Моделирование блокчейнов для машинного обучения с целью выявления возможной отслеживаемости и определения сумм транзакций в сети Monero

Натан Борггрен (Nathan Borggren), Хён-Юн Ким (Hyoung-yoon Kim), Лихан Яо (Lihan Yao) и Гэри Коплик (Gary Koplik)
Geometric Data Analytics, Inc., Дарем, Северная Каролина, 27701

Аннотация

Monero — популярная криптовалюта, ориентированная на обеспечение приватности пользователей. Блокчейн Monero использует криптографические технологии, позволяющие скрыть суммы транзакций, а также протокол кольцевых конфиденциальных транзакций (Ring Confidential Transactions), скрывающий реальную транзакцию среди различного количества ложных транзакций. Нами был разработан ряд наборов моделей блокчейнов, по 10 и 50 агентов, которые мы полностью контролируем и ключами которых обладаем, чтобы протестировать эти возможности. Мы выводим свойства транзакции путём определения локальной структуры публично доступных блокчейнов и используем полученные при моделировании маркеры для машинного обучения. Результаты машинного обучения по полученным нами характеристикам смоделированных блокчейнов говорит о том, что предлагаемый метод может использоваться в качестве вспомогательного для идентификации отдельных лиц и групп пользователей, несмотря на то, что не позволяет успешно раскрыть скрытые суммы транзакций. Мы также применяем указанный метод к реальному блокчейну Monero с целью выявления транзакций криптовалютной биржи ShapeShift, допустившей утечку информации через свой API, который обеспечивал соответствующими метками как саму биржу, так и её пользователей.

1 Введение

Мы уже успешно применяли машинное обучение (ML) в прошлом, объединяя свойства, полученные напрямую из блокчейна, с метками, взятыми из источников вне блокчейна [1, 2]. Другие исследователи также использовали методы ML для деанонимизации бирж [3], идентификации вредоносного программного обеспечения и прочих объектов [5]. За исключением работы [2], основной целью анализа всегда был Bitcoin. Причина состояла в простоте получения меток, даже несмотря на сложность их верификации [6]. Однако, и другие криптовалюты уже достаточно хорошо развились и обладают достаточной ликвидностью для использования в схемах межвалютного смешивания [2, 7].

Несмотря на то, что многие криптовалюты по сути являются производными Bitcoin и появились в результате форка репозитория кода с последующими косметическими изменениями, некоторые монеты всё же реализовали различные базовые допуски, имеют в значительной степени обновлённые кодовые базы и используют новые криптографические методы, позволяющие обойти некоторые ограничения, связанные с обеспечением приватности, характерные для Bitcoin. Одной из таких монет является Monero [8]. Monero имеет некоторые структурные отличия от Bitcoin, которых оказалось вполне достаточно для того, чтобы сделать механизмы ML в стиле Bitcoin неприменимыми.

В частности, следующие свойства блокчейна Monero не позволили нам ранее провести необходимый анализ:

- адреса не сохраняются в блокчейне, что не позволяет идентифицировать получателя транзакции;
- суммы транзакций скрываются;
- входы транзакций комбинируются с ложными транзакциями (благодаря протоколу RingCT), что позволяет скрыть источник транзакции.

Несмотря на такие усовершенствования, Monero всё же остаётся уязвимой на определённом уровне отслеживания с применением эвристического анализа [9] и разбором остаточных эффектов хардфорков [10].

Но возможно ли вновь попробовать применить подход ML для деанонимизации Monero? Для этого придётся преодолеть несколько препятствий. Во-первых, выделение свойств блокчейна Monero будет сильно топологическим по своей природе; временные метки, количество входов, размер RingCT и связь между транзакциями — всё это тот самый сырой материал, на основе которого предстоит определить необходимые свойства. Во-вторых, из клиента кошелька monerod, если сравнивать его с клиентом кошелька bitcoind, были удалены несколько свойств, которые используются аналитиками данных блокчейнов, но в которых нет никакой необходимости при повседневном использовании монеты. Наконец, большие репозитории меток транзакций будут либо недоступны, либо ненадёжны, либо засекречены.

Чтобы справиться с первым препятствием, о котором будет говориться в Разделе 3, мы количественно оценили некоторые топологические аспекты окружения определённой транзакции в блокчейне. В случае со вторым препятствием мы воспользовались открытым репозиторием, стоящим за xmrchain.net [11]. Для преодоления третьего препятствия нами были собраны и верифицированы метки, собранные с помощью API

биржи ShapeShift. Эти метки позволили нам деанонимизировать только одно лицо, ShapeShift, и несмотря на то, что в Разделе 5 мы используем их, чтобы продемонстрировать, как можно использовать свойства применительно к реальному блокчейну Monero, из соображений сохранения общности и повышения уровня достоверности нами был выбран иной курс, который, надеемся, вдохновит других на проведение дополнительного анализа приватности Monero и других блокчейнов.

В исследовательских работах, посвящённых проблемам кибербезопасности, как правило, генерируются искусственные наборы данных, позволяющие разработать и оценить методы обнаружения угрозы [12]. Мы используем эту аналогию для выработки концепции, которая обеспечит понимание принципов работы Monero на практике; нами были созданы тестовые сети, включающие сеть из десяти и сеть из пятидесяти пользователей, позволившие создать набор данных для анализа ML. Для оценки ML нами использовался блокчейн, то есть, публично доступный сегмент сети, что, в свою очередь, позволило разработать свойства, а на основе содержания кошельков нами были получены метки, как показано на рис. 1.

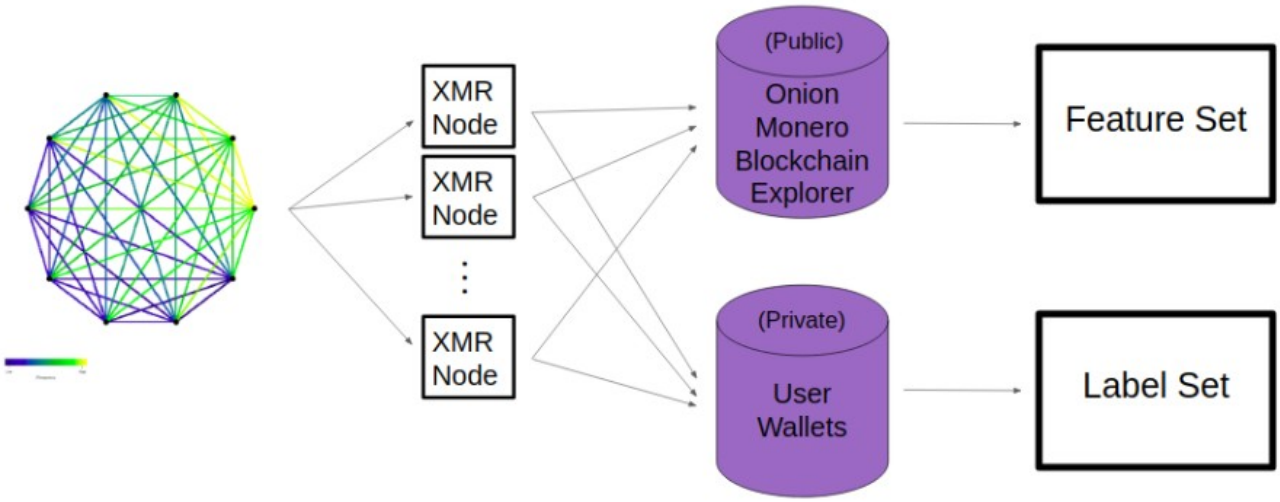


Рисунок 1. Тестовая сеть, использованная для формирования блокчейна и получения меток

Monero также даёт те новые возможности машинного обучения, которые были недоступны или неприменимы в случае с подобными Bitcoin криптовалютами. В Разделе 4, помимо исследования задачи идентификации агентов, нами также рассматриваются и другие случаи. Например, мы можем использовать полученные свойства в рамках регрессионного анализа, чтобы найти отсутствующую сумму транзакции. Несмотря на то, что в этом случае попытка не увенчалась успехом, нам удалось достичь результата с точки зрения восстановления реального входа среди ложных транзакций.

2 Моделирование блокчейнов

Чтобы смоделировать экономическую систему Monero, в которой присутствует множество агентов, взаимодействующих друг с другом, одновременно занимаясь при этом майнингом, нами было использовано приложение tmux, позволяющее создавать на одной машине множество диалоговых сеансов и управлять ими. Для создания tmux-сети агентов Monero на базе жёстко закодированной информации кошельков нами был запущен сценарий командной строки. В этой сети каждый агент запускает процесс майнинга и удалённый вызов процедуры (RPC) кошелька в соответствии с назначенным набором портов. Все агенты сети соединены между собой и поддерживают синхронизацию блокчейна.

В файлах экономической системы указано, как она должна работать. Каждый такой файл содержит список данные сумм транзакций, адресов получателей, а также значения времени ожидания между транзакциями, которым должен следовать определённый агент. Эти файлы были сгенерированными нами с помощью стохастической модели. Для обработки данных файлов и создания HTTP-запросов транзакций, передаваемых RPC кошельков Monero, мы, соответственно, использовали программу Python.

Помимо обязательного ограничения, связанного с тем, что время ожидания между транзакциями должно быть достаточным для того, чтобы задержка при обработке транзакций была относительно небольшой, содержание файлов экономической системы может быть настолько разнообразным, насколько позволит воображение. Нами было рассмотрено несколько экономических сценариев с различными допусками, в соответствии с которыми и были созданы соответствующие файлы экономической системы.

В рамках трёх из пяти сценариев нам и рассматривается экономическая система, в которой участвуют десять агентов. В остальных двух случаях присутствует пятьдесят агентов. Интервалы ожидания между транзакциями и суммы транзакций генерируются с помощью дистрибутивов Poisson. В случае со сценарием s03 интервалы ожидания агентов определяются различными параметрами Poisson в диапазоне от 45 до 90 000, в то время как суммы транзакций зависят от одного и того же параметра Poisson для всех агентов. В рамках сценария s04 для создания сумм транзакций также используются различные параметры Poisson. Сценарий s05 был таким же, как и s03, но в этом случае файлы экономической системы были разделены на два пула. То есть, агенты взаимодействовали только с тем пулом, к которому принадлежали. Пулы были разделены таким образом по причине того, что публичные данные транзакций являются одной из проблем машинного обучения, которую мы стремимся решить.

В случае с одним из двух сценариев с пятьюдесятью агентами (s06) нами было создано два пула с равным количеством агентов. Кроме того, для этих двух пулов мы создали два разных цикла транзакций, смоделировав разницу между временными зонами, которая существует в реальном мире. Наконец, в последнем сценарии с пятьюдесятью агентами (s07) мы поделили сеть транзакций на пять пулов по десять агентов в каждом. На рис. 2 представлены наши экспериментальные настройки.

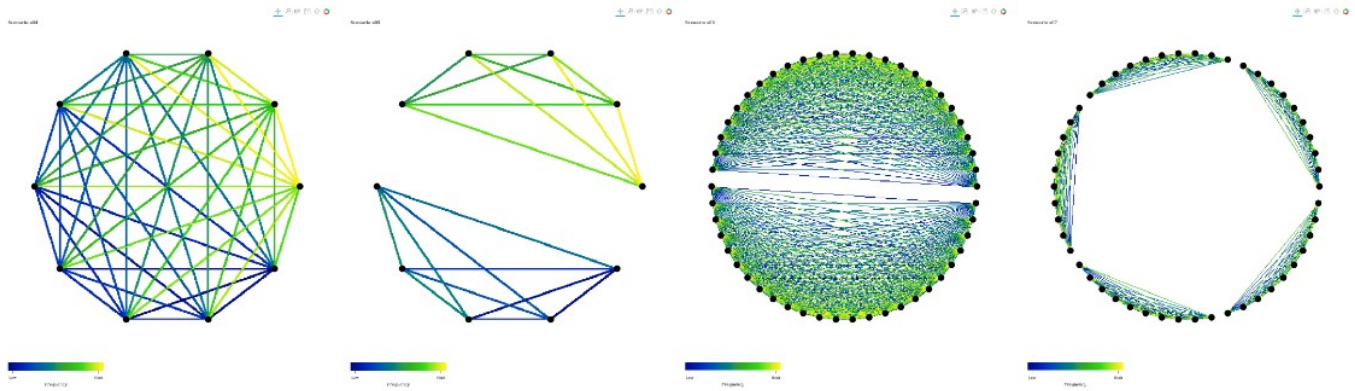


Рисунок 2. Графы файлов экономической системы, демонстрирующие взаимосвязь партнёров по совершению сделок, расположенных по рёбрам, цвет при этом отражает размер транзакции

Чтобы исследовать данные блокчейна, мы подключили луковый блокчейн-эксплорер Monero (Onion Monero Blockchain Explorer) к блокчейну одного из агентов. В таблице кратко описаны сгенерированные наборы данных, а рис. 3 свидетельствует о невероятной разрежённости, которую можно достичь, удалив ложные транзакции из блокчейна.

Модель	Количество блоков	Количество транзакций
s03	23 812	4898
s04	25 509	4923
s05	41 583	4923
s06	37 281	24 807
s07	58 551	7070

Таблица 1. Краткое описание наборов данных моделей

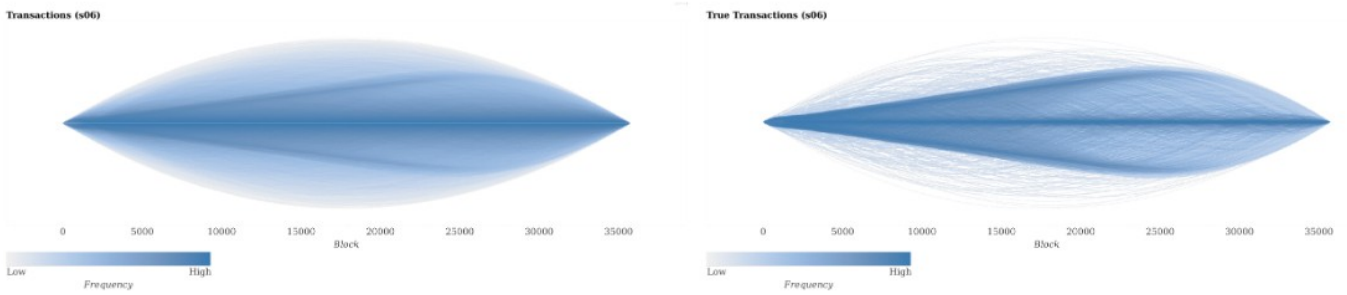


Рисунок 3. Удаление ложных транзакций даёт более чёткую картину активности блокчейна

3 Определение свойств блокчейнов Monero

Реальный блокчейн Monero. В случае с большинством реальных криптовалютных блокчейнов сложность выполнения задач машинного обучения состоит в том, что получение меток, необходимых для контролируемого обучения, остаётся основной проблемой, даже несмотря на доступность информации о том, что происходит в блокчейне. Это особенно актуально, когда речь заходит о Monero.

В определённых ситуациях метки и скрытая информация блокчейна могут быть получены правоохранительными органами, хедж-фондами и просто теми, кто по-настоящему интересуется этой проблемой. Нами были использованы данные, взятые с сервиса Shapeshift. Набор данных содержал записи примерно 1 700 000 транзакций, 20 000 из которых являлись транзакциями ShapeShift, переводимыми в Monero.

В случае с каждой записью транзакции данные содержали 7 свойств с нулевым переходом (0-hop), то есть, свойств, характерных для транзакций в том виде, в котором они появляются в интерфейсе эксплорера Monero, например, временные метки, размер кольца, день недели, час и так далее. Из свойств с нулевым переходом также собираются 175 свойств с единичным переходом (1-hop), совокупная статистика, включая среднее, максимальное и стандартное отклонение, связанные со свойствами с нулевым переходом из соседства транзакции. 182 свойства прошли Z-нормализацию до выполнения задачи предиктивного анализа.

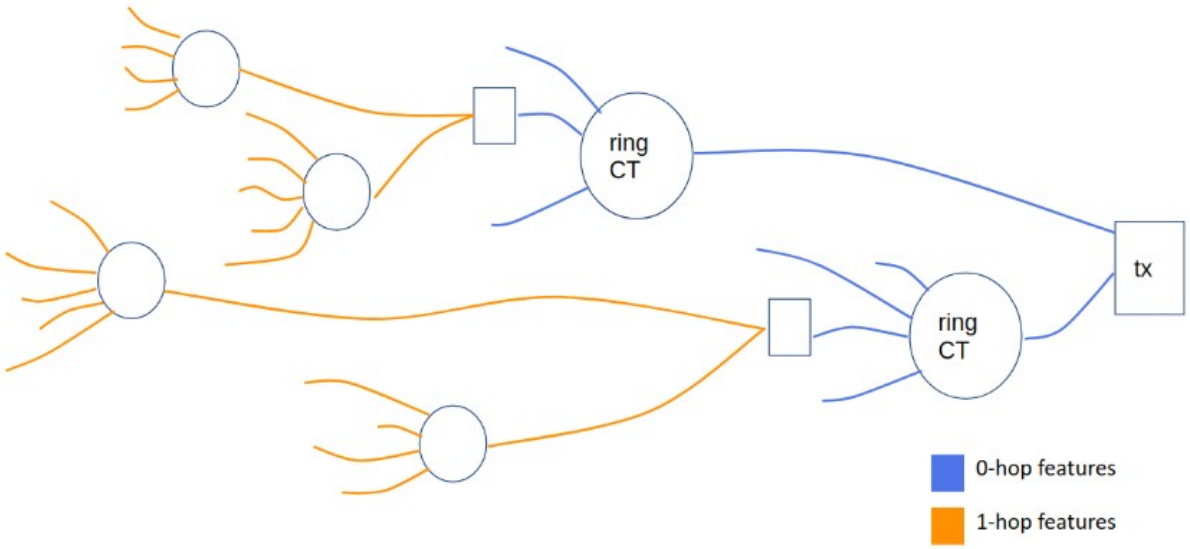


Рисунок 4. Свойства соседства транзакции определяются путём сбора статистических данных транзакций с нулевым и единичным переходом

Данный подход учитывает информацию соседства определённой транзакции, не требуя при этом извлечения данных всей сети. Мы выяснили, что свойства с одним переходом оказались довольно информативными с точки зрения прогнозирования транзакций Shapeshift. В тестовых сетях наличие информации из соседства транзакции также упрощает прогнозирование участников группы и задачи регрессии значений транзакций.

В связанной с этой работе [13] нами предлагается использовать корреляционную статистику, чтобы избежать помех, вызываемых применения механизма смешивания. Это свойство вытекает из гипотезы, связанной с многократным использованием кошелька: если транзакция содержит множество RingCT, реальные входы в каждом кольце, добавляемые одним и тем же пользователем или пользователями, создают некий поведенческий шаблон. Например, получатель может ожидать, что сумма токена будет превышать сумму, имеющуюся в одном кошельке, поэтому отправитель использует входы из множества кошельков, с которыми ранее заключал сделки с подобной частотой. Мы количественно оценивали такое интуитивное поведение путём составления таблицы, содержащей статистику транзакций, содержащих ровно два кольца. Запись (*i, j*) корреляционной матрицы указывает на временную корреляцию между самой старой записью *i* в первом кольце и самой старой записью *j* во втором. Биннинг матрицы может быть представлен в часах или с помощью разделения относительной разности временных меток между входами *i* и *j*. Несмотря на то, что наши свойства не связаны с измеренной корреляционной матрицей напрямую, они были выбраны как чувствительные к подобным эффектам.

Тестовые сети. Каждая и з пяти тестовых сетей s03, s04, s05, s06 и s07 послужила источником вводных данных для решения задачи регрессионного анализа суммы транзакции, в то время как прогнозирование участников группы изучалось с помощью трёх последних сетей. Предварительная обработка информации блокчейн-эксплорера Monero и выделение свойств следуют за процессом выделения свойств реального блокчейна Monero.

4 Машинное обучение

4(a) Машинное обучение распознаванию ложных/реальных транзакций

Новая интересная задача ML, с точки зрения которой Monero является уникальной криптовалютой, состоит в распознавании, какой из заданного набора элементов RingCT является реальной предыдущей транзакцией. Предлагаемый нами корреляционный анализ предполагает, что поведенческие шаблоны, например, то, в какое время дня пользователь обычно проводит транзакции, позволяют идентифицировать реальный вход. Вычисленные нами свойства чувствительны к этим различиям и позволяют предположить, что ML распознает реальный сигнал, несмотря на наличие миксинов.

На рис. 5 показано, как решается такая задача. Чёрным выделен скриншот блокчейна, где красная стрелка указывает на реальный вход. Его получилось идентифицировать путём анализа кошельков (показаны белым) по окончании моделирования.

Рисунок 5. Чёрным выделен публичный блокчейн, а белым — транзакции определённого пользователя. Задача ML состоит в восстановлении индекса реальной транзакции

4(b) Машинное обучение идентификации личности пользователей и участия в группе

Для решения задачи идентификации пользователей на основе файлов регистрации транзакций мы задействовали нейронные сети и алгоритм Random Forest для их классификации. В данном случае наблюдение

велось за транзакцией — мы классифицировали группу, к которой принадлежал получатель транзакции. Группы во всех сценариях взаимодействовали строго в пределах группы и только в течение заданных временных интервалов. Данные временные интервалы повторялись циклически, как бы имитируя временные зоны.

В случае с обеими моделями нами был проведён рандомизированный поиск по всем гиперпараметрам. Нейронная сеть состояла из двух плотных слоёв со 182 и *n* нейронов, соответственно, где второй слой находился в диапазоне от 10 до 30.

Рандомизированный поиск по гиперпараметрам Random Forest включал в себя: количество деревьев решений, минимальный размер выборки при разделении и критерии такого разделения. Поскольку алгоритм Random Forest подразумевает интерпретируемое взвешивание важности свойств, тремя самыми информативными свойствами для определения участия в группе будут: **S05** — информативные свойства S05 включают в себя минуту появления входа среди других входов транзакции. Они перечислены в убывающем порядке по их важности: по сумме среднего значения минут появления входов, среднее значение максимального количества минут появления входов, среднее значение минимального количества минут появления входов. **S06** — информативными свойствами S06 являются: сумма секунд появления входов, среднее значение максимального количества часов появления входов и сумма максимального количества часов появления входов. **S07** — информативными свойствами S07 являются: среднее значение минут появления входов, максимальное количество минут появления входов и средне значение максимального количества минут появления входов.

4(c) Машинное обучение идентификации сумм транзакций

Эффективность модели при выполнении задачи регрессионного анализа суммы обозначается коэффициентом R^2.

$$R^2 = 1 - \frac{\sum_i (y_i - \tilde{y}_i)^2}{\square}$$

где μ_y является ожидаемой суммой транзакции. Для решения данной задачи мы ввели базовый предиктор, который прогнозировал ожидаемые суммы транзакций ShapeShift, а значение эффективности его R^2 было равно 0. Идеальным результатом является 1, а модели, улучшающие наши базовые показатели эффективности, находятся в диапазоне ¿.

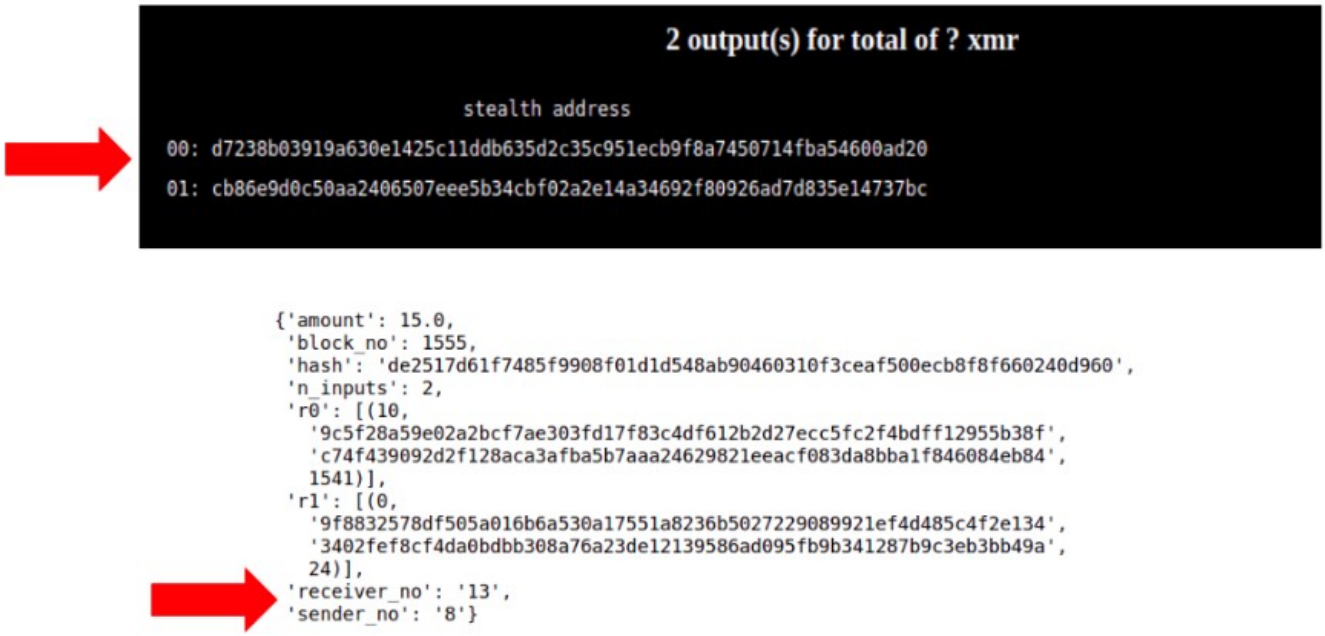


Рисунок 6. Участие в группе определяется на основе файлов кошелька и экономических команд

Чтобы спрогнозировать суммы транзакций в записи *i*, нами были задействованы две модели: эпсилон регрессии опорных векторов (SVR) и нейронная сеть.

В случае с обеими моделями был проведён рандомизированный поиск по гиперпараметрам. Гиперпараметры SVR включали в себя ядро, параметр регуляризации C и параметр допусков ϵ. В случае с обеими моделями нами был проведён рандомизированный поиск по всем гиперпараметрам. Нейронная сеть состояла из двух плотных слоёв со 182 и n нейронов, соответственно, где n находилась в диапазоне [10, 30]. Скорость обучения отличалась от гиперпараметра, по которому осуществлялся поиск. При наличии заданного набора гиперпараметров рабочие характеристики модели вычисляются по пятикратным значениям данных, а затем усредняются.

Мы выяснили, что для прогнозирования сумм требуется больше информации о транзакциях. Значение R^2 модели SVR составило -0,16, а значение нейронной сети -0,1, то есть, оба находились ниже базового параметра, равного 0. Дальнейшие направления регрессионного анализа включают в себя:

1. При наличии множества колец — определение корреляции на уровне входов между кольцами.
2. Использование известной информации на уровне пользователя с целью прогнозирования суммы.

4(d) Результаты ML

На рис. 8 показаны результаты применения нашего подхода к машинному обучению. Мы полагаем, что единообразие пользователей и серьёзная задержка во времени проведения транзакций, если сравнивать с завяленными характеристиками решения, сыграли свою роль в случае со сценарием 6 и сценарием 7.

5 Машинное обучение на базе реальной сети Monero, идентификация транзакции ShapeShift

Ранее в работе [2] мы собирали и проводили валидацию транзакций ShapeShift с помощью API биржи. Сбор данных позволил получить метки транзакций Bitcoin, ZCash, Litecoin и Dash. Определение свойств в случае с криптовалютами, подобными Bitcoin, позволило нам успешно восстановить 73% транзакций ShapeShift, проанализировав при этом всего 20% блокчейна Bitcoin.

В таблице 2 мы расширили данный анализ, чтобы идентифицировать те транзакции ShapeShift, в которых целевой криптовалютой являлась Monero. К нашему удивлению, результаты применения классификатора в случае с Monero превзошли результаты Bitcoin. Мы считаем, что причиной является тот факт, что на пике в 2018 году на долю ShapeShift приходилось 4% всего блокчейна Monero, в то время как всего одна из тысячи транзакций Bitcoin совершалась через ShapeShift. И несмотря на то, что наш набор данных был в значительной степени не сбалансирован, степень этой несбалансированности была не столь уж и серьёзной.

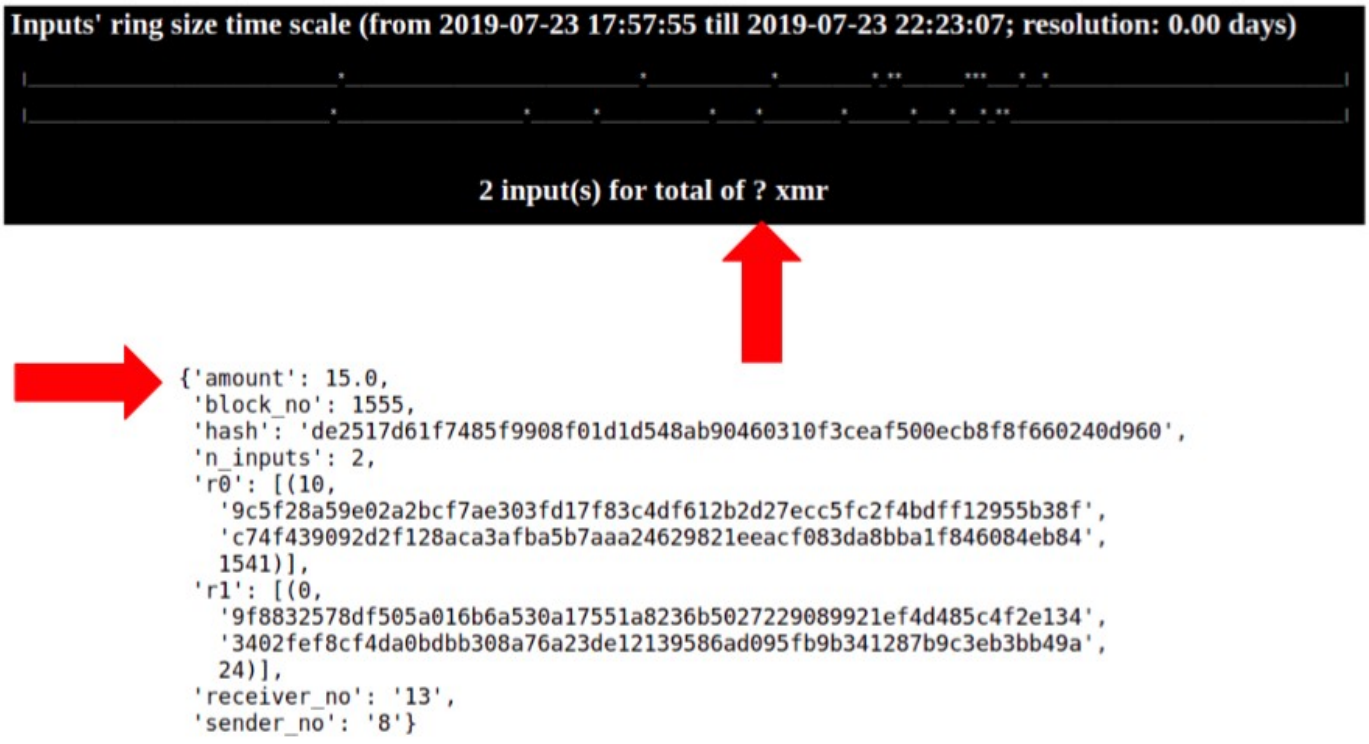


Рисунок 7. Значение суммы транзакции отсутствует (показано чёрным). Мы попытались восстановить его на основе данных кошелька (показано белым) путём регрессионного анализа свойств

Параметр	Краткая статистика	Транзакции ShapeShift	Другие транзакции
Точность	**В среднем**	0,050062	0,999161
	Среднеквадратическое отклонение	0,000129	0,000030
Восстановление	**В среднем**	0,941982	0,794504
	Среднеквадратическое отклонение	0,002058	0,000548

Таблица 2. Выделенные нами свойства позволили с высокой точностью восстановить транзакции ShapeShift. Путём анализа всего 20% блокчейна было идентифицировано 94% транзакций

Несомненно, самым важным свойством стало количество колец, использованных в транзакции, при этом семь из остальных девяти самых важных свойств также использовали краткую статистику количества колец, что показано на рис. 9.

6 Заключение

Мы выяснили, что, несмотря на все сложности, возникшие в силу повышенного уровня приватности Monero, блокчейн этой криптовалюты всё же уязвим с точки зрения деанонимизации и сбора информации с применением методов машинного обучения. С целью подтверждения этого факта нами были созданы тестовые сети, необходимые для моделирования блокчейнов. Затем, нами были выделены свойства этих блокчейнов, а данные кошельков были проанализированы так, чтобы выделить соответствующие метки. Monero доказала свою устойчивость к восстановлению скрытых сумм транзакций, в то время как применение классификаторов для идентификации ложных транзакций и идентификации групп/отдельных пользователей позволило получить определённую информацию.

Кроме того, нами было продемонстрировано, что ML применительно к реальному блокчейну Monero может быть использовано для идентификации заданной стороны при условии наличия достаточного количества меток. В качестве примера нами была взята криптовалютная биржа ShapeShift, но мы полагаем, что те же результаты будут действительны и в других случаях, если удастся восстановить большое количество меток, например, на основе данных кошельков, полученных при расследовании преступлений. Несмотря на наличие данных сумм транзакций ShapeShift, попытки регрессионного анализа с целью восстановления этих сумм оказались безуспешными. Мы полагаем, что в будущем можно будет использовать более глубокие свойства, например, осуществлять отслеживание вплоть до coinbase-транзакций, принимать во внимание версию Monero, использованную при проведении предыдущей транзакции, а также задействовать будущие свойства, а не исключительно те, что уже были известны в прошлом, как делалось в этот раз. Возможно, такие свойства позволят заполучить информацию суммы и обеспечат успешность регрессионного анализа, но выполнение такой задачи, безусловно, потребует значительных вычислительных ресурсов.

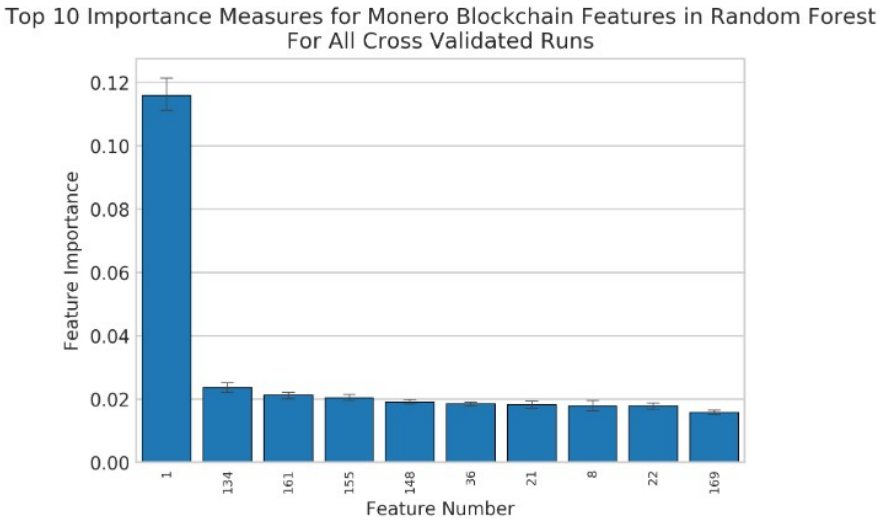
Можно отметить, что в последних версиях по умолчанию в обязательном порядке задан размер колец RingCT равный одиннадцати. Поскольку свойства, полученные на основе размера колец, были самыми информативными в случае с анализом транзакций ShapeShift, мы полагаем, что такое изменение вы значительной степени повысит уровень приватности Monero. Тем не менее, прошлые исследования показали, что история проведения прошлых транзакций ещё какое-то время будет актуальной, вероятно, до того момента, пока эти новые характеристики не изменят лежащего в основе распределения с войств.

Classification Валидация точности нейронной сети с помощью Random Forest	**Сценарий 4** 10 участников	**Сценарий 5** 10 участников в 2 группах	**Сценарий 6** 50 участников в 2 группах	**Сценарий 7** 50 участников в 5 группах
Ring CT (прогнозирование реального входа среди миксинов)	22,1% 18,7%	32,4% 28,9%	31,9% 31,4%	31,4% 32,8%
Прогнозирование ID участника	42,9% 39,8%	41,6% 40,7%	1,8% 2,1%	4,9% 4%
Прогнозирование группы	Отсутствует	96,5% 96%	52,9% 53,1%	25% 22,7%
Классификаторы, превысившие значения случайного угадывания выделены зелёным				

Рисунок 8. Результаты ML при решении различных задач. При прогнозировании сумм транзакций результаты регрессионного анализа не превысили базового значения определения средних значений сумм транзакций

Нам бы хотелось поблагодарить Лабораторию аналитических наук (LAS) при Университете штата Северная Каролина за финансирование и непрерывную поддержку. Ирония состоит в том, что, несмотря на

обилие кода и данных Monero, лукового блокчейн-эксплорера и API ShapeShift, учитывая, что все участники привержены принципам приватности и безопасности в сфере блокчейн-технологии, всё это пригодилось для проведения исследования, направленного на снижение уровня приватности, и мы благодарны им за их вклад.



(a) the features were ranked in accordance to their importance

1: num_rings
134: min_min_num_rings
161: sum_median_second
155: sum_median_num_rings
148: sum_mean_num_rings
36: mean_sum_num_rings
21: mean_median_second
8: mean_mean_num_rings
22: mean_max_num_rings
169: sum_min_num_rings

(b) Descriptors for the important features

Рисунок 9. Характеристики топологии RingCT были довольно информативными

Ссылки

[1] Натан Борггрен (*Nathan Borggren*). Глубокое обучение идентификации поведения лиц в экономике Bitcoin (*Deep learning of entity behavior in the bitcoin economy*). https://ncsu-las.org/wp-content/uploads/2017/12/borggren-gda_bitcoin.pdf, 2017.

[2] Натан Борггрен (*Nathan Borggren*), Гэри Коплик (*Gary Koplik*), Пол Бендич (*Paul Bendich*) и Джон Харер (*John Harer*). Деанонимизация ShapeShift: связывание транзакций из нескольких блокчейнов (*Deanonymizing shapeshift: Linking transactions across multiple blockchains*). https://ncsu-las.org/wp-content/uploads/2019/01/LAS_Shapeshift_Poster_1543186217.pdf, 2017.

[3] Стефан Рэншус (*Stephen Ranshous*), Клифф А. Джослин (*Cliff A. Joslyn*), Шон Крейлинг (*Sean Kreyling*), Кэтлин Новак (*Kathleen Nowak*), Нагиза Ф Саматова (*Nagiza F. Samatova*), Кёртис Л. Уэст (*Curtis L. West*) и Сэмюель Уинтерс (*Samuel Winters*). Направленный гиперграф шаблона майнинга в транзакциях Bitcoin (*Exchange pattern mining in the bitcoin transaction directed hypergraph*). Lecture Notes in Computer Science (включая серии Lecture Notes in Artificial Intelligence и Lecture Notes in Bioinformatics), 10323 LNCS:248{263, 2017.

[4] Кюнейт Гуркан Аккора (*Cuneyt Gurcan Akcora*), Итао Ли (*Yitao Li*), Юлия Р. Гель (*Yulia R. Gel*), и Мюрат Кантаршиоглу (*Murat Kantarcioglu*). Биткоингейст: топологический анализ данных с целью обнаружения вредоносного программного обеспечения в блокчейнах Bitcoin (*Bitcoinheist: Topological data analysis for ransomware detection on the bitcoin blockchain*). CoRR, abs/1906.07852, 2019.

[5] Франческо Зола (*Francesco Zola*), Мария Эгумендиа (*Maria Eguimendia*), Ян Лукас Брюс (*Jan Lukas Bruse*) и Рауль Ордунья Уррутия (*Raul Orduna Urrutia*). Каскадное миашинное обучение с целью проведения атаки на механизмы обеспечения анонимности Bitcoin (*Cascading machine learning to attack bitcoin anonymity*), 2019.

[6] Алес Янда (*Ales Janda*). Кошелёк-эксплорер (*Wallet explorer*). https://www.walletexplorer.com, 2013-2017.

[7] Арон Юсаф (*Haaroon Yousaf*), Джорлд Капос (*George Kappos*) и Сара Микльджон (*Sarah Meiklejohn*). Отслеживание транзакций в криптовалютных реестрах (*Tracing transactions across cryptocurrency ledgers*). CoRR, abs/1810.12786, 2018.

[8] fluffypony и др. Проект Monero (*Monero project*). https://github.com/monero-project/monero/blob/master/src/wallet/wallet2.cpp, 2015.

[9] Эндрю Миллер (*Andrew Miller*), Мальти Мёзер (*Malte Moser*), Кевин Ли (*Kevin Lee*) и Эрвинд Нараянан (*Arvind Narayanan*). Эмпирический анализ связываемости в блокчейне Monero (*An empirical analysis of linkability in the monero blockchain*). CoRR, abs/1704.04299, 2017.

[10] Абрахам Хинтереггер (*Abraham Hinteregger*) и Бернард Хасльхофер (*Bernhard Haslhofer*). Эмпирический анализ отслеживаемости Monero между блокчейнами (*An empirical analysis of monero cross-chain traceability*). CoRR, abs/1812.02808, 2018.

[11] moneroexamples и др. onion-monero-blockchain-explorer. https://github.com/moneroexamples/onion-monero-blockchain-explorer, 2016.

[12] Дж.. Глассер (*J. Glasser*) и Б. Линдауэр (*B. Lindauer*). Заполняем пробел: прагматический подход к генерировнию внутренних данных угрозы (*Bridging the gap: A pragmatic approach to generating insider threat data*). Семинары по безопасности и конфиденциальности данных IEEE 2013 (*2013 IEEE Security and Privacy Workshops*), стр. 98-104, май 2013.

[13] Натан Борггрен (*Nathan Borggren*) и Лихан Яо (*Lihan Yao*). Корреляция транзакций Monero с множеством входов (*Correlations of multi-input monero transactions*), 2019.