Повышение уровня обфускации в рамках протокола CryptoNote

Адам Маккензи (Adam Mackenzie)* Шурэ Ноезер (Surae Noether)† и Monero Core Team
Исследовательская лаборатория Monero (Monero Research Lab)
26 Января 2015

Аннотация

Нами определяется несколько видов атак посредством анализа блокчейна, позволяющих снизить уровень неотслеживаемости, обеспечиваемый протоколом CryptoNote 2. Также нами предлагается анализ возможных решений, рассматриваются соответствующие достоинства и недостатки этих решений, а также рекомендуемые улучшения протокола Monero, которые, как мы надеемся, поспособствуют долгосрочной защите криптовалюты от атак путём анализа блокчейна. Рекомендуемые нами улучшения протокола Monero включают в себя реализацию политики минимального смешивания $n = 2$ сторонних выходов на одну подпись на общесетевом уровне и уровне протокола, повышение этого значения до $n = 4$ на уровне протокола спустя два года, а также введение значения $n = 4$ в промежуточной перспективе. Нами также предлагается метод отправки выходов Monero, подобный торрент-протоколу. Помимо прочего, мы рассматриваем метод неравномерного, привязанного к «возрасту» выбора миксинов, позволяющий избежать других форм анализа блокчейна, которые также указаны в этой работе, но в силу различных причин нами не даётся каких-либо формальных рекомендаций по реализации. Также нами довольно подробно обсуждаются последствия предлагаемых улучшений.

Настоящий исследовательский бюллетень не проходил независимого анализа и отражает исключительно результаты внутренних исследований.

1 Обзор

Нами рассматриваются несколько проблем, связанных с безопасностью и отслеживаемостью при использовании кольцевых подписей, с целью обеспечения анонимности транзакций цифровых валют. Следует помнить о том, что протокол CryptoNote доказывает безусловное отсутствие связываемости, но неотслеживаемость остаётся недоказанной; в действительности CryptoNote предполагает довольно высокий уровень отслеживаемости, что будет показано нами в этом документе.

Нам бы хотелось напомнить читателю, что две транзакции считаются *несвязываемыми*, если наблюдатель не может доказать, что эти транзакции были отправлены одному и тому же пользователю. Мы считаем транзакцию *неотслеживаемой*, если все возможные отправители транзакции являются равновероятными. Таким образом, проблемы, связанные с неотслеживаемостью и характерные для протокола CryptoNote, подразумевают, что, даже несмотря на то, что пользователи без какого-либо риска для собственной анонимности могут *принимать* криптовалюты, в основе которых лежит CryptoNote, совершенно не гарантировано, что они смогут *потратить* такую валюту, не раскрыв какой-то информации, касающейся их предшествующих транзакций. Безусловно, если не использовать подход неинтерактивного доказательства с нулевым разглашением (NIZK) (подобного тому, что предлагается

*lab@getmonero.org
†surae.noether@protonmail.com

в работе [2]), то отслеживаемость станет неизбежной. Тем не менее подходы, в основе которых лежат доказательства NIZK, до этого момента имели довольно ограниченное применение, а сложность вычислений, связанных с ними, являлась серьёзным недостатком. Решение некоторых вопросов, связанных с возможностью отслеживания при отсутствии доказательств с нулевым разглашением, является основной задачей настоящего документа. Критически важно, чтобы читатель отметил, что предлагаемые нами методы никак не препятствуют возможности предоставления пользователями историй их транзакций для проведения аудита, если это потребуется в силу действия местного законодательства (см. Раздел 6).

В Разделе 2 мы показываем, как *открытая трата* или *трата при нулевом смешивании* обеспечивает возможность отслеживания, что также подробно рассматривается в работе [3]. В Разделе 3 нами указываются три потенциальных направления атаки, которые могут быть использованы с целью снижения свойства неотслеживаемости, обеспечиваемого протоколом CryptoNote. При этом мы не утверждаем, что предлагаемый список является полным. В частности, в подразделе 3.1 подробно описано, как информация о возрасте выходов транзакции может использоваться для нейтрализации свойства неотслеживаемости CryptoNote. В подразделе 3.3 также подробно показано, как открытые данные, касающиеся транзакций и используемые в кольцевой подписи, могут способствовать проведению комбинаторного анализа, позволяющего связать выходы транзакции, которые в противном случае связать было бы невозможно. В подразделе 3.2 мы демонстрируем, как неосторожное использование выходов транзакций в составе какой-либо новой транзакции может позволить злоумышленнику прийти к определённым выводам в отношении принадлежности выходов оригинальной транзакции; мы называем это «атакой с ассоциацией по использованию».

В Разделе 4 мы даём несколько подробных рекомендаций, направленных на то, чтобы избежать подобных видов атаки, а также несколько менее подробных рекомендаций, касающихся других возможных вариантов атаки. В подразделе 4.1 нами указывается обязательное минимальное количество миксинов $n = 2$, которое следует использовать на сетевом уровне, а также обосновывается выбранное нами значение. Подразделы 4.2 и 4.3 содержат подробные рекомендации, которые позволят избежать атак с применением комбинаторного анализа и атак с ассоциацией по использованию, описанных в подразделах 3.3 и 3.2, соответственно.

В Разделе 5 мы подробно рассматриваем планы по реализации наших предложений. Наконец, в Разделе 6, нами предлагается метод, позволяющий соблюсти местные законы без ущерба для анонимности пользователя, которую мы всеми силами стараемся поддерживать.

2 Отслеживаемость при тратах с нулевым смешиванием

В этом разделе нами затрагиваются несколько проблем, связанных с неотслеживаемостью в рамках протокола CryptoNote и других схем кольцевой подписи. Вопросы, подробно рассмотренные в работе [3] , конечно же, не являются всеохватывающими, и в этом документе мы коснемся других проблем. Мы точно так же не утверждаем, что проблемы, затронутые нами в документе, являются исчерпывающими. В подразделе 3.1 нами описаны атаки, связанные с возрастом выходов транзакций, в подразделе 3.3— определённые комбинаторные виды атак, которые могут использоваться для обеспечения возможности отслеживания при помощи внешних данных для кольцевых подписей, таких как высота блока, и в подразделе 3.2 мы показываем, как неосторожное совместное использование выходов транзакций может дать злоумышленнику, перехватившему эти данные, информацию относительно принадлежности транзакции. Процесс цепной реакции в результате открытой траты, описанный в работе [3], является хорошей отправной точкой для начала обсуждения проблемы.

При использовании протокола кольцевой подписи, подобного CryptoNote, с целью защиты от отслеживания и связываемости, злоумышленник, обладающий большим количеством выходов транзакций, может вызвать цепную реакцию, связанную с возможностью отслеживания, пожертвовав собственной анонимностью. Это возможно в случае с так называемыми «транзакциями с нулевым смешиванием», не использующими в качестве миксинов чужие выходы при создании соответствующей кольцевой подписи. Перед тем как продолжить, хотелось бы, чтобы читатель определился с некоторой терминологией. Во-первых, следует помнить о том, что в случае с прозрачными криптовалютами, такими как Bitcoin, выходы транзакций можно проверить и определить, были они уже потрачены или нет. На практике часто используется акроним UTXO, который расшифровывается как *unspent transaction output (непотраченный выход транзакции)*. В случае же с протоколом CryptoNote этот акроним использовать будет неверно, так как довольно трудно определить, был ли потрачен выход, не прибегая к какой-либо форме анализа блокчейна или атаки, направленной на обеспечение возможности отслеживания, подобной тем, что мы описываем в этой работе. Как следствие, в настоящем документе мы не будем обозначать выходы транзакций как потраченные или непотраченные. Во-вторых, нами будет рассмотрена практическая трата выходов транзакций в составе кольцевой подписи, сгенерированной без использования выходов других транзакций в качестве миксинов, то есть *открытая трата* или *трата с нулевым смешиванием*. Открытая трата подразумевает создание кольцевой подписи без миксинов; кольцевая подпись в данном случае эквивалентна обычной цифровой подписи.

Изначально поднимаемые нами вопросы безопасности описать довольно просто. Как и в оригинальном документе по кольцевым подписям [4], рассмотрим ситуацию, когда правительственный служащий организует утечку секретной информации путём сбора публичных ключей n других сотрудников (миксинов). Затем он использует эти ключи и свой собственный приватный ключ, чтобы создать кольцевую подпись, которой будет подписан документ, утечка которого им планируется. У внешнего наблюдателя, очевидно, будет 1 из $n+1$ шансов угадать, кто из сотрудников организовал утечку секретной информации. Что произойдёт, если сотрудники, владеющие ключами к миксинам, станут отрицать свою причастность к утечке? Им придётся одному за другим добровольно раскрыть свои приватные ключи, чтобы доказать начальству или друг другу, что они этого не делали? Сначала, после того как первый сотрудник раскроет свой приватный ключ, уровень обфускации снизится с 1 из $n+1$ до 1 of n. После того как это же сделает второй сотрудник, уровень обфускации упадёт с 1 из n до 1 of $n-1$ и так далее. При таком сценарии кольцевые подписи становятся значительно слабее. Как было указано в работе [3], при итеративном (в транзакции за транзакцией) использовании кольцевых подписей проблема только усугубляется. В силу цепной реакции безопасность одной транзакции может зависеть от безопасности другой транзакции даже при отсутствии напрямую связанных между собой сторон. Безусловно, кольцевые подписи будут по-прежнему считаться действительными как *цифровые подписи*, и их роль в качестве основы работы криптовалютной сети не будет вызывать сомнения. Вопрос заключается в решении проблемы с подобным снижением уровня неотслеживаемости.

Рассмотрим в качестве примера следующую ситуацию. Допустим, Элис создаёт кольцевую подпись, используя миксины {Боб, Синтия, Даг}. Значит, для любого наблюдателя список возможных отправителей транзакции, очевидно, будет выглядеть как {Элис, Боб, Синтия, Даг}. Если Боб, Синтия и Даг затем потратят выходы своих транзакций без миксинов, то любому внешнему наблюдателю станет ясно, что они более не владеют этими выходами транзакций. Следовательно, их можно исключить из списка возможных отправителей транзакции Элис, и теперь ясно, что отправителем транзакции является именно она. Всё будет выглядеть не так, как будто только три пользователя (Боб, Синтия и Даг) потратили свои выходы в открытую, а будет видно, что все четверо провели транзакции таким образом. Следовательно, любые подписи, использующие выходы транзакций Элис в качестве миксинов, теперь

также будут скомпрометированы. При наличии достаточного количества таких скомпрометированных транзакций могут возникнуть эффекты второго порядка, что приведёт к цепной реакции.

Нами было смоделировано создание кольцевых подписей, как оно представлено в работе [3]. При этом нами был использован вариант модели урны Поля, подробно описанной в работе [1], что позволило нам продемонстрировать, что цепная реакция с точки зрения отслеживания вполне реальна. Следовательно, если все миксины, которые были использованы для создания новой подписи, контролируются злоумышленником, значит, такой злоумышленник контролирует и неотслеживаемость этой кольцевой подписи. Модель урны Поля имеет недостаток, связанный с тем, что *«богатый становится ещё богаче»*; то есть злоумышленник, контролирующий множество выходов в начале, вероятно, с течением времени будет контролировать всё больше и больше выходов, если не соблюдать осторожность при выборе соответствующих параметров.

Положительным свойством CryptoNote в данном случае является отсутствие возможности установления прямой связи между пользователями и выходами транзакций. Можно только выяснить, связан ли образ ключа с конкретным выходом транзакции. По сути, так называемые *скрытые адреса*, используемые протоколом CryptoNote, гарантируют, что анонимность не будет нейтрализована полностью. Точнее, когда Боб, Синтия и Даг тратят свои выходы транзакций в открытую, любой может связать эти выходы с образами ключей посредством блокчейна. В этом случае, когда будет очевидно, что образ ключа Элис не использовался ни в одной из этих транзакций, этот *образ ключа* можно будет связать с соответствующим выходом, но не с личностью Элис.

2.1 Сдача и пыль как причина траты с нулевым смешиванием

Теперь мы покажем, как пользователям регулярно приходится проводить транзакции в открытую, вопреки собственному желанию.

Кто-то может разумно предположить, что разработчиками оригинальной версии протокола CryptoNote было реализовано стандартное требование к номиналу всех выходов транзакций, но это не так. Например, в то время как сумма в размере 102.5 XMR может быть разбита на 100 XMR, 2 XMR и 0.5 XMR, протокол вовсе не запрещает разбить её на суммы с необычным номиналом 100 XMR, 2 XMR, 0.499999999 XMR и 0.000000001 XMR. Выход также нельзя будет смешать, поскольку, как правило, он будет деноминирован, как в случае с 0.499999999 XMR (а следовательно, набор выходов транзакций, связанных с деноминацией, будет пустым), или же сумма выхода будет меньше, чем размер комиссии, как в случае с номиналом 0.000000001 XMR.

Ниже приводится простой пример того, как при использовании существующего кода деноминации генерируется пыль. В предлагаемой ситуации у Боба в кошельке имеется 103.32111111, и он решает отправить Элис 53.32111111 XMR, воспользовавшись двумя сторонними выходами транзакций в качестве миксинов, и получить 50 XMR сдачи. Предположим, 103.32111111 XMR Боба разбиваются на выходы с номиналом 100 XMR, 3 XMR, 0.3 XMR, 0.02 XMR и пыль 0.00111111 XMR. Существующий код деноминации разобьёт выход для 100 XMR, принадлежащий Бобу, на два выхода по 50 XMR каждый, и отправит один выход Элис и один Бобу. Затем код отправит все остающиеся в кошельке Боба выходы (3, 0.3, 0.02, 0.00111111) Элис, при этом каждый выход будет подписан собственной кольцевой подписью с минимально возможным уровнем смешивания. К сожалению, выход с пылью (.00111111) представляет собой уникальную категорию в том смысле, что набор выходов транзакции, отвечающий за деноминацию .00111111, будет пустым. Следовательно, использование миксинов будет невозможным. Всё это можно представить в виде следующей диаграммы:

$$100 \text{ XMR (Боб) (кол-во миксинов (mix-in)=2)} \longrightarrow 50 \text{ XMR (Элис)}, 50 \text{ XMR (Боб)} \quad (1)$$

$$3 \text{ XMR (Боб) (кол-во миксинов (mix-in)=2)} \longrightarrow 3 \text{ XMR (Элис)}$$

$$0.3 \text{ XMR (Боб) (кол-во миксинов (mix-in)=2)} \longrightarrow 0.3 \text{ XMR (Элис)}$$

$$0.02 \text{ XMR (Боб) (кол-во миксинов (mix-in)=2)} \longrightarrow 0.02 \text{ XMR (Элис)}$$

$$0.00111111 \text{ XMR (Боб) (кол-во миксинов (mix-in)=0)} \longrightarrow 0.00111111 \text{ XMR (Элис)}$$

В случае с этой транзакцией уже очевидны некоторые проблемы, связанные с анонимностью; так как CryptoNote не требует, чтобы для целого числа $1 \leq A \leq 9$ выходы транзакций строго соответствовали форме $A \times 10^B$, мы имеем так называемый выход с пылью 0.00111111 XMR, и этот выход нельзя смешать, если только не найдутся другие выходы с таким же необычным номиналом. Несмотря на то, что анонимность уже обеспечивается кольцевыми подписями со всеми остальными выходами, Боб, по сути, снимает с себя завесу анонимности, включая несмешиваемый выход с пылью в транзакцию. Любой наблюдатель заметит этот странный, несмешанный выход, связанный с данной транзакцией и включённый в тот же блок. Следовательно, наблюдатель может прийти к определённому заключению относительно принадлежности средств.

Что ещё хуже, теперь Элис также является обладателем выхода с пылью, который нельзя потратить в составе кольцевой подписи. И это дополняет вышеописанную проблему. Несмотря на то, что наблюдатель всё же не сможет определить, кто является адресатом транзакции благодаря свойству безусловной несвязываемости, обеспечиваемому протоколом CryptoNote, Элис лишится анонимности собственных транзакций, если когда-либо решит использовать эту пыль в качестве входа, из-за вышеуказанных проблем с неотслеживаемостью.

3 Другие проблемы с отслеживаемостью

В этом разделе нами будут рассмотрены другие проблемы, связанные с безопасностью и отслеживаемостью в рамках протокола CryptoNote. Это не полный список путей проведения атаки с целью нарушения свойства неотслеживаемости CryptoNote, но нами рассматриваются самые насущные проблемы, выявленные командой ведущих разработчиков (Monero Core Team) и членами Исследовательской лаборатории Monero (Monero Research Lab), позволяющие провести атаку, направленную на нарушение неотслеживаемости путём анализа блокчейна.

3.1 Временная связь

В этом подразделе нами будет показано, как возраст выхода транзакции позволяет отследить его в других транзакциях. Существуют по крайней мере две проблемы, связанные с возрастом транзакций. Как и до этого, начнём с эффекта цепной реакции. Предполагается, что более старые выходы использовались в большем количестве кольцевых подписей, чем новые; просто они чаще появляются

в наборах выходов транзакций, и вероятность того, что их выберут из кольцевых подписей, выше. Любая схема создания кольцевой подписи, игнорирующая это, обеспечивает огромный потенциал для эффекта цепной реакции по причине наличия слишком часто используемых выходов. Кроме того, чем раньше злоумышленник получит контроль над набором транзакций, тем выше вероятность того, что он сможет провести атаку путём цепной реакции.

Вторая проблема связана с отслеживаемостью, но не цепной реакцией. При проведении каждой транзакции в составе кольцевой подписи тратится выход, а также соответствующий ряд других выходов транзакции в качестве миксинов. Настоящий выход транзакции, который тратится, также входит в этот ряд. Если в кольцевой подписи использовалось n выходов других транзакций, то ряд будет состоять из $n+1$ элементов. Наблюдатель не может точно сказать, какой из выходов в этом ряду, связанном с кольцевой подписью, является действительным, если не прибегнет к какой-либо форме анализа блокчейна. Тем не менее на практике при наличии определённого выхода транзакции злоумышленник может смоделировать накопленную вероятность, согласно которой выход уже был потрачен по возрастающей функции времени. Выход, который присутствовал в блокчейне на протяжении 202612 блоков, вероятнее всего, уже тратился, в отличие от выхода, который находился в блокчейне всего 202 блока. Следовательно, при отсутствии какой-либо иной внешней информации и при наличии любого выхода, самый «молодой» среди других использованных для создания кольцевой подписи выходов, наиболее вероятно, будет тем самым реальным выходом, который тратится. И наоборот, самый старый выход будет реальным с наименьшей вероятностью.

В этой работе по ряду причин мы не пишем о том, как избежать подобной проблемы. Если было бы возможно определить, были потрачены выходы или нет, мы бы просто выбирали те выходы, которые оставались до этого момента непотраченными. Однако это не так просто сделать, не прибегая к анализу блокчейна. Наше решение предполагает использование метода неравномерного выбора выходов транзакций при создании кольцевой подписи; выбирать выходы транзакций следует на основе их возраста таким образом, чтобы вероятность того, что эти выходы будут выбраны для создания подписи, была обратно пропорциональна вероятности того, что они уже были потрачены. Это предполагает необходимость оценки нами, разработчиками Monero, распределения вероятности, от которого будет зависеть возраст выходов транзакции.

Это, в свою очередь, также является проблемой. Когда биржевой курс испытывает долгосрочное падение (происходит инфляция), рационально мыслящие пользователи, как правило, начинают тратить выходы транзакций, так как уже завтра их выходы будут стоить меньше, чем сегодня, с точки зрения возможности приобретения товаров и услуг, и копить их было бы экономически невыгодно. Когда же биржевой курс испытывает долгосрочный устойчивый рост (то есть происходит дефляция), *те же* пользователи, скорее всего, будут накапливать выходы по обратной причине. Следовательно, распределение выходов транзакций по возрасту будет, по крайней мере, разным с течением времени, учитывая, что какая-то часть пользователей будет рационально исходить из экономических показателей валюты. Было бы неразумно давать рекомендации, касающиеся безопасности, основываясь на экономических характеристиках протокола.

Хитрому наблюдателю можно порекомендовать собирать данные Bitcoin, блокчейн которого прозрачен, и выработать долгосрочное маргинальное распределение на основе полученной таким образом информации. Тем не менее в этом случае мы делаем несколько проблематичных допусков. Мы допускаем, что экономические показатели Bitcoin эквивалентны экономическим показателям Monero или даже подобному эквиваленту, а это не совсем верно из-за относительного объёма экономик. Кроме того, выбирая один определённый тип распределения, который не меняется со временем, мы позволяем злоумышленнику определить различие между постоянным распределением выбора и изменяющимся

со временем распределением, используемым программным обеспечением кошелька. Кто-то предпочтёт изменяющееся со временем распределение, возможно, даже с применением некоторого метода итеративного обновления, но для этого потребуется некая форма данных наблюдения, а также последующая обработка таких данных; как уже было отмечено ранее, не так просто отследить долгосрочное поведение, когда речь заходит о CryptoNote.

Таким образом, выбор фиксированного распределения с целью генерирования случайных чисел для выбора выходов, используемых для создания кольцевой подписи, представляется субоптимальным решением. Альтернативой фиксированному распределению является вероятностный выбор из семейства распределений; простыми словами, мы смогли бы кинуть кости дважды, то есть сначала выбрать распределение, а вторым броском выбрать выходы, воспользовавшись ранее выбранным для создания кольцевой подписи типом распределения. Однако такие смешанные подходы к распределению в конечном счёте эквивалентны выбору единственного фиксированного распределения, а связанные с ними проблемы сводятся к сценарию с первым фиксированным распределением.

Парадоксально то, что невнесение изменений в программное обеспечение равноценно выбору равномерного распределения для каждой кольцевой подписи! То есть в случае с долгосрочным распределением выходов транзакций мы оказываемся в безвыходном положении, так как любой сделанный нами выбор, определённо, станет субоптимальным решением. Команде ведущих разработчиков Monero и Исследовательской лаборатории Monero было бы хорошо следовать той философии разработки, которая была разумно выбрана и предполагает внесение небольших модификаций с их последующим развитием, а не реализацию глобальных изменений с дальнейшей попыткой их обратного масштабирования. Таким образом, несмотря на то, что мы определили саму проблему, связанную с безопасностью, мы не станем давать каких-либо рекомендаций до тех пор, пока у нас не появится дополнительных данных, после чего можно будет определиться с выбором.

3.2 Ассоциация по использованию выходов в пределах транзакции

В этом подразделе нами будет описано, как неосторожное совместное использование выходов позволяет связать соответствующие транзакции с одним пользователем. Другими словами, используемые вместе выходы транзакций, вероятно, будут принадлежать одному и тому же владельцу. Согласно принципу, описанному в подразделе 3.3, мы рассматриваем такой вариант в качестве комбинаторной атаки. Несмотря на то, что кольцевая подпись обеспечивает равновероятность всех возможных подписантов с точки зрения наблюдателя, внешние относительно *кольцевой подписи* данные могут использоваться как дополнительная информация.

Рассмотрим следующий общий пример. Если Боб отправляет Элис несколько Monero, сумма разбивается на несколько выходов транзакции, как в примере 1. Когда Элис решит потратить эти средства, она может использовать пару или большее количество выходов из той же начальной транзакции уже в своей кольцевой подписи. Если она так и сделает, наблюдатель сможет прийти к выводу, что выходы, источником которых является общая базовая транзакция, наиболее вероятно, принадлежат действительному подписанту транзакции. В результате возникает вероятностная проблема, связанная с отслеживаемостью. Это трудно как-либо представить, поэтому рассмотрим вопрос на примере конкретной ситуации, в которой в кошельке у Боба имеется 1 XMR.

Боб хочет отправить Элис 0.75 XMR с комиссией в размере 0.01 XMR. Monero, которые получит Элис, 0.75 XMR, будут ей переведены как два новых непотраченных выхода номиналом 0.7 XMR и 0.05 XMR. Кроме того, выход на сумму 0.01 XMR должен быть передан в качестве комиссии. В результате остаётся 0.24 XMR сдачи, которые Боб получит как два непотраченных выхода транзакции номиналом

0.2 и 0.04 XMR. В какой-то момент после этого Элис вспомнит, что у неё есть 0.75 XMR и решит потратить их где-нибудь. Когда она сделает это, оба её выхода, 0.7 XMR и 0.05 XMR, будут включены в её кольцевую подпись. Наблюдатель может взглянуть на эту подпись и прийти к заключению, что кто-то, создавший кольцевую подпись, вероятно, и является обладателем выходов 0.7 XMR и 0.05 XMR. Даже если бы Элис использовала подпись с множеством миксинов, то, так как она использовала два выхода, источником которых является одна и та же транзакция, эти выходы, наиболее вероятно, можно было идентифицировать как принадлежащие действительному подписанту. Транзакция могла бы выглядеть подобно той, что показана в Таблице 1.

Таблица 1: Связывание выходов транзакции по ассоциации

(a) Транзакция 1, проведённая Бобом

Вход	Выход
1.0 XMR (Боб)	0.7 XMR (Элис)
	0.05 XMR (Элис)
	0.2 XMR (Боб)
	0.04 XMR (Боб)
	0.01 XMR (Комиссия)

(b) Транзакция 2, проведённая Бобом

Вход	Выход
0.7 XMR (Элис)	···
0.05 XMR (Элис)	···

Следует отметить несколько моментов, связанных с этим примером. Во-первых, сам пример практически идентичен комбинаторной атаке с использованием данных высоты блока, описанной нами в подразделе 3.3. Однако в данном примере эти выходы не просто имеют общую высоту блока, но, по сути, и одну корневую транзакцию. Во-вторых, эта проблема симметрична относительно Боба и Элис. Боб также подвергается опасности, когда пытается потратить свой выход сдачи, 0.2 или 0.04, в своей транзакции. Его кольцевые подписи точно так же укажут на его исходную транзакцию. Поэтому, если он решит потратить свои выходы 0.2 и 0.04 XMR в одной транзакции, эти выходы можно будет связать с ним.

3.3 Комбинационные атаки с целью раскрытия выходов

В этом подразделе нами приводится общее описание атаки, о которой говорилось в подразделе 3.2, так называемой комбинаторной атаки, предполагающей проведение комбинаторного анализа, направленного на выявление принадлежности выходов. Основным выводом к изложенному в данном разделе является то, что транзакции следует просто разбивать на несколько меньших транзакций, проводимых в течение какого-то периода времени, как в случае с торрентом, а также, что пользователям следует время от времени переводить свои средства самим себе.

Любые данные, доступные наблюдателю, касающиеся транзакций и необходимые для формирования кольцевой подписи, могут быть использованы для определения принадлежности средств. Рассмотрим следующий пример с применением данных высоты блока. Допустим, Синтией отдельными транзакциями было получено 10, 20 и 30 XMR на высоте блока 39,859. На высоте гораздо более позднего блока наблюдатель, Ив, видит три транзакции на сумму 10, 20 и 30 XMR, соответственно. Ив видит, что в транзакции на сумму 10 XMR используются миксины из блока 39,859 и что ситуация с транзакциями на сумму 20 и 30 XMR та же. Проверив все возможные комбинации, Ив приходит к выводу, что кто-то, получивший 60 XMR на блоке с высотой 39,859, наиболее вероятно, и тратит эти средства, несмотря на то, что Синтия и Ив при этом могут не знать друг друга.

Эта информация может использоваться в тандеме с другими способами атаки. В частности, *если*

ряд отдельных транзакций имеет какие-либо общие данные (такие как общая высота блока) *и если все транзакции в этом ряду используют миксины с какими-то другими общими данными, значит, есть вероятность, что эти транзакции связаны между собой одним общим пользователем.* Это даёт злоумышленникам некоторую информацию, связанную с отслеживаемостью. Предлагаемое нами определение комбинаторной атаки кажется невероятно общим и, по сути, включает в себя атаку с выявлением связи по использованию, о которой говорилось в подразделе 3.2. Комбинаторный анализ ряда транзакций, обладающих некоторыми общими данными, согласно этому принципу, позволяет прийти лишь к некоторым выводам. То есть, несмотря на то, что такая *комбинаторная атака* в большинстве случаев может не сработать, когда её всё же удаётся провести, она работает довольно хорошо.

4 Наши рекомендации

В данном разделе нами приводятся рекомендации, предназначенные для Команды ведущих разработчиков и касающиеся разработки Monero. Наши рекомендации, несмотря на отсутствие в них совершенного решения вышеуказанных проблем, позволяют пользователям избежать атак, связанных с отслеживаемостью транзакций. Мы рекомендуем любому и всем разработчикам протокола CryptoNote ввести по крайней мере общесетевое требование к использованию минимального количества миксинов, равного $n \geq 3$, и чем больше, тем лучше.

4.1 Требование к использованию минимального количества миксинов

В этом подразделе мы пытаемся избежать явления цепной реакции и рекомендуем изменить протокол Monero путём введения *требования* к использованию минимального количества миксинов, равного $n \geq 2$, на транзакцию при наличии по крайней мере 2 сторонних выходов, доступных для смешивания, а также установки обязательного требования к использованию миксинов в количестве $n \geq 4$ по умолчанию на уровне программного обеспечения кошельков. Нами также приводится обоснование нашей рекомендации.

В частности, нами рекомендуется такое изменение протокола Monero, при котором узел будет отказывать в ретрансляции транзакции с кольцевой подписью, не содержащей обязательного количества миксинов, если только такая транзакция не будет деноминирована таким образом, что попадёт в набор выходов транзакций без достаточного количества выходов. Нами также рекомендуется изменить протокол Monero так, чтобы блок отклонялся при наличии в нём каких-либо транзакций с кольцевой подписью, не содержащей обязательного количества миксинов.

Безусловно, любое обязательное минимальное количество миксинов, равное $n \geq 2$, даёт вероятностную гарантию предотвращения любой цепной реакции, приводящей к отслеживаемости. В частности, злоумышленнику придётся потратить больше средств, чем будет им раскрыто, что было показано в работе [3]. Единственным вопросом остаётся, какое значение n следует сделать обязательным? Это горячо обсуждалось членами Исследовательской лаборатории Monero и Командой ведущих разработчиков. Более высокие значения n обеспечат более высокий уровень безопасности, но за счёт неоправданного разрастания блокчейна и роста размера комиссий. Как уже было упомянуто, нами не используются методы NIZK, и никакой метод создания кольцевой подписи не позволяет достичь соответствующего уровня безопасности, поэтому любой размер блокчейна, превышающий размер блокчейна NIZK (например, Zerocoin) будет неприемлемым. С другой стороны, более низкие значения n позволят сократить место, занимаемое блокчейном, но, возможно, в ущерб безопасности. К счастью, размер кольцевой подписи, выраженный в байтах, возрастает практически линейно относительно количества сторонних

ключей, используемых в кольцевой подписи, и увеличение значения n равномерно улучшает устойчивость системы к атакам, вызывающим цепную реакцию, что было описано в работе [3]. Такое повышение устойчивости, однако, не является линейным, и оптимальное решение отсутствует. С точки зрения безопасности значение $n = 3$ лучше, чем $n = 2$, $n = 4$ лучше, чем $n = 3$, и так далее.

Путём моделирования методом Монте-Карло [5] мы пришли к выводу, что отдельные пользователи могут защитить себя от атак с применением цепной реакции, используя по крайней мере $n = 3$ миксинов в своих кольцевых подписях. Количественно измеренные метрики безопасности могли бы подтвердить эти результаты по крайней мере двумя путями. Первая метрика, использованная нами для подтверждения этих результатов, требовала, чтобы успешность атаки с использованием цепной реакции зависела от размера набора выходов (предполагается выбор значения $n \geq 3$). Действительно, первая метрика позволяет нам увидеть, что, независимо от размера набора выходов, злоумышленник, контролирующий 50% набора выходов, с 95% вероятностью будут контролировать одну из десяти новых транзакций, если значение $n = 2$, но если $n \geq 3$, успех злоумышленника будет значительно зависеть от размера набора выходов. Вторая метрика, использованная нами, требовала, чтобы злоумышленнику, контролирующему 50% набора выходов произвольного размера, пришлось подождать десятки тысяч транзакций перед тем, как он мог бы достичь 95% вероятности получения контроля над одной новой транзакцией (предполагается, что было выбрано значение $n \geq 4$).

Однако действительно и то, что при значении $n = 2$ цепная реакция вероятностно гарантированно исчерпает себя, что говорит о постепенном снижении целостности системы, а также заставляет злоумышленника тратить больше средств, чем им будет раскрыто. После всестороннего обсуждения нами было принято решение установить по умолчанию значение $n = 4$ на уровне программного обеспечения кошельков, но позволить тем пользователям, которых более собственной анонимности заботит размер комиссии, снизить количество используемых миксинов до значения $n = 2$, что кажется благоразумным. Как упоминалось ранее, команде ведущих разработчиков Monero и Исследовательской лаборатории Monero было бы хорошо следовать той философии разработки, которая предполагает внесение небольших модификаций с их последующим развитием, а не реализацию глобальных изменений с дальнейшей попыткой их обратного масштабирования. Мы также официально рекомендуем изменить протокол таким образом, чтобы после того как блокчейн достаточно вырастет, обязательное минимальное количество миксинов выросло до значения $n = 4$. Спустя 2 года протокол Monero автоматически станет более безопасным.

В качестве отступления хотелось бы отметить, что это не единственный способ решить проблему цепной реакции. В качестве альтернативного решения можно разрешить проводить транзакции в открытую, но с использованием тегов. Такое решение предполагает маркировку некоторых выходов транзакций как безопасных с точки зрения смешивания, а некоторых как опасных для смешивания; эквивалентно это предусматривает обрезание наборов миксинов. Перед валидацией блока майнер мог бы по тегам проверять, чтобы в блоке ни одна кольцевая подпись в качестве входа не использовала выхода транзакции с нулевым смешиванием.

Тем не менее такие решения, основанные на использовании тегов, также имеют некоторые недостатки. Например, несмотря на то, что отдельно взятый блок может и не содержать кольцевых подписей, использующих в качестве входа выход транзакции с нулевым смешиванием, из-за цепной реакции, описанной в работе [3], блок всё же может включать в себя отслеживаемые транзакции. На самом деле транзакции, присутствующие в блоке, могут быть отделены от опасных «открытых» транзакций несколькими другими транзакциями. Кроме того, в случае с основанным на применении тегов решением, допускающим наличие транзакций с нулевым смешиванием, пользователи всё же могут отказаться от преимуществ, обеспечиваемых кольцевыми подписями, связанных с анонимностью. Было

бы предпочтительней найти решение, которое поддерживало бы сохранение анонимности как единую политику на уровне сети, что позволило бы пользователям отдавать предпочтение более высокому уровню анонимности, соблюдая при этом местные законы, о чём говорится в Разделе 6.

4.2 Как избежать комбинационных атак

Для Боба самым простым способом скрыть то, что он является владельцем средств, при перехвате информации Ив во время проведения ею комбинаторной атаки была бы периодическая отправка своих выходов себе же самому со случайными интервалами времени между такими отправками. Это запутывает анализ блокчейна, проводимый Ив, о котором говорилось в подразделе 3.3 и фактически в подразделе 3.1. В этом подразделе нам просто хотелось бы указать на то, что ни один пользователь не должен отправлять все свои выходы самому себе одновременно. Более того, любой получатель средств, напротив, должен попросить отправителя разбить транзакцию на части, как в случае с торрентом, и отправить необходимую сумму через какое-то время. Таким образом, если Ив будет ожидать появления транзакции с определённой суммой в каком-то определённом временном окне, она уже не сможет так просто определить, соответствует ли некоторая комбинация выходов из всех транзакций, входящих в блок, точной сумме, которая, как она ожидает, будет отправлена получателю.

Итеративное отправление транзакций самому себе с некоторыми случайными интервалами времени между отправками и разбиение всех транзакций (включая пересылаемые себе транзакции) на множество малых транзакций, также отправляемых через разные промежутки времени, значительно затрудняют перехват информации в блокчейне, в основе которого лежат исключительно данные высоты блока. Следует отметить, что эта рекомендация может быть реализована на уровне кошелька, но не на уровне протокола.

4.3 Как избежать ассоциации по использованию

Чтобы избежать атаки путём так называемой «ассоциации по использованию», можно прибегнуть к альтернативной схеме. Вместо того чтобы создавать одну транзакцию с несколькими непотраченными выходами, мы можем сгенерировать по одной отдельной транзакции для каждого выхода. Таким образом, если Боб хочет отправить Элис 0.75 XMR из своего кошелька, где у него имеется 1 XMR, то ему по факту будет необходимо провести две транзакции. В первой транзакции будет отправлено 0.7 XMR , а во второй - 0.05 XMR. Прежде всего, до отправки 0.7 XMR из кошелька, в котором есть 1 XMR, необходимо применить обычный метод деноминации и разбиения сдачи, как показано в таблице 2. Затем, получив свою сдачу, Боб завершит транзакцию. После проведения первой транзакции Бобом у него останется резервный выход на 0.09 XMR, в отношении которого он также может применить обычный метод деноминации и разбиения сдачи и отправить 0.05 XMR Элис, как показано в таблице 3.

Таблица 2: Транзакция 1

(a) Транзакция $1A$, проведённая Бобом

Вход	Выход
1.0 XMR (Боб)	0.7 XMR (Элис)
	0.2 XMR (Боб)
	0.09 XMR (Боб)
	0.01 XMR (комиссия)

(b) Транзакция $1B$, проведённая Элис

Вход	Выход
0.7 XMR (Элис)	$\cdots$

Таблица 3: Транзакция 2

(a) Транзакция $2A$, проведённая Бобом

Вход	Выход
0.09 XMR (Боб)	0.05 XMR (Элис)
	0.03 XMR (Боб)
	0.01 XMR (комиссия)

(b) Транзакция $2B$, проведённая Элис

Вход	Выход
0.05 XMR (Элис)	$\cdots$

При использовании вышеуказанной схемы Боб генерирует только один дополнительный выход, но и это позволяет защитить Элис от раскрытия в будущем путём ассоциации выходов, входящих в кольцевые подписи. Комиссии за майнинг повышаются в зависимости от количества полученных выходов транзакций, но при этом сохраняется анонимность.

Тем не менее следует отметить, что в этой альтернативной схеме есть одна тонкость. Если Боб использует свой собственный вход, чтобы отправить 0.05 XMR в транзакции $2A$, это даст возможность провести тот же обстоятельственный анализ подобия, которого мы пытались избежать в первую очередь. Выходы могут быть связаны как потраченные либо из одной и той же транзакции или общего *дерева* транзакций, либо из общего блока или же в течение короткого промежутка времени. Этот список может быть и не полным. Безусловно, такие сценарии всё в большей степени затрудняют проведение анализа Ив, если сравнивать их с простым сценарием, описанным в подразделе 3.2.

Два выхода, взятые из непересекающихся транзакций с двумя разными высотами блока, наименее вероятно будут принадлежать одному владельцу, чем два выхода из одной и той же транзакции или чем два выхода из двух транзакций, связанных одним этапом, или чем два выхода с одной высотой блока. Поэтому, поскольку Боб также может использовать любой другой имеющийся у него выход, чтобы отправить 0.05, а не только его выход сдачи, по возможности он так и должен сделать. Это приводит нас к естественному решению, которое несложно реализовать в простой форме: для каждой транзакции следует использовать всего один выход и отправлять получателю только один выход. И этот процесс должен повторяться до тех пор, пока получателю не будут отправлены все средства.

Безусловно, это неэффективный метод генерирования выходов транзакций. Чтобы повысить эффективность для создания транзакции также можно использовать множество входов, но только если соответствующие выходы не будут находиться в пределах нескольких блоков друг от друга.

5 План реализации наших рекомендаций

Реализацию требования к использованию минимального количества миксинов $n = 2$ на общесетевом уровне в рамках архитектуры Monero можно произвести несколькими способами.

(1) Можно потребовать от пользователей при использовании кошелька отправлять средства через RPC между кошельками с минимальным количеством миксинов, если только проводимые транзакции не попадают в набор выходов транзакций без достаточного количества миксинов.

(2) Майнеров можно попросить или потребовать включать в блоки только транзакции с минимальным количество миксинов, если только такие транзакции не попадают в набор выходов транзакций без достаточного количества миксинов.

(3) Одноранговый протокол можно изменить таким образом, чтобы транзакции, не отвечающие требованию к использованию минимального количества миксинов во всех своих кольцевых подписях, не

ретранслировались, если только такие транзакции не будут попадать в набор выходов транзакций без достаточного количества миксинов.

Кроме того, если транзакция будет включена в блок, высота которого будет выше определённого критического значения, то количество миксинов $n = 2$ не будет достаточным, и, по сути, оно должно будет составлять $n = 4$. Мы рекомендуем обеспечить соответствие такой высоте блока через два или три года после публикации этого документа, но тут важны детали. Предлагаемые нами рекомендации по использованию минимального количества миксинов на общесетевом уровне следует реализовать (1) скоро, пропустить пункт (2) и запланировать выполнение пункта (3) на более позднюю дату после тщательного тестирования в тестовой сети. Основным препятствием для немедленной реализации является *пыль*, передаваемая при проведении транзакций. Чтобы преодолеть его, нами определяется форма *легального выхода транзакции* в качестве любого числа $A \times 10^B$, где A является целым числом $1 \leq A \leq 9$ а B является целым числом $-12 \leq B$. Таким образом, выход транзакции на сумму 0.00111111 XMR не соответствует форме легального выхода транзакции, а выходы $0.001, 0.0001, 0.00001, 0.000001, 0.0000001$ и 0.00000001 XMR являются легальными. Нами предлагается следующая политика:

(i) Потребовать от пользователей, чтобы при использовании кошельков ими генерировались только те выходы транзакций, которые бы соответствовали форме легального выхода транзакции;

(ii) Майнеров можно попросить или потребовать включать в блоки только те транзакции, в которых все выходы соответствуют форме легального выхода транзакции;

(iii) Одноранговый протокол можно изменить таким образом, чтобы транзакции, в которых все выходы не соответствуют форме легального выхода транзакции, не ретранслировались.

Реализация этих рекомендаций позволит пользователю использовать нелегальный выход транзакции в качестве входа при условии, что все выходы такой транзакции будут соответствовать этой строгой форме. Таким образом, в системе более не будет создаваться пыли, а уже имеющаяся в ней пыль пропадёт с течением времени. Опять же, наш план реализации формы легального выхода транзакции состоит в выполнении пункта (*i*) в первую очередь, пропуске пункта (*ii*) и выполнении пункта (*iii*) позднее после тщательного тестирования.

В целях мотивации пользователей к итеративному проведению транзакций самим себе с временными интервалами различной длины нами рекомендуется, чтобы на уровне кошелька выводилось предупреждение, если в блокчейн было добавлено случайное, но минимальное количество блоков с тех пор, как кошельком был получен соответствующий выход. Спустя 6 или 8 недель, например, все выходы в кошельке будет необходимо переслать.

Кроме того, в программном обеспечении кошелька должно быть заложено следующее:

(a) одновременно получателю можно отправить только один выход транзакции;

(b) транзакция, включающая в себя несколько выходов, должна быть разбита на несколько меньших транзакций, передаваемых в течение какого-то времени, подобно тому как это происходит в случае с торрентом;

(c) каждый выход транзакции, принятый получателем, должен использовать входы транзакций с различной высотой блока и из разных корневых транзакций.

6 Проведение аудита в целях соблюдения законности

Коммерсантам, использующим кольцевые подписи и скрытые адреса Monero, по вполне понятным причинам хочется соблюдать соответствующие местные законы, действующие в их юрисдикциях. Как правило, такие законы требуют, чтобы коммерсанты были в состоянии подтвердить принадлежность собственных средств, а также продемонстрировать, куда были переведены эти средства. Чтобы подтвердить обладание непотраченными средствами в блокчейне, владелец счёта Monero может опубликовать кольцевую подпись к своему публичному ключу с нулевым количеством миксинов. Следует помнить о том, что кольцевая подпись, не содержащая миксинов, является обычной цифровой подписью. После этого аудитор сможет просканировать блокчейн на предмет наличия в нём публичного ключа к выходу и убедиться в том, что образ ключа, созданный на основе кольцевой подписи с нулевым смешиванием, ещё не использовался в блокчейне. Чтобы защитить анонимность предприятия и клиентов, после проверки аудитор должен уничтожить подпись и образ ключа.

7 Заключение

Нами были продемонстрированы несколько способов, используя которые злоумышленник может провести анализ блокчейна с целью снижения уровня неотслеживаемости, обеспечиваемого протоколом CryptoNote 2. Проблема трат с нулевым смешиванием, описанная в работе [3], проблема ассоциации выходов по возрасту, проблема ассоциации по использованию, а также более общая проблема комбинаторного анализа — все эти проблемы были обозначены нами. Для решения проблемы цепной реакции при нулевом смешивании мы рекомендуем ввести обязательно минимальное количество миксинов $n = 2$ на общесетевом уровне, а также нами обосновывается, почему количество миксинов, равное $n = 4$, вероятнее всего, будет слишком большим с точки зрения использования в практических целях. Мы рекомендуем применять метод неравномерного выбора выходов транзакций при создании подписи, чтобы избежать проблем, связанных со связыванием по возрасту выходов. Также мы рекомендуем разбивать все транзакции на множество малых транзакций таким образом, чтобы они проводились через какое-то время, как в случае с торрентом, и так, чтобы каждая субтранзакция включала в себя единственный выход, так как это позволит избежать проблем, связанных с проведением комбинаторного анализа.

Список литературы

[1] Norman Lloyd Johnson and Samuel Kotz. *Urn models and their application: an approach to modern discrete probability theory (Модели урн и их применение: подход к современной теории дискретной вероятности)*. Wiley New York, 1977.

[2] Ian Miers, Christina Garman, Matthew Green, and Aviel D Rubin. Zerocoin: Anonymous distributed e-cash from bitcoin (zerocoin: анонимная распределённая электронная валюта на базе bitcoin). In *Security and Privacy (SP), 2013 IEEE Symposium on*, pages 397–411. IEEE, 2013.

[3] Surae Noether, Sarang Noether, and Adam Mackenzie. Mrl-0001: A note on chain reactions in traceability in cryptonote 2.0 (mrl-0001: Возможная цепная реакция, связанная с отслеживаемостью, в рамках протокола cryptonote 2.0). Technical report.

[4] Ronald L Rivest, Adi Shamir, and Yael Tauman. How to leak a secret (Как узнать секрет). In *Advances in Cryptology-ASIACRYPT 2001*, pages 552–565. Springer, 2001.

[5] Tacotime. Saturation of unspent transaction outputs (Накопление непотраченных выходов транзакций), September 2014.