



Collaborative Deanonymization

Patrik Keller¹(✉), Martin Florian², and Rainer Böhme¹

¹ University of Innsbruck, Innsbruck, Austria
{patrik.keller,rainer.boehme}@uibk.ac.at

² Weizenbaum Institute, Humboldt University of Berlin, Berlin, Germany
martin.florian@hu-berlin.de

Abstract. Privacy-seeking cryptocurrency users rely on anonymization techniques like CoinJoin and ring transactions. By using such technologies benign users potentially provide anonymity to bad actors. We propose overlay protocols to resolve the tension between anonymity and accountability in a peer-to-peer manner. Cryptocurrencies can adopt this approach to enable prosecution of publicly recognized crimes. We illustrate how the protocols could apply to Monero rings and CoinJoin transactions in Bitcoin.

1 Introduction

“Anonymity loves company.” [6] It is well-established that anonymity is co-created by the members of an anonymity set, who share the same intention and employ technical systems and protocols to make them appear indistinguishable to outside observers [16]. Inherently, benign members seeking privacy assist bad actors avoiding law enforcement.

Previously, the tension between privacy and law enforcement has been studied for mixes in communication networks [4, 5, 11]. The proposed solutions rely on putting backdoors into systems or the supporting cryptography, such that designated parties can revoke the anonymity in justified cases. Access to the backdoor is made transparent, which holds law enforcement accountable and impedes mass surveillance. With the advent of privacy-hardened cryptocurrencies, the tension is instantiated for money flows. While backdoors seem technically feasible, it is unlikely that they can be sustained in decentralized systems, whose *raison d’être* is the rejection of privileged parties with special access rights.

Another, more widely acceptable idea to combat money laundering specifically are threshold schemes. Small payments would enjoy unlinkability while larger transactions require identification or are traceable by design [10, 19]. The downsides of this approach include the need to agree on a threshold and, more importantly, it would require strong identities in order to prevent “smurfing” attacks, which split a large sum into many small payments.

We explore a different approach. In many cases, the parties forming the anonymity set can retain some private information, which can help deanonymize other members of the set. Collaborative deanonymization means that some parties, henceforth called *witnesses*, share information on request for the purpose

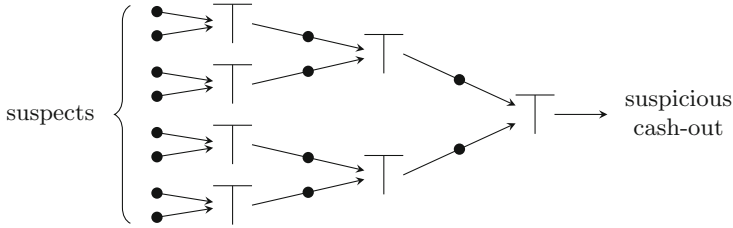


Fig. 1. Example entity graph of 7 ring-type transactions, $m = 2$. Dots are entities, arrows denote possible payments. Observe the exponential growth of suspects.

of solving a crime. In a nutshell, law enforcement publicly shares information requests for specific crimes. Then users check whether they are involved, decide whether the crime should be prosecuted, and potentially reveal private information to support deanonymization.

We argue that this approach is compatible with the peer-to-peer spirit of decentralized systems because every witness decides if she supports the investigation or not. This limits the method to felonies that are universally disapproved, such as extortion (ransomware) or the financing of child sexual abuse. For the method to be effective, it is not required that every witness collaborates. Every collaborating witness reduces the search space. Law enforcement might leverage a range of incentives to induce collaboration: alibi, altruism, bounties, and—in justified cases—force (e.g., seizure and use of a private key). Unlike traffic or blockchain analyses, collaborative deanonymization does not scale, hence the risk of secret mass surveillance is small. Moreover, as search requests are announced publicly, law enforcement can be held accountable. The very fact that anonymity is conditional can deter crime.

In the following we develop a scenario (Sect. 2), formulate desiderata, and sketch overlay protocols (Sects. 3 and 4) that enable collaborative deanonymization of two relevant privacy techniques, CoinJoin and Monero rings, without changing the target systems. Section 5 concludes.

2 Scenario and Model

Consider a scenario where a law enforcement agency (LEA) identifies a suspicious cash-out from a cryptocurrency address. The objective of an investigation is to find an identifiable source, i.e., *backtracking*. After employing known blockchain analysis methods, like state-of-the-art clustering [8], the LEA obtains an entity graph where backtracking is ambiguous only due to mixing transactions.

We model such transactions as collections of m inputs and n outputs. The LEA has no information about the relation.¹ Without loss of generality, we

¹ Conversely, if the LEA has some information (e.g. due to non-uniformly valued inputs and outputs), it can partition the transaction and proceed as described.

assume that each output of a transaction is funded by exactly one input. Backtracking links the entity associated with the targeted t -th output to the entity of the funding input. Between transactions, each input references exactly one output of a previous transaction.

general	join-type	ring-type
$\begin{array}{ c } \hline \\ \hline m \\ \hline \\ \hline \end{array}$	$\begin{array}{ c } \hline 1 \\ \hline 2 \\ \hline \vdots \\ \hline m \\ \hline \end{array}$	$\begin{array}{ c } \hline 1 \\ \hline 2 \\ \hline \vdots \\ \hline m \\ \hline \end{array}$
$\begin{array}{ c } \hline \\ \hline n \\ \hline \\ \hline \end{array}$	$\begin{array}{ c } \hline \psi(1) \\ \hline \psi(2) \\ \hline \vdots \\ \hline \psi(m) \\ \hline \end{array}$	$\begin{array}{ c } \hline \sigma \\ \hline \\ \hline \\ \hline \end{array}$

We consider two of the most relevant types of mixing transactions: join-type as used in CoinJoin [12] and ring-type as used in Monero [7]. Join-type transactions are formed collaboratively by m parties, potentially facilitated by an intermediary such as JoinMarket [13]. We model this using m inputs, each funding a distinct output ($n = m$). A join-type transaction can then be expressed as a permutation ψ on $\{1, \dots, m\}$. The LEA’s problem is to find the funding input $\psi(t)$ of the t -th output. In practice, CoinJoin transactions vary in size. A study estimates the modal value of inputs for CoinJoins on Bitcoin at $m = 3$ [13]. Transactions with $m > 10$ are rare.²

In contrast to join-type transactions, ring-type transactions can be formed *without* the cooperation of other entities. Moreover, a ring-type transaction does not spend all outputs referenced on its input side. In our simplified model, ring transactions have m inputs and a single output ($n = t = 1$). The LEA’s goal is to learn the true input σ .³ At the time of writing, the Monero reference implementation fixes the number of inputs to $m = 11$.

For both types of mixing transactions, the anonymity of the participants is based on the observer’s uncertainty about ψ and σ , respectively. If multiple mixing transactions are cascaded, the number of possible funding sources (*suspects*) increases exponentially in the number of layers (see Fig. 1). We propose protocols that allow the LEA to reduce the number of suspects in collaboration with a subset of the involved parties.

3 Collaborative Backtracking

We assume an authenticated one-way communication channel from the LEA to the protocol participants. The LEA uses this channel to announce inquiries on targeted transaction outputs. Each inquiry conveys enough information so that a potential witness can decide whether she supports the request, i. e., whether she approves prosecution of the specific case, or not.

We further assume an unauthenticated but confidential communication channel from the witnesses to the LEA and, for group testimonies, communication

² A CoinJoin with $m = 100$ made headlines in June 2019: <https://www.coindesk.com/bitcoin-users-perform-what-might-be-the-largest-coinjoin-ever>.

³ We depart from Monero’s terminology, which calls an entire ring “input.”

channels between the witnesses. Witnesses willing to support an inquiry use these channels to give testimonies that facilitate backtracking for a single transaction.

3.1 Individual Testimony

An individual testimony is a protocol between a single witness and the LEA. It results in ruling out one of the possible inputs. Formally speaking, the witness associated with the i -th input should prove that $\psi(t) \neq i$ or $\sigma \neq i$, respectively.

For join-type transactions, the witness can testify by signing a challenge with the private keys belonging to the i -th input and the j -th output (obviously $t \neq j$).

Ring-type transactions hide the true input using traceable ring signatures [7]. By design, these ring signatures reveal attempts to spend an input more than once. The spending of an input yields a transaction-independent *key image* that must be included in a valid signature—transactions attempting to spend the same input will contain identical key images [17]. Let o be the output of a *preceding* transaction that links the witness to the suspicious transaction T . The witness prepares a phantom transaction T' for the LEA. It has one input referencing o and one output. The output could be invalid in order to avoid accidental inclusion in the blockchain. For example, T' could spend more funds than available in o . Crucially, the phantom transaction unambiguously spends o . If the key image associated with T' is different to the key image of T , it must hold that $i \neq \sigma$.

3.2 Group Testimony

The LEA is interested in a single input to output relationship, but it learns one relationship per individual testimony. Group testimonies can avoid this unnecessary privacy loss. Multiple witnesses controlling the set of inputs S collaboratively testify $\psi(t) \notin S$ or $\sigma \notin S$, while maintaining their anonymity within S .

For join-type transactions, this can be realized by signing a challenge with all $2 \cdot |S|$ private keys belonging to the witnesses' inputs and outputs. In the best case, all $m - 1$ witnesses cooperate ($S = \{1, \dots, m\} \setminus \{t\}$) and identify the true suspect. If $|S| < m$ witnesses participate in the protocol, for example because private keys are deleted or witnesses unreachable, the search space is reduced to $m - |S|$ suspects. Join-type group testimonies retain S as the anonymity set of witnesses. Cases where $S = \{1, \dots, m\} \setminus \{t\}$ minimize the anonymity loss for witnesses when testifying that $\psi(t) \notin S$.

For ring-type transactions, it is possible to implement group testimonies with the construction of a provably spent set [18, 20]. For example, each cooperating witness can individually form a new transaction T' like for an individual testimony, however this time referencing not only its own input but all inputs S of cooperating witnesses. Given $|S|$ transactions that all have the same set of inputs S and yet differing key images, the LEA gains evidence that $\sigma \notin S$. If an output o referenced by an input $i \in S$ is *unspent* at the time of the testimony, the respective witness can achieve an anonymity set of S for o by referencing

all S when spending o . Conversely, if o has already been spent in a transaction T'' with input set S'' , the anonymity set of the witness reduces to $S \cap S''$.

Notably, each of the cooperative protocols can be executed jointly for multiple mixing transactions. This testifies that the owners of S (now generalized to the enumeration of all inputs in all transactions involved) initiated none of these transactions. This approach is especially interesting for ring-type transactions, as larger S increase the overlap with the anonymity sets of outputs that have already been spent elsewhere.

3.3 Dealing with the Risk of False Testimonies

A general question is how much confidence the LEA can place in the testimonies. This calls for a closer look at how collaborative deanonymization can fail, and in the worst case produce false or misleading evidence. We observe crucial differences between join-type and ring-type transactions.

Monero stores σ on the blockchain, however in encrypted form. This should reduce the risk of false testimonies to the security of the cryptography used, even if private keys are leaked or stolen.

By contrast, CoinJoin does not commit ψ to the blockchain. Even computationally unbounded observers cannot decide about the relation. The resulting deniability bears a risk of false testimonies. For example, if the perpetrator has access to the private keys of a witness, he could obtain a false alibi by signing a false input–output relation. If the victim among the witnesses does not participate in the collaborative deanonymization, she is falsely accused. If she does participate, the LEA receives two conflicting statements. This concentrates the suspicion on both the perpetrator and the victim, hence perpetrators have little to gain from false statements—unless their victims are unavailable.

The sketched situation highlights that parties engaging in CoinJoins might be exposed to physical risks under collaborative deanonymization. A potential direction of research is to modify the protocols used for CoinJoin formation in such a way that ψ is committed to the blockchain at the time of the transaction. This would obviate false accusations and reduce the incentives to attack other witnesses. The key question to answer is under which conditions what part of ψ should be revealable. For example, should every party commit to one relation individually? Would a threshold scheme make sense? Moreover, it would be desirable to make the commitment coercion-resistant. Otherwise, the risk could reappear at the time of the CoinJoin formation, rather than be mitigated.

Another approach for increasing the credibility of testimonies could be based on witnesses proving that addresses belong to the same wallet, e.g., if their wallet generates addresses deterministically from a common secret. It is an open question how such a proof can efficiently be completed without revealing more information about the wallet than necessary.

4 Forward Tracking

A variant of the scenario presented in Sect. 2 is *forward tracking*. Here, the LEA has identified a suspicious origin and wishes to trace the money flow to its (current) destination or until it hits a known cash-out point. We sketch how our approach can be adapted to this case.

4.1 Testimonies for Forward Tracking

Due to the symmetry of join-type transactions, the backtracking protocols (Sect. 3) can be repurposed for forward tracking. Since ψ is bijective, testimonies which rule out assignments of ψ also rule out assignments of ψ^{-1} .

Ring-type transactions are less straightforward. The protocols given in Sect. 3 enable collaborating witnesses to testify that a set of inputs S does not contain the funding input for a given transaction T , i. e., $\sigma \notin S$. For the case of forward tracking, they must instead prove that only one specific suspicious input s is not a funding input, i. e., $\sigma \neq s$. Individual witnesses can accomplish this by creating a phantom transaction T' , which include all but the suspicious input s . As T' and T share the same funding input i , they will produce identical key images. By comparing the key images of T and T' , the LEA can verify that $\sigma \neq s$ without learning i .

4.2 Blacklisting and Cover Transactions

Forward tracking is related to transaction *blacklisting* previously proposed (and controversially debated) as a regulatory instrument [2, 15]. Specifically the “poison” policy [14], where taint of a single input is propagated to all outputs, mimics the proliferation of a priori suspicion. An interesting question is whether the threat of blacklisting can foster collaboration. For example, the propagation policy could terminate at whitelisted transactions after sufficient evidence has been collected to disambiguate the entity graph (for forward and backtracking).

Forward tracking on Monero rings comes with two caveats. First, it might be hard to decide about when to terminate (unsuccessfully), because it is often unknown whether a given output has been spent at all. Second, the method is susceptible to cover transactions placed by a perpetrator. Such transactions reference the investigated money flow in order to increase the search space and with it the number of witnesses needed.

Blacklisting might be a defense against this behavior because it would devalue the funds in cover transactions and thus raise the cost of creating them. However, the effectiveness of this method as well as other defenses are open research questions. We note that backtracking is not affected by the threat of cover transactions because funding transactions cannot be added after the spending transaction.

5 Conclusion and Outlook

We have outlined a novel way to investigate criminal money flows in cryptocurrencies even if the perpetrators use anonymization techniques. Our approach

requires collaboration of witnesses, which keeps the method costly enough to prevent mass surveillance or the prosecution of petty crimes. Specifically, we have given protocols for backtracking and forward tracking of CoinJoin transactions in Bitcoin as well as Monero rings. Several techniques ensure that the information shared with law enforcement can be limited to the necessary minimum. The new risk of false accusations has been discussed. A general consequence of collaborative deanonymization is that old private keys remain sensitive even if they do not control any funds anymore.

We shall also pinpoint future work. Obviously, the protocols for secure testimonies need to be further developed and their properties formalized and proven. A proof-of-concept implementation for the most relevant types of mixing transactions could demonstrate the practicality of our approach. Whether and under which condition LEAs can deploy collaborative deanonymization, must be subject of more interdisciplinary work with legal scholars. Adapting the approach to less common types of mixing transactions (see for instance Table 1 of [9] for an overview) would help to complete the picture.

The topic also lends itself to economic studies. One could investigate the incentives of witnesses to collaborate, presumably with cooperative game theory [3]. In addition, potential knock-on effects on the participation in mixing transactions call for a model in the tradition of competitive game theory [1].

Two broader technical directions are to explore collaborative deanonymization for anonymous communication systems, and to research deniable privacy techniques, which could protect potential witnesses from any pressure to testify or release deanonymizing information.

In summary, collaborative deanonymization appears not only under-researched, but also under-estimated for its potential to balance the conflicting goals of privacy and law enforcement in future digital currency systems. This short paper sets out to make a case for this promising tool.

Acknowledgements. We thank our colleagues Michael Fröwis, Malte Möser, Tim Ruffing, and a number of anonymous reviewers for helpful discussions of earlier versions of this work. Rainer Böhme's and Patrik Keller's work on this topic is supported by the Austrian FFG's KIRAS programme under project VIRT-CRIME.

References

1. Abramova, S., Schöttle, P., Böhme, R.: Mixing coins of different quality: a game-theoretic approach. In: Brenner, M., et al. (eds.) FC 2017. LNCS, vol. 10323, pp. 280–297. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70278-0_18
2. Anderson, R., Shumailov, I., Ahmed, M.: Making bitcoin legal. In: Matyáš, V., Švenda, P., Stajano, F., Christianson, B., Anderson, J. (eds.) Security Protocols 2018. LNCS, vol. 11286, pp. 243–253. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03251-7_29
3. Arce, D.G., Böhme, R.: Pricing anonymity. In: Meiklejohn, S., Sako, K. (eds.) FC 2018. LNCS, vol. 10957, pp. 349–368. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-662-58387-6_19

4. Backes, M., Clark, J., Kate, A., Simeonovski, M., Druschel, P.: BackRef: accountability in anonymous communication networks. In: Boureau, I., Owesarski, P., Vaudenay, S. (eds.) ACNS 2014. LNCS, vol. 8479, pp. 380–400. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07536-5_23
5. Claessens, J., Díaz, C., Goemans, C., Dumortier, J., Preneel, B., Vandewalle, J.: Revocable anonymous access to the Internet? *Internet Res.* **13**(4), 242–258 (2003)
6. Dingledine, R., Mathewson, N.: Anonymity loves company: usability and the network effect. In: Workshop on the Economics of Information Security (2006)
7. Fujisaki, E., Suzuki, K.: Traceable ring signature. In: Okamoto, T., Wang, X. (eds.) PKC 2007. LNCS, vol. 4450, pp. 181–200. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-71677-8_13
8. Goldfeder, S., Kalodner, H., Reisman, D., Narayanan, A.: When the cookie meets the blockchain: privacy risks of web payments via cryptocurrencies. *Privacy Enhancing Technol.* **4**, 179–199 (2018)
9. Heilman, E., AlShenibr, L., Baldimtsi, F., Scafuro, A., Goldberg, S.: TumbleBit: an untrusted bitcoin-compatible anonymous payment hub. In: Network and Distributed System Security Symposium. Internet Society (2017)
10. Jarecki, S., Shmatikov, V.: Probabilistic escrow of financial transactions with cumulative threshold disclosure. In: Patrick, A.S., Yung, M. (eds.) FC 2005. LNCS, vol. 3570, pp. 172–187. Springer, Heidelberg (2005). https://doi.org/10.1007/11507840_17
11. Köpsell, S., Wendolsky, R., Federrath, H.: Revocable anonymity. In: Müller, G. (ed.) ETRICS 2006. LNCS, vol. 3995, pp. 206–220. Springer, Heidelberg (2006). https://doi.org/10.1007/11766155_15
12. Maxwell, G.: CoinJoin: Bitcoin privacy for the real world. Forum post (2013)
13. Möser, M., Böhme, R.: Join me on a market for anonymity. In: Workshop on the Economics of Information Security (2016)
14. Möser, M., Böhme, R., Breuker, D.: Towards risk scoring of bitcoin transactions. In: Böhme, R., Brenner, M., Moore, T., Smith, M. (eds.) FC 2014. LNCS, vol. 8438, pp. 16–32. Springer, Heidelberg (2014). https://doi.org/10.1007/978-3-662-44774-1_2
15. Möser, M., Narayanan, A.: Effective cryptocurrency regulation through blacklisting. Preprint (2019)
16. Pfitzmann, A., Köhntopp, M.: Anonymity, unobservability, and pseudonymity — a proposal for terminology. In: Federrath, H. (ed.) Designing Privacy Enhancing Technologies. LNCS, vol. 2009, pp. 1–9. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-44702-4_1
17. van Saberhagen, N.: CryptoNote v2.0. Whitepaper (2013)
18. Wijaya, D.A., Liu, J., Steinfeld, R., Liu, D.: Monero ring attack: recreating zero mix-in transaction effect. In: Trust, Security And Privacy In Computing And Communications, pp. 1196–1201. IEEE (2018)
19. Wüst, K., Kostianen, K., Čapkun, V., Čapkun, S.: PRCash: fast, private and regulated transactions for digital currencies. In: Goldberg, I., Moore, T. (eds.) FC 2019. LNCS, vol. 11598, pp. 158–178. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32101-7_11
20. Yu, Z., Au, M.H., Yu, J., Yang, R., Xu, Q., Lau, W.F.: New empirical traceability analysis of CryptoNote-style blockchains. In: Goldberg, I., Moore, T. (eds.) FC 2019. LNCS, vol. 11598, pp. 133–149. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32101-7_9