с одного адреса на другой. Этот метод включает сравнение адресов со списком, представленным как результат кампаний по борьбе с киберпреступностью. В [6] применены неразмеченные данные, приближенные к реальной ситуации, когда нет доступа к ярлыкам. Используются методы обучения без учителя для обнаружения транзакций, связанных с отмывание денег.

Исследование [7] посвящено обнаружению программ-вымогателей. Для этой цели проводится сравнение эффективности логистической регрессии, алгоритма *RF* и экстремального градиентного бустинга, при этом последняя модель оказалась наиболее точной.

[4] *Chainabuse*. https://www.chainabuse.com.
[5] *WalletExplorer. Smart Bitcoin block explorer*. https://www.walletexplorer.com.

В [8] предложена модель обнаружения кражи биткоинов на основе признаков транзакций при краже криптовалюты. Сравнивались различные алгоритмы, основанные на обучении с учителем и без него. Результаты использования методов на основе обучения с учителем в целом оказались лучше. В качестве набора данных использовался направленный граф транзакций, где вершины – это хэш транзакции, рёбра – потоки криптовалюты, вес ребра – количество переведённой криптовалюты. При этом временной интервал между созданием транзакции и переводом максимально большого количества криптовалюты был взят в качестве основного признака.

Система обнаружения мошенничества, представленная в [9], основана на алгоритмах *RF* и градиентного бустинга. С помощью построенной модели, обученной на основе данных, содержащих паттерны мошеннических транзакций, спроектирована система, которая может прогнозировать связанность новых транзакций с мошенничеством. Рассматриваются такие мошеннические схемы, как, например, возможность потратить одни и те же биткоины дважды [10].

Исследуется также выявление бесфайловых атак для осуществления майнинга [11]. В предложенном классификаторе используются модели: метод опорных векторов, *k-NN* и *RF*.

Часть работ посвящена поиску аномалий в транзакциях с помощью искусственных НС. Чаще всего подобные методы основаны на анализе графа транзакций. Для обнаружения незаконных транзакций используются графовые НС и графовые сети на внимании [12].

В [13] для анализа поведения транзакций биткоина используются графовые НС, в которые добавляются линейные слои. Предложено использовать модифицированную версию графовой свёрточной сети, которая способна работать с ориентированными графами.

В [14] для выявления отличительных черт транзакций программ-вымогателей используются алгоритмы ИИ, такие как НС и оптимизируемые деревья решений.

В [15] собран набор данных и построен граф транзакций, к которому применён алгоритм глубокого обучения, основанный на графовой свёрточной сети. В [16] используется метод, основанный на графовой свёрточной сети [17], позволяющий осуществить поиск схожих узлов на одном уровне вложенности с получением векторов, в которые входят такие узлы, а с ними можно обращаться как с отдельным признаком данных. Затем эти признаки вместе с полученными векторами классифицируются с помощью алгоритма *RF*.

Для нарушения анонимности транзакций биткоина в целях обнаружения нелегальных транзакций в [18] используется НС обратного распространения ошибки.

Работа [19] включает поведенческий анализ данных о криптовалюте, т.е. обнаружение подозрительных адресов, основываясь на структуре транзакций и некоторых их признаках — поиск адресов, совпадающих с паттерном. Примером такого признака может служить большое количество входящих и исходящих адресов. Для повышения точности предполагается использование визуализации при устранении ложноположительных результатов.

В большинстве работ визуализация используется в качестве вспомогательного средства для иллюстрации применяемых методов (например, графовых НС), а не как инструмент обнаружения аномалий [20, 21]. В [22] данные о транзакциях преобразуются в графовую структуру, после чего возможен визуальный анализ полученных графов с целью выявления подозрительных паттернов. Так, мошеннические кошельки, как правило, характеризуются множеством входящих транзакций и одной выходящей, тогда как для легитимных кошельков типично наличие нескольких выходных транзакций.

В [23] применены методы визуализации для анализа сетей транзакций. В [24] графы транзакций визуализируются, а затем полученные изображения анализируются различными алгоритмами. В работах [25–28] используются различные методы визуализации к выявлению мошенничества в сфере мобильных денежных переводов.

Анализ показывает, что в большинстве работ применяется граф транзакций. Однако использование графовой модели может быть неэффективно в системах, функционирующих в режиме реального времени. В связи с этим актуальной задачей является создание подхода, способного прогнозировать легитимность транзакции на основе данных об одной отдельной операции, без необходимости формирования полной графовой структуры.

3 Предлагаемый метод

Ключевыми характеристиками криптовалютного обмена являются *децентрализованность* и *псевдонимность*. Поскольку одним из принципов блокчейна является его прозрачность, все транзакции доступны для просмотра другими пользователями, что исключает полную анонимность. Тем не менее, идентификаторы не содержат персональных данных, что позволяет говорить именно о псевдонимности транзакций.

Благодаря псевдонимности операций с биткоином, транзакции сложно отследить. Этим пользуются злоумышленники и используют криптовалюту для реализации различных противозаконных схем. Поскольку данные блоков в блокчейне не связаны с реальными юридическими данными людей, транзакция сама по себе может ничего не говорить об уровне легитимности действия пользователя. Однако цепочки транзакций могут свидетельствовать о применении мошеннических схем. Например, пользователь может использовать цепочку вводов и выводов криптовалюты для перевода средств на счёт запрещённой организации, скрыв при этом источник и назначение средств. Схожие схемы применяются при легализации доходов, полученных преступным путём.

При наличии знаний о типичных паттернах криптовалютных транзакций становится возможным с определённой вероятностью предположить назначение конкретной транзакции. Например, в случае оплаты множество входящих транзакций может консолидироваться в два выходных адреса, в то время как при операциях на криптовалютных биржах один адрес может инициировать одновременную отправку большого числа транзакций с целью снижения затрат на комиссии[6]. Для выявления нелегальных цепочек в большинстве исследований используется граф транзакций. Однако такой подход затрудняет обнаружение аномальных транзакций в режиме реального времени и не позволяет своевременно реагировать на инциденты. Возможным компромиссным решением может стать обогащение информации о транзакциях дополнительными графовыми признаками.

3.1 Описание набора данных

Выявление интервала и количества транзакций (отношение 5 см. рисунок 1) происходит на основе набора данных о транзакциях. Набор данных *BABD (Bitcoin Address Behavior Dataset,* набор данных о поведении адресов сети биткоин*)*, использованный в данной работе, загружен с сайта *Kaggle* [7]. Размеченные данные этого набора можно разделить на две условные группы [1]: данные о легитимных адресах (*Cyber-Security Service – Сервисы кибербезопасности; Centralized Exchange* – централизованный обмен*; P2P Financial Infrastructure Service* – сервис финансовой инфраструктуры сети *Peer-to-peer, где Peer-to-peer* – децентрализованная сеть, в которой все участники равноправны*; P2P Financial Service* – финансовый сервис *P2P; Individual Wallet* – индивидуальные кошельки) и нелегитимных (*Blackmail* – шантаж*; Darknet Market* – рынок сети *Darknet; Gambling* – азартные игры*; Government Criminal Blacklist* – государственный чёрный список*; Money Laundering* – отмы-

[6] Типы криптовалютных адресов и их транзакций в сети биткоин. 22.02.2023. https://shard.ru/article/types_of_cryptocurrency_addresses_and_transactions_in_the_bitcoin_network.
[7] *Kaggle*. Bitcoin Address Behavior Dataset (BABD-13). https://www.kaggle.com/datasets/lemonx/babd13.

вание денег; *Ponzi Scheme* – финансовые пирамиды). Майнеры (*Mining Pool*) можно отнести к условно легитимным, а криптовалютные тумблеры (*Tumbler*) к условно подозрительным, так как последние, повышая анонимность транзакций, могут использоваться для реализации нелегальных схем.

Предобработка данных заключалась в следующих преобразованиях.

Были удалены поля данных, которые не являются признаками категории транзакции, а именно *account* и *SW*. Поле *account* позволяет отличить один кошелёк от другого и строить графы транзакций. Поле *SW* отражает тип данных – имеет слабое (*weak address, WA*) или сильное (*strong address, SA*) подтверждение предоставленной авторами набора данных разметки. Удалены дубликаты экземпляров данных и убраны классы транзакций, которые содержат меньше 500 экземпляров данных – *P2P Financial Infrastructure Service, Government Criminal Blacklist, Money Laundering* и *Ponzi Scheme*. Собранный набор данных содержит девять классов транзакций, для которых в совокупности предоставлено 502416 экземпляров данных. Количество и распределение экземпляров набора данных по различным классам транзакций до и после предобработки представлено в таблице 1.

Таблица 1 – Распределение экземпляров данных по классам транзакций

Класс транзакции	До предварительной обработки		После предварительной обработки	
	Количество	Процент	Количество	Процент
Blackmail	8686	1.5953	8148	1.6218
Cyber-Security Service	91617	16.8271	90995	18.1115
Darknet Market	13861	2.5458	13857	2.7581
Centralized Exchange	300000	55.1003	260039	51.7577
P2P Financial Infrastructure Service	180	0.0331	0	0.0000
P2P Financial Service	9309	1.7098	9284	1.8479
Gambling	105257	19.3323	104605	20.8204
Government Criminal Blacklist	27	0.0050	0	0.0000
Money Laundering	16	0.0029	0	0.0000
Ponzi Scheme	15	0.0028	0	0.0000
Mining Pool	1580	0.2902	1580	0.3145
Tumbler	12412	2.2797	12406	2.4693
Individual Wallet	1502	0.2759	1502	0.2990
Всего	544462	1.0000	502416	1.0000

Анализируемый набор данных является несбалансированным и в основном представлен тремя классами транзакций – *Centralized Exchange* (51.76%), *Gambling* (20.82%) и *Cyber-Security Service* (18.11%), которые представляют всего 90.69 % данных. Весь набор данных содержит 151 признак.

3.2 Описание подходов

Предлагаемый метод для обнаружения аномальных транзакций криптовалют основан на онтологической схеме связей, представленной в разделе 1, и включает два основных подхода, направленных на отбор моделей НС и их совместное применение для обнаружения.

Подход для отбора моделей в задаче классификации транзакций криптовалюты.

Шаг 1. Предобработка данных о криптографических транзакциях.

Как правило, данные о криптографических транзакциях представляют собой структуру типа ключ-значение. В основе данных лежат различные поля, связанные с отправителем и получателем транзакций. На основе всей совокупности транзакций возможно извлечение дополнительных признаков, связанных с различными статистическими или графовыми метриками каждой отдельной транзакции. Итоговый набор данных представляет собой набор численных признаков и метку, связанные с каждой из транзакций.

Шаг 2. Разделение данных на обучающую, валидационную и тестовую выборки.

Для проверки однородности данных используются методы кросс-валидации. Поэтому для оценки эффективности принято решение использовать 80% данных для кросс-валидации моделей, и 20% – для их тестирования. Данные делятся в соответствии с распределением по классам транзакций. Тестовые данные недоступны моделям в процессе обучения. При кросс-валидации данные делились на части: 60% от изначального набора данных использовались для обучения, а 20% для валидации. Т.е. итоговое соотношение данных: 60/20/20.

Шаг 3. Использование обучающей и валидационной выборок для оптимизации гиперпараметров моделей.

Оптимизация гиперпараметров моделей происходит на 80% данных с помощью кросс-валидации. Каждая итерация данного процесса ставится на моделях с фиксированными гиперпараметрами. Результатом каждого эксперимента являются усреднённые результаты метрик эффективности и их среднеквадратичное отклонение. По результатам проверки различных комбинаций гиперпараметров выбираются их лучшие комбинации для каждой модели.

Шаг 4. Проверка оптимизированных моделей на тестовой выборке.

Лучшие модели по результатам кросс-валидации проверяются на тестовых данных, составляющих 20% от начального набора данных. Именно метрики эффективности, полученные на тестовых данных, являются итоговой оценкой производительности моделей.

Шаг 5. Отбор и выгрузка лучших моделей, их передача в систему обнаружения аномальных транзакций криптовалюты.

Лучшие модели сохраняются для дальнейшего использования при обнаружении подозрительных транзакций. Формат сохранения моделей зависит от итоговой реализации системы обнаружения, однако для предотвращения возможных атак предпочтительнее использовать защищённые форматы данных.

Подход для обнаружения аномальных транзакций криптовалюты.

Шаг 1. Предобработка данных.

Обученные модели предоставляют ответ на входные данные определённого типа. Они ожидают получить некоторый вектор, содержащий признаки транзакции в определённом порядке со значениями заданного типа (отношение 5 на рисунке 1). Важным является преобразование с помощью специальных адаптеров данных о транзакциях в формат, понятный моделям.

Шаг 2. Получение ответов от моделей.

Модели, получив на вход данные с шага 1, предоставляют ответ, содержащий вероятность соответствия полученной транзакции одному из известных модели классов – т.е. определяют, есть ли аномалия (отношение 6 на рисунке 1). В соответствии с определёнными в системе обнаружения пороговыми значениями определяется итоговый класс транзакции (отношение 7 на рисунке 1). Если транзакция является аномальной, информация об этом передается на шаг 3.

Шаг 3. Оценка обнаруженных транзакций оператором.

Оператор получает информацию обо всех транзакциях, признанных аномальными. Если оператор согласен с выводами модели, формируется задача для расследования инцидента (отношение 8 на рисунке 1). Мнение оператора используется для разметки и расширения наборов данных. Оценка оператора используется для анализа производительности моделей.

Шаг 4. Расширение обучающих данных, дообучение или переобучение моделей.

Этот шаг направлен на мониторинг производительности моделей с последующим их дообучения или переобучения. Данные процессы могут быть связаны со значительным расширением обучающего набора данных и значительным падением эффективности обнаружения аномальных транзакций.

Для оценки качества процессов выбраны следующие метрики: аккуратность, полнота, точность и *F*-мера [29].

4 Эксперименты

В рамках экспериментов[8] проанализированы следующие модели:

- *CNN* (*Convolutional Neural Network*) – свёрточная НС;
- *DNN* (*Deep Neural Network*) – глубокая НС;
- *GRU* (*Gated Recurrent Units*) – управляемые рекуррентные блоки.

[8] Все эксперименты выполнялись на компьютере с *Windows* 11 *Pro* 23H2. Для экспериментальной проверки моделей написан скрипт на языке *Python* в среде разработки *PyCharm* 2024.1. В качестве интерпретатора использовался *Python* 3.12.

- *CNN-DNN* – НС, представляющая собой комбинацию *CNN* и слоёв *DNN*;
- *CNN-GRU* – НС, представляющая собой комбинацию *CNN* и слоёв *GRU*.

Выбор моделей *CNN, DNN* и *GRU* обусловлен тем, что *CNN* эффективно выявляет пространственные зависимости, *DNN* хорошо обобщает сложные структуры данных, а *GRU* эффективно работает с последовательными временными данными. Гиперпараметры НС моделей и их значения представлены в таблице 2.

Таблица 2 – Гиперпараметры нейросетевых моделей

Модель	Гиперпараметры и их диапазоны
CNN	dropout: [0.00; 0.30], шаг 0.05; blocks: [1; 6], шаг 1; units: [64; 512], шаг 64
DNN	dropout: [0.00; 0.30], шаг 0.05; blocks: [1; 6], шаг 1; units: [64; 512], шаг 64
GRU	dropout: [0.00; 0.30], шаг 0.05; blocks: [1; 6], шаг 1; units: [64; 512], шаг 64
CNN-DNN	dropout: [0.00; 0.30], шаг 0.05; cnn_blocks: [1; 3], шаг 1; units: [64; 512], шаг 64; dnn_blocks: [1; 3], шаг 1; units: [64; 512], шаг 64
CNN-GRU	dropout: [0.00; 0.30], шаг 0.05; cnn_blocks: [1; 3], шаг 1; units: [64; 512], шаг 64; gru_blocks: [1; 3], шаг 1; units: [64; 512], шаг 64

Для анализа обобщающей способности моделей данные разделены в соотношении 60/20/20 на обучающую, валидационную и тестовую выборки. Оптимизация гиперпараметров моделей осуществлялась на обучающей и валидационной выборках. Данные из тестовой выборки моделям не предоставлялись. Эта выборка использовалась только для проверки эффективности моделей с наилучшими параметрами. Лучшие результаты работы НС моделей и значения гиперпараметров, при которых они были получены, представлены в таблице 3.

Таблица 3 – Результаты работы нейросетевых моделей в задаче классификации транзакций

Модель	Гиперпараметры	Процесс	Аккуратность	Точность	Полнота	F-мера
CNN	dropout: 0.05, blocks: 1, units: 384	Обучение	0.8235	0.8817	0.7355	0.8020
		Тест	0.8224	0.8821	0.7344	0.8015
DNN	dropout: 0.05, blocks: 6, units: 256	Обучение	0.8860	0.8788	0.8999	0.8869
		Тест	0.8894	0.9026	0.8772	0.8898
GRU	dropout: 0.05, blocks: 5, units: 512	Обучение	0.9361	0.9461	0.9288	0.9374
		Тест	**0.9352**	**0.9456**	**0.9280**	**0.9367**
CNN-DNN	dropout: 0.10, cnn_blocks: 1, dnn_blocks: 3, cnn_units: 128, dnn_units: 256	Обучение	0.8745	0.8988	0.8475	0.8724
		Тест	0.8758	0.9009	0.8502	0.8749
CNN-GRU	dropout: 0.10, cnn_blocks: 1, gru_blocks: 3, cnn_units: 256, gru_units: 256	Обучение	0.8780	0.9051	0.8526	0.8780
		Тест	0.8791	0.9051	0.8536	0.8786

В задаче классификации транзакций лучшие результаты показала модель *GRU*. *GRU* представляет собой тип рекуррентной НС с относительно простой структурой, которая позволяет эффективно решать проблему исчезающего градиента за счёт улучшения сходимости и производительности. Наиболее близкие к *GRU* результаты у *DNN*, самые слабые – у *CNN*. Комбинирование моделей показало средние результаты, что позволяет сделать вывод о преимуществе более простых моделей в решаемой задаче.

Для сравнения результатов НС с результатами методов ансамблевого МО, представленных в оригинальной работе по составлению набора данных *BABD*, были проанализированы следующие методы: *RF*; *CB* (*CatBoost*) – градиентный бустинг в деревьях решений; *XGB* (*Extreme Gradient Boosting*) – оптимизированный распределённый градиентный бустинг.

Гиперпараметры моделей и их значения представлены в таблице 4. Условия экспериментов для ансамблевых методов были аналогичны тем, что применялись для нейросетевых.

Лучшие результаты работы ансамблевых моделей и значения гиперпараметров, при которых они были получены, представлены в таблице 5.

Таблица 4 – Гиперпараметры ансамблевых моделей

Модель	Гиперпараметры и их диапазоны
RF	n_estimators: [100, 500], шаг 100; criterion: [gini, entropy, log_loss]; max_features: [sqrt, log2]
CB	iterations: [1000, 3000], шаг 500; learning_rate: [0.001, 0.03, 0.1]; grow_policy: [SymmetricTree, Depthwise, Lossguide]
XGB	n_estimators: [100, 500], шаг 100; learning_rate: [0.1, 0.01, 0.001]; booster: [gbtree, gblinear, dart]

Таблица 5 – Результаты работы ансамблевых моделей в задаче классификации транзакций

Модель	Гиперпараметры	Процесс	Аккуратность	Точность	Полнота	F-мера
RF	criterion: gini, max_features: sqrt, n_estimators: 400	Обучение	0.9535	0.9497	0.9535	0.9510
		Тест	0.9570	0.9570	0.9570	0.9570
CB	grow_policy: SymmetricTree, iterations: 3000, learning_rate: 0.1	Обучение	0.9630	0.9602	0.9630	0.9614
		Тест	0.9643	0.9643	0.9643	0.9643
XGB	booster: gbtree, learning_rate: 0.1, n_estimators: 500	Обучение	0.9659	0.9632	0.9659	0.9644
		Тест	**0.9673**	**0.9673**	**0.9673**	**0.9673**

Лучшие результаты показала модель *XGB*. Особенностью данной модели является обеспечение параллельного бустинга деревьев. Близкие результаты у модели *CB*, худшие – у *RF*. Результаты *RF* выше результатов *GRU*, что позволяет сделать вывод о преимуществе ансамблевых методов над нейросетевыми в проведённом эксперименте.

5 Обсуждение результатов

Метрики эффективности моделей *XGB* и *GRU* представлены в таблице 6.

Таблица 6 – Эффективность обнаружения отдельных классов транзакций

Класс транзакции	Модель	Точность	Полнота	F-мера	Количество экземпляров данных
Blackmail	*XGB*	0.61	0.53	0.57	1 641
	GRU	0.53	0.48	0.51	
Cyber-Security Service	*XGB*	0.94	0.96	0.95	18 130
	GRU	0.88	0.90	0.89	
Darknet Market	*XGB*	0.98	0.99	0.99	2 701
	GRU	0.89	0.81	0.85	
Centralized Exchange	*XGB*	0.98	0.99	0.99	51 848
	GRU	0.96	0.97	0.97	
P2P Financial Service	*XGB*	0.99	0.93	0.96	1 870
	GRU	0.93	0.88	0.90	
Gambling	*XGB*	0.98	0.98	0.98	21 105
	GRU	0.94	0.95	0.94	
Individual Wallet	*XGB*	1.00	0.96	0.98	327
	GRU	0.78	0.80	0.79	
Tumbler	*XGB*	0.89	0.85	0.87	2 535
	GRU	0.90	0.75	0.82	
Mining Pool	*XGB*	0.08	0.02	0.03	327
	GRU	0.67	0.01	0.01	
Среднее (макро)	*XGB*	0.83	0.80	0.81	100 484
	GRU	0.83	0.73	0.74	
Среднее (взвешенное)	*XGB*	0.96	0.97	0.97	
	GRU	0.94	0.93	0.93	

Из таблицы 6 видно, что добиться высокой эффективности обнаружения для всех представленных классов не удалось. Это явно проявляется на таких типах транзакций, как *Mining Pool* и *Blackmail*, а также *Tumbler*. Точность классификации *Mining Pool* у *GRU* значительно выше, чем у *XGB*, однако данное преимущество нивелируется низкой полнотой обнаружения. Для улучшения результатов в этих классах необходимо рассмотреть методы балансировки данных или разработать специализированные архитектуры. Эффективность *GRU* схожа с *XGB*, а результаты *XGB* по *F*-мере выше для всех классов транзакций.

В [30] применены статистические методы (критерий согласия Колмогорова, критерий Андерсона-Дарлинга, критерий Крамера-Мизеса-Смирнова [31]) при сравнении двух наборов данных: только нелегальные транзакции и легальные вместе с нелегальными. Таблица 7 содержит результаты исследований с использованием набора данных *BABD*.

Таблица 7 – Результаты исследований, использующих набор данных *BABD*

Метод	Аккуратность	Точность	Полнота	F-мера
критерий согласия Колмогорова	0.85	1.00	0.85	0.92
критерий Андерсона-Дарлинга	0.79	1.00	0.79	0.88
критерий Крамера-Мизеса-Смирнова	0.80	1.00	0.80	0.89
Наш нейросетевой метод (*GRU*)	0.94	0.95	0.93	0.94
Наш ансамблевый метод (*XGB*)	0.97	0.97	0.97	0.97

Результаты проведённых экспериментов показали, что предложенные в данной статье решения показывают более сбалансированные результаты по полноте и точности, чем рассмотренные аналоги. Результаты работы нейросетевых методов уступают ансамблевым, но их более высокая обобщающая способность может дать им преимущество.

Заключение

Предложен подход к обнаружению аномальных транзакций в сети криптовалют с применением ИИ. Экспериментально проведено сравнение трёх подходов к анализу данных о криптовалютных транзакциях на примере биткоина: классификаторы, основанные на алгоритмах МО; статистические алгоритмы; предложенный подход, основанный на НС. Получены следующие результаты для модели *GRU* НС: аккуратность 0.94, точность 0.95, полнота 0.93 и *F*-мера 0.94, что свидетельствует о высокой эффективности предложенной модели по сравнению с альтернативными решениями, основанными на традиционных алгоритмах МО.

Выявлены ограничения, связанные с классификацией отдельных типов транзакций (*Blackmail, Tumbler*), что указывает на необходимость дальнейших исследований, направленных на повышение точности и полноты обнаружения нелегальных операций.

Предложенный метод может быть использован в системах цифровой криминалистики, в частности для мониторинга подозрительной активности в режиме реального времени.

Список источников

[1] ***Xiang Y., Lei Y., Bao D., Li T., Yang Q., Liu W., Ren W., Choo K.K.R.*** BABD: A Bitcoin Address Behavior Dataset for Pattern Analysis. *IEEE Transactions on Information Forensics and Security*, 2024. Vol.19. P.2171-2185. DOI: 10.1109/TIFS.2023.3347894.

[2] ***Kovalchuk O., Shevchuk R., Banakh S.*** Cryptocurrency Crime Risks Modeling: Environment, E-Commerce, and Cybersecurity Issue. *IEEE Access*, 2024. Vol.12. P.50673- 50688. DOI: 10.1109/ACCESS.2024.3386428.

[3] ***Nayyer N., Javaid N., Akbar M., Aldegheishem A., Alrajeh N., Jamil M.*** A new framework for fraud detection in bitcoin transactions through ensemble stacking model in smart cities. *IEEE Access*. 2023. Vol.11. P.90916- 90938. DOI: 10.1109/ACCESS.2023.3308298.

[4] ***Nerurkar P., Busnel Y., Ludinard R., Shah K., Bhirud S., Patel D.*** Detecting illicit entities in bitcoin using supervised learning of ensemble decision trees // In: Proceedings of the 10th international conference on information communication and management (Paris, France, 2020). 2020. P.25-30. DOI: 10.1145/3418981.3418984.

[5] ***Gomez G., Moreno-Sanchez P., Caballero J.*** Watch your back: identifying cybercrime financial relationships in bitcoin through back-and-forth exploration // In: Proceedings of the 2022 ACM SIGSAC conference on computer and communications security (Los Angeles, CA, USA, 2022). P.1291-1305. DOI: 10.1145/3548606.3560587.

[6] ***Lorenz J., Silva M.I., Aparício D., Ascensão J.T., Bizarro P.*** Machine learning methods to detect money laundering in the bitcoin blockchain in the presence of label scarcity // In: Proceedings of the first ACM international conference on AI in finance (New York, NY, USA, 2020). P.1-8. DOI: 10.1145/3383455.3422549.

[7] ***Alsaif S.A.*** Machine Learning- Based Ransomware Classification of Bitcoin Transactions. *Applied Computational Intelligence and Soft Computing*. 2023. Vol.2023. №1. P.6274260. DOI: 10.1155/2023/6274260.

[8] ***Chen B., Wei F., Gu C.*** Bitcoin theft detection based on supervised machine learning algorithms. *Security and Communication Networks*. 2021. Vol.2021. №1. P.6643763. DOI: 10.1155/2021/6643763.

[9] ***Ashfaq T., Khalid R., Yahaya A.S., Aslam S., Azar A.T., Alsafari S., Hameed I.A.*** A machine learning and blockchain based efficient fraud detection mechanism. *Sensors*. 2022. Vol.22. №19. P.7162. DOI: 10.3390/s22197162.

[10] ***Jang J., Lee H.N.*** Profitable double-spending attacks. *Applied Sciences*. 2020. Vol.10. №23. P.8477. DOI: 10.3390/app10238477.

[11] ***Handaya W.B.T., Yusoff M.N., Jantan A.*** Machine learning approach for detection of fileless cryptocurrency mining malware. *Journal of Physics: Conference Series*. IOP Publishing, 2020. Vol.1450. №1. P.012075. DOI: 10.1088/1742-6596/1450/1/012075.

[12] ***Pocher N., Zichichi M., Merizzi F., Shafiq M.Z., Ferretti S.*** Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics. *Electronic Markets*. 2023. Vol.33. №1. P.37. DOI: 10.1007/s12525-023-00654-3.

[13] ***Alarab I., Prakoonwit S., Nacer M.I.*** Competence of graph convolutional networks for anti-money laundering in bitcoin blockchain // In: Proceedings of the 2020 5th international conference on machine learning technologies (New York, NY, USA, 2020). P.23-27. DOI: 10.1145/3409073.

[14] ***Al-Haija Q.A., Alsulami A.A.*** High performance classification model to identify ransomware payments for heterogeneous bitcoin networks. *Electronics*. 2021. Vol.10. №17. P.2113. DOI: 10.3390/electronics10172113.

[15] ***Nerurkar P.*** Illegal activity detection on bitcoin transaction using deep learning. *Soft Computing*. 2023. Vol.27. №9. P.5503-5520. DOI: 10.21203/rs.3.rs-1454891/v1.

[16] ***Lo W.W., Kulatilleke G.K., Sarhan M., Layeghy S., Portmann M.*** Inspection-L: self-supervised GNN node embeddings for money laundering detection in bitcoin. *Applied Intelligence*. 2023. Vol.53. №16. P.19406-19417. https://arxiv.org/pdf/2203.10465.

[17] ***Velickovic P., Fedus W., Hamilton W.L., Lio P., Bengio Y., Hjelm R.D.*** Deep graph infomax // In: Proceedings of International Conference on Learning Representations (New Orleans, Louisiana, USA, 2019), 2019. P.1-17. https://openreview.net/pdf?id=rklz9iAcKQ.

[18] ***Saxena R., Arora D., Nagar V.*** Integration of back-propagation neural network to classify of cybercriminal entities in blockchain // In: Proceedings of Trends in Electronics and Health Informatics: TEHI 2021. Singapore: Springer Nature Singapore, 2022. P.523-532. DOI: 10.1007/978-981-16-8826-3_45.

[19] ***Wu Y., Tao F., Liu L., Gu J., Panneerselvam J., Zhu R., Shahzad M.N.*** A bitcoin transaction network analytic method for future blockchain forensic investigation. *IEEE Transactions on Network Science and Engineering*. 2020. Vol.8. №2. P.1230-1241. DOI: 10.1109/TNSE.2020.2970113.

[20] ***Wu J., Liu J., Chen W., Huang H., Zheng Z., Zhang Y.*** Detecting mixing services via mining bitcoin transaction network with hybrid motifs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 2021. Vol.52. №4. P.2237-2249. https://arxiv.org/pdf/2001.05233.

[21] ***Nicholls J., Kuppa A., Le-Khac N.A.*** FraudLens: Graph Structural Learning for Bitcoin Illicit Activity Identification // In: Proceedings of the 39th Annual Computer Security Applications Conference (Austin, TX, USA, 2023). 2023. P.324-336. DOI: 10.1145/3627106.3627200.

[22] ***Tharani J. S., Charles E.Y.A., Hóu Z., Palaniswami M., Muthukkumarasamy V.*** Graph based visualisation techniques for analysis of blockchain transactions // In: 2021 IEEE 46th Conference on Local Computer Networks (LCN) (Edmonton, AB, Canada, 2021). IEEE, 2021. P.427-430. DOI: 10.1109/LCN52139.2021.9524878.

[23] ***Xiong H., Yiu S. M., Lam, K.Y.*** Bitanalysis: A visualization system for bitcoin wallet investigation // IEEE Transactions on Big Data. 2022. Vol.9. №2. P.621-636. DOI: 10.1109/TBDATA.2022.3188660.

[24] ***Wu F., Wei Y., Luo X.*** Abnormal Trading Visualized Detection on Bitcoin Transaction Based on Semi-Supervised Machine Learning and Graph Database. 2024. https://ssrn.com/abstract=4769024. DOI: 10.2139/ssrn.4769024.

[25] ***Novikova E., Kotenko I., Fedotov E.*** Interactive Multiview Visualization for Fraud Detection in Mobile Money Transfer Services. *International Journal of Mobile Computing and Multimedia Communications*, Vol.6, No.4, 2015. P.72-97. DOI: 10.4018/IJMCMC.2014100105.

[26] ***Novikova E., Kotenko I.*** Visualization-Driven Approach to Fraud Detection in the Mobile Money Transfer Services. *Algorithms, Methods and Applications in Mobile Computing and Communications*. IGI Global. 2019. P.205-236. DOI: 10.4018/978-1-5225-5693-0.ch009.

[27] ***Novikova E., Bestuzhev M., Kotenko I.*** Anomaly Detection in the HVAC System Operation by a RadViz Based Visualization-Driven Approach // Katsikas S. et al. (eds) Computer Security. ESORICS 2019 International Workshops, CyberICPS, SECPRE, SPOSE, and ADIoT, Luxembourg, September 26–27, 2019 Revised Selected Papers / Lecture Notes in Computer Sciences, Vol.11980, 2020. Springer, Cham. P.402-418. DOI: 10.1007/978-3-030-42048-2_26.

[28] ***Novikova E., Kotenko I., Murenin I.*** The Visual Analytics Approach for Analyzing Trajectories of Critical Infrastructure Employers. *Energies* (MDPI) 2020, Vol.13, №15, P.3936. DOI: 10.3390/en13153936.

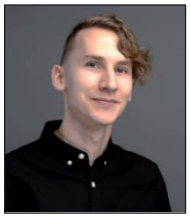
[29] ***Левшун Д.А., Левшун Д.С., Котенко И.В.*** Обнаружение и объяснение аномалий в индустриальных системах Интернета вещей на основе автокодировщика. *Онтология проектирования*. 2025. Т.15, №1(55). С.96-113. DOI:10.18287/2223-9537-2025-15-1-96-113.

[30] ***Maheshwari R., Sriram Praveen V.A., Shobha G., Shetty J., Chala A., Watanuki H.*** Illicit activity detection in bitcoin transactions using timeseries analysis. *International Journal of Advanced Computer Science and Applications*. 2023. Vol.14. №3. P.13-18. DOI: 10.14569/IJACSA.2023.0140302.

[31] ***Samsudeen F., Perera H.*** Behavioral analysis of bitcoin users on illegal transactions. *Advances in Science Technology and Engineering Systems Journal*. 2019. Vol.4. №2. P.402-412. DOI: 10.25046/aj040250.

Сведения об авторах

Котенко Игорь Витальевич, 1961 г. рождения. С отличием окончил Военный инженерный Краснознамённый институт им. А.Ф. Можайского в 1983 г. и Военную академию связи в 1987 г., д.т.н. (1999), профессор (2001), заслуженный деятель науки Российской Федерации (2023). Главный научный сотрудник и руководитель Лаборатории проблем компьютерной безопасности СПб ФИЦ РАН, профессор Университета ИТМО, УрФУ, Харбинского политехнического университета (КНР) и Хэйлунцзянского университета (КНР), заместитель директора Института информационной безопасности Университета Иннополис. В списке научных трудов более 800 работ в области защиты информации и искусственного интеллекта, включая 25 монографий, и более 100 патентов на изобретения и зарегистрированных программ. ORCID: 0000-0001-6859-7120; Author ID (РИНЦ): 110102; Author ID (Scopus): 15925268000. *ivkote@comsec.spb.ru.* ✉.



Левшун Дмитрий Сергеевич, 1993 г. рождения. Выпускник СПбГЭТУ «ЛЭТИ» (2017), к.т.н. (2021). Старший научный сотрудник лаборатории проблем компьютерной безопасности СПб ФИЦ РАН. Доцент Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПбГУТ) и Европейского университета в Санкт-Петербурге. В списке научных трудов более 100 работ в области информационной безопасности, проектирования защищённых систем, Интернета вещей, искусственного интеллекта. ORCID: 0000-0003-1898-6624; Author ID (РИНЦ): 840344; Author ID (Scopus): 57189306576; Researcher ID (WoS): C-1566-2018. *levshun@comsec.spb.ru.*

Жернова Ксения Николаевна, 1994 г. рождения. Окончила СПбГУТ в 2018 г., к.т.н. (2022). Старший научный сотрудник Международного центра цифровой криминалистики СПб ФИЦ РАН. Доцент Европейского университета в Санкт-Петербурге. В списке научных трудов более 50 работ в области информационной безопасности, визуального анализа данных, проблем искусственного интеллекта. ORCID: 0000-0003-0161-3645; Author ID (РИНЦ): 1035947; Author ID (Scopus): 57216946658. *zhernova@comsec.spb.ru.*

Чечулин Андрей Алексеевич, 1983 г. рождения. Окончил Санкт-Петербургский государственный политехнический университет им. Петра Великого в 2005 г., к.т.н. (2013), доцент (2018). Руководитель Международного центра цифровой криминалистики СПб ФИЦ РАН. Доцент СПбГУТ и Университета ИТМО. Автор более 300 научных публикаций в области сетевой безопасности, обнаружения вторжений, цифровой криминалистики, анализа социальных сетей и визуализации данных. ORCID: 0000-0001-7056-6972; Author ID (РИНЦ): 608765; Author ID (Scopus): 55248184200; Researcher ID (WoS): K-7971-2012. *chechulin@comsec.spb.ru*.

Поступила в редакцию 12.11.2024, после рецензирования 17.04.2025. Принята к публикации 26.05.2025.

Scientific article DOI: 10.18287/2223-9537-2025-15-3-334-350

Detection of anomalous cryptocurrency transactions using neural networks and ontologies

© 2025, I.V. Kotenko ✉, D.S. Levshun, K.N. Zhernova, A.A. Chechulin

Saint-Petersburg Federal Research Scientific Center of the Russian Academy of Science, Saint-Petersburg, Russia

Abstract

The article explores an effective approach to detecting anomalies in cryptocurrency transactions using neural network models, including convolutional, deep, and gated recurrent units (GRUs), and compares their performance with other existing methods for identifying illicit transactions in cryptocurrency networks. A research of relevant studies is conducted in the fields of transaction data analysis in the cryptocurrency network, data visualization for transaction analysis, and the use of computer vision techniques for detecting anomalous behavior. The subject area of the study is defined. The problem of detecting anomalies in cryptocurrency transactions is based on the fact that these transactions are pseudonymous, i.e. there are no direct indications of the identity of the sender and recipient. The relevance and contribution of this work lie in the development of a method capable of identifying anomalous transactions with high accuracy in near real-time. Experimental studies were conducted using a dataset of cryptocurrency transactions, applying both neural and non-neural classifiers. The results are compared against existing approaches in the field. The experiments demonstrated that gated recurrent units outperformed other neural models in this task, achieving an accuracy of 0.94, precision of 0.95, recall of 0.93, and F1-score of 0.94, indicating the high effectiveness of the proposed model. Nonetheless, this approach showed slightly lower performance compared to traditional machine learning algorithms, such as optimized distributed gradient boosting. The novelty of the proposed approach lies in its use of statistical characteristics derived from the transaction graph, combined with deep learning and gradient boosting techniques. The approach can be applied in the development of software tools for detecting illicit cryptocurrency transactions within information security systems and digital forensics.

Keywords: *anomaly detection, computer security, machine learning, neural networks, visual data analysis, cryptocurrency, digital forensics.*

For citation: *Kotenko IV, Levshun DS, Zhernova KN, Chechulin AA.* Detection of anomalous cryptocurrency transactions using neural networks and ontologies [In Russian]. *Ontology of designing*. 2025; 15(3): 334-350. DOI: 10.18287/2223-9537-2025-15-3-334-350.

Financial Support: The study was supported by a joint grant from the Russian Science Foundation and the St. Petersburg Science Foundation No. 24-21-20058, https://rscf.ru/project/24-21-20058/.

Authors' contributions: *Kotenko I.V.* – problem statement, general guidance on the work and planning of experiments; *Levshun D.S.* – conducting the experiment and interpreting the results; *Zhernova K.N.* – search and description of the data set for the experiment, review of relevant works; *Chechulin A.A.* - development of the research concept, preparation of the data set.

Conflict of interest: The authors declare no conflict of interest.

List of figures and tables

References

[1] ***Xiang Y, Lei Y, Bao D, Li T, Yang Q, Liu W, Ren W, Choo KKR.*** BABD: A Bitcoin Address Behavior Dataset for Pattern Analysis. *IEEE Transactions on Information Forensics and Security*, 2024; 19: 2171-2185. DOI: 10.1109/TIFS.2023.3347894.

[2] ***Kovalchuk O, Shevchuk R, Banakh S.*** Cryptocurrency Crime Risks Modeling: Environment, E-Commerce, and Cybersecurity Issue. *IEEE Access*. 2024; 12: 50673- 50688. DOI: 10.1109/ACCESS.2024.3386428.

[3] ***Nayyer N, Javaid N, Akbar M, Aldegheishem A, Alrajeh N, Jamil M.*** A new framework for fraud detection in bitcoin transactions through ensemble stacking model in smart cities. *IEEE Access*. 2023; 11: 90916- 90938. DOI: 10.1109/ACCESS.2023.3308298.

[4] ***Nerurkar P, Busnel Y, Ludinard R, Shah K, Bhirud S, Patel D.*** Detecting illicit entities in bitcoin using supervised learning of ensemble decision trees. *Proceedings of the 10th international conference on information communication and management* (Paris, France, 2020). 2020: 25-30. DOI: 10.1145/3418981.3418984.

[5] ***Gomez G, Moreno-Sanchez P, Caballero J.*** Watch your back: identifying cybercrime financial relationships in bitcoin through back-and-forth exploration. *Proceedings of the 2022 ACM SIGSAC conference on computer and communications security* (Los Angeles, CA, USA, 2022). 2022: 1291-1305. DOI: 10.1145/3548606.3560587.

[6] ***Lorenz J, Silva MI, Aparício D, Ascensão JT, Bizarro P.*** Machine learning methods to detect money laundering in the bitcoin blockchain in the presence of label scarcity. *Proceedings of the first ACM international conference on AI in finance* (New York, NY, USA, 2020). 2020: 1-8. DOI: 10.1145/3383455.3422549.

[7] ***Alsaif SA.*** Machine Learning- Based Ransomware Classification of Bitcoin Transactions. *Applied Computational Intelligence and Soft Computing*. 2023; 2023(1): 6274260. DOI: 10.1155/2023/6274260.

[8] ***Chen B, Wei F, Gu C.*** Bitcoin theft detection based on supervised machine learning algorithms. *Security and Communication Networks*. 2021; 2021(1): 6643763. DOI: 10.1155/2021/6643763.

[9] ***Ashfaq T, Khalid R, Yahaya AS, Aslam S, Azar AT, Alsafari S, Hameed IA.*** A machine learning and blockchain based efficient fraud detection mechanism. *Sensors*. 2022; 22(19): 7162. DOI: 10.3390/s22197162.

[10] ***Jang J, Lee HN.*** Profitable double-spending attacks. *Applied Sciences*. 2020; 10(23): 8477. DOI: 10.3390/app10238477.

[11] ***Handaya WBT, Yusoff MN, Jantan A.*** Machine learning approach for detection of fileless cryptocurrency mining malware. *Journal of Physics: Conference Series*. IOP Publishing, 2020; 1450(1): 012075. DOI: 10.1088/1742-6596/1450/1/012075.

[12] ***Pocher N, Zichichi M, Merizzi F, Shafiq MZ, Ferretti S.*** Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics. *Electronic Markets*. 2023; 33(1): 37. DOI: 10.1007/s12525-023-00654-3.

[13] ***Alarab I, Prakoonwit S, Nacer MI.*** Competence of graph convolutional networks for anti-money laundering in bitcoin blockchain. *Proceedings of the 2020 5th international conference on machine learning technologies* (New York, NY, USA, 2020). 2020: 23-27. DOI: 10.1145/3409073.

[14] ***Al-Haija QA, Alsulami AA.*** High performance classification model to identify ransomware payments for heterogeneous bitcoin networks.*Electronics*. 2021; 10(17): 2113. DOI: 10.3390/electronics10172113.

[15] ***Nerurkar P.*** Illegal activity detection on bitcoin transaction using deep learning. *Soft Computing*. 2023; 27(9): 5503-5520. DOI: 10.21203/rs.3.rs-1454891/v1.

[16] ***Lo WW, Kulatilleke GK, Sarhan M, Layeghy S, Portmann M.*** Inspection-L: self-supervised GNN node embeddings for money laundering detection in bitcoin. *Applied Intelligence*. 2023; 53(16): 19406-19417. https://arxiv.org/pdf/2203.10465.

[17] ***Velickovic P, Fedus W, Hamilton WL, Lio P, Bengio Y, Hjelm RD.*** Deep graph infomax. *Proceedings of International Conference on Learning Representations* (New Orleans, Louisiana, USA, 2019), 2019: 1-17. https://openreview.net/pdf?id=rklz9iAcKQ.

[18] ***Saxena R, Arora D, Nagar V.*** Integration of back-propagation neural network to classify of cybercriminal entities in blockchain. *Proceedings of Trends in Electronics and Health Informatics*: TEHI 2021. Singapore: Springer Nature Singapore, 2022: 523-532. DOI: 10.1007/978-981-16-8826-3_45.

[19] ***Wu Y, Tao F, Liu L, Gu J, Panneerselvam J, Zhu R, Shahzad MN.*** A bitcoin transaction network analytic method for future blockchain forensic investigation. *IEEE Transactions on Network Science and Engineering*. 2020; 8(2): 1230-1241. DOI: 10.1109/TNSE.2020.2970113.

[20] ***Wu J, Liu J, Chen W, Huang H, Zheng Z, Zhang Y.*** Detecting mixing services via mining bitcoin transaction network with hybrid motifs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*. 2021; 52(4): 2237-2249. https://arxiv.org/pdf/2001.05233.

[21] ***Nicholls J, Kuppa A, Le-Khac NA.*** FraudLens: Graph Structural Learning for Bitcoin Illicit Activity Identification. *Proceedings of the 39th Annual Computer Security Applications Conference* (Austin, TX, USA, 2023). 2023: 324-336. DOI: 10.1145/3627106.3627200.

[22] ***Tharani JS, Charles EYA, Hóu Z, Palaniswami M, Muthukkumarasamy V.*** Graph based visualisation techniques for analysis of blockchain transactions. *IEEE 46th Conference on Local Computer Networks* (LCN). IEEE, 2021: 427-430. DOI: 10.1109/LCN52139.2021.9524878.

[23] ***Xiong H, Yiu SM, Lam KY.*** Bitanalysis: A visualization system for bitcoin wallet investigation. *IEEE Transactions on Big Data*. 2022; 9(2): 621-636. DOI: 10.1109/TBDATA.2022.3188660.

[24] ***Wu F, Wei Y, Luo X.*** Abnormal Trading Visualized Detection on Bitcoin Transaction Based on Semi-Supervised Machine Learning and Graph Database. 2024. https://ssrn.com/abstract=4769024. DOI: 10.2139/ssrn.4769024.

[25] ***Novikova E, Kotenko I, Fedotov E.*** Interactive Multiview Visualization for Fraud Detection in Mobile Money Transfer Services. *International Journal of Mobile Computing and Multimedia Commmunications*, 2015; 6(4): 72-97. DOI: 10.4018/IJMCMC.2014100105.

[26] ***Novikova E, Kotenko I.*** Visualization-Driven Approach to Fraud Detection in the Mobile Money Transfer Services. *Algorithms, Methods and Applications in Mobile Computing and Communications*. IGI Global. 2019: 205-236. DOI: 10.4018/978-1-5225-5693-0.ch009.

[27] ***Novikova E, Bestuzhev M, Kotenko I.*** Anomaly Detection in the HVAC System Operation by a RadViz Based Visualization-Driven Approach // Katsikas S. et al. (eds) Computer Security. *ESORICS 2019 International Workshops, CyberICPS, SECPRE, SPOSE, and ADIoT*, Luxembourg, September 26–27, 2019 Revised Selected Papers / Lecture Notes in Computer Sciences. Springer, Cham. 2020; 11980: 402-418. DOI: 10.1007/978-3-030-42048-2_26.

[28] ***Novikova E, Kotenko I, Murenin I.*** The Visual Analytics Approach for Analyzing Trajectories of Critical Infrastructure Employers. *Energies* (MDPI). 2020; 13(15): 3936. DOI: 10.3390/en13153936.

[29] ***Levshun D, Levshun D, Kotenko I.*** Detecting and explaining anomalies in industrial Internet of things systems using an autoencoder [In Russian]. *Ontology of Designing*, 2025; 1(55): 96-113. DOI:10.18287/2223-9537-2025-15-1-96-113.

[30] ***Maheshwari R, Sriram Praveen VA, Shobha G, Shetty J, Chala A, Watanuki H.*** Illicit activity detection in bitcoin transactions using timeseries analysis. *International Journal of Advanced Computer Science and Applications*. 2023; 14(3): 13-18. DOI: 10.14569/IJACSA.2023.0140302.

[31] ***Samsudeen F, Perera H.*** Behavioral analysis of bitcoin users on illegal transactions. *Advances in Science Technology and Engineering Systems Journal*. 2019; 4(2): 402-412. DOI: 10.25046/aj040250.

About the authors

Igor Vitalyevich Kotenko (b. 1961) graduated with honors from the St. Petersburg Academy of Space Engineering in 1983 and St. Petersburg Signal Academy in 1987, D. Sc. Eng. (1999), professor (2021), Honored Scientist of the Russian Federation (2023), Chief Researcher and Head of the Laboratory of Computer Security Problems at St. Petersburg Federal Research Center of the Russian Academy of Sciences, professor at ITMO University, Saint Petersburg State University of Telecommunications, Ural Federal University, Harbin Institute of Technology (China) and Heilongjiang University (China), Deputy Director of the Institute of Information Security at Innopolis University. He is a co-author of more than 800 publications in the field of information security and artificial intelligence, including 25 monographs, and more than 100 patents for inventions and registered programs. ORCID: 0000-0001-6859-7120; Author ID (RSCI): 110102; Author ID (Scopus): 15925268000. *ivkote@comsec.spb.ru.* ✉.

Dmitry Sergeevich Levshun (b 1993) graduated from the Saint Petersburg Electrotechnical University "LETI" in 2017, Ph.D. (2021). He is a senior researcher at the Laboratory of Computer Security Problems at the St. Petersburg Federal Research Center of the Russian Academy of Sciences. He is an associate professor at the Saint Petersburg State University of Telecommunications named after prof. M.A. Bonch-Bruevich and at the European University at St. Petersburg. His list of scientific publications includes more than 100 works in the field of information security, design of security systems, Internet of Things, and artificial intelligence. ORCID: 0000-0003-1898-6624; Author ID (RSCI): 840344; Author ID (Scopus): 57189306576; Researcher ID (WoS): C-1566-2018. *levshun@comsec.spb.ru.*

Ksenia Nikolayevna Zhernova (b. 1994) graduated from the Saint Petersburg State University of Telecommunications named after prof. M.A. Bonch-Bruevich (Saint-Petersburg, Russia) in 2018, Can. Sc. Eng. in Information Security (2022). Senior Researcher at the International Digital Forensics Center at the Saint Petersburg Federal Research Center of the Russian Academy of Sciences. Associate Professor at the European University at Saint-Petersburg. The list of scientific publications includes more than 50 works in the field of information security, visual data analysis, and artificial intelligence problems. ORCID: 0000-0003-0161-3645; Author ID (RSCI): 1035947; Author ID (Scopus): 57216946658. *zhernova@comsec.spb.ru*.

Andrey Alekseevich Chechulin (b. 1983), graduated from Peter the Great St. Petersburg Polytechnic University in 2005, holds a Can.Sc. Eng. in Information Security (2013), and Associate Professor (2018). Lead Researcher and Head

of the International Digital Forensics Center at the St. Petersburg Federal Research Center of the Russian Academy of Sciences. Associate Professor at St. Petersburg State University of Telecommunications and ITMO University. He is the author of over 300 scientific publications in the fields of network security, intrusion detection, digital forensics, social network analysis, and data visualization. ORCID: 0000-0001-7056-6972; Author ID (RSCI): 608765; Author ID (Scopus): 55248184200; Researcher ID (WoS): K-7971-2012. *chechulin@comsec.spb.ru.*

Received November 12, 2024. Revised April 17, 2025. Accepted May 26, 2025.