Dominic Deuber, Viktoria Ronge, and Christian Rückert

# SoK: Assumptions Underlying Cryptocurrency Deanonymizations

**Abstract:** In recent years, cryptocurrencies have increasingly been used in cybercrime and have become the key means of payment in darknet marketplaces, partly due to their alleged anonymity. Furthermore, the research attacking the anonymity of even those cryptocurrencies that claim to offer anonymity by design is growing and is being applied by law enforcement agencies in the fight against cybercrime. Their investigative measures require a certain degree of suspicion and it is unclear whether findings resulting from attacks on cryptocurrencies' anonymity can indeed establish that required degree of suspicion. The reason for this is that these attacks are partly based upon uncertain assumptions which are often not properly addressed in the corresponding papers. To close this gap, we extract the assumptions in papers that are attacking Bitcoin, Monero and Zcash, major cryptocurrencies used in darknet markets which have also received the most attention from researchers. We develop a taxonomy to capture the different nature of those assumptions in order to help investigators to better assess whether the required degree of suspicion for specific investigative measures could be established. We found that assumptions based on user behaviour are in general the most unreliable and thus any findings of attacks based on them might not allow for intense investigative measures such as pre-trial detention. We hope to raise awareness of the problem so that in the future there will be fewer unlawful investigations based upon uncertain assumptions and thus fewer human rights violations.

**Keywords:** anonymity, cryptocurrencies, assumptions

DOI 10.56553/popets-2022-0091
Received 2021-11-30; revised 2022-03-15; accepted 2022-03-16.

## 1 Introduction

Over the past few years, the analyses of cryptocurrency data have become common investigative measures and are now part of the daily business of law enforcement agencies [22]. Such analyses played a key role in the seizure of the prominent darknet marketplace Wall Street Market [31]. US law enforcement agencies pay millions of dollars every year to companies [28] such as Chainalysis [14], Elliptic [21] and CipherTrace [16] which claim to be able to deanonymize transactions in major cryptocurrencies. Using techniques to analyse cryptocurrency data always carries the risk of resulting in false positives. In an academic context, this does not pose a huge problem as false positives have in general no direct consequences. In contrast, when used by law enforcement the problem is far more serious. The reason for this is that investigations based on findings obtained from anonymity analyses of cryptocurrencies might lead to severe interferences with human rights. As investigative measures require a certain degree of suspicion, it is crucial to know how reliable these findings are. A low reliability might not establish the required degree of suspicion and thus might result in unlawful investigations. The reliability strongly depends upon the assumptions underlying the analyses.

There has not been much related work regarding the reliability of assumptions in the context of cryptocurrency attacks. Goldwasser and Kalai identify assumptions as being crucial to cryptography as any security proof is only as good as the underlying assumptions [27]. They notice that there are an increasing number of assumptions that restrict the possibilities of the attacker or depend heavily upon the construction which is to be proved secure. To save the value of cryptographic proofs, Goldwasser and Kalai propose a classification for cryptographic hardness assumptions. While the classification is highly recommendable for hardness assumptions, it is not sufficient to deal with all the different assumptions underlying cryptocurrency attacks. The reason for this is that the assumptions underlying cryptocurrency attacks are extremely diverse and range from computational hardness assumptions to protocol assumptions to assumptions about user behaviour. On the other side of related work, Conti et al. provided a survey on security and privacy attacks on Bitcoin. In a similar vein, Kus Khalilov and Levi focus on the anonymity and privacy of Bitcoin-like cryptocurrencies. However, both surveys

**Dominic Deuber, Viktoria Ronge:** Friedrich-Alexander-Universität Erlangen-Nürnberg, firstname.lastname@fau.de
**Christian Rückert:** Universität Mannheim, christian.rueckert@uni-mannheim.de

focus on the attacks and not on their underlying assumptions [17, 39]. Against this background, this SoK paper

1. illustrates the legal problems with uncertain assumptions and offers recommendations on how to deal with them (Sections 3 and 5);
2. surveys the underlying assumptions of cryptocurrency attack papers and makes them explicit as the papers often do not mention what those assumptions are or how reliable they are (Section 4); and
3. proposes a taxonomy of assumptions for attacks depending upon their reliability, that can be used by expert witnesses and understood by legal decision-makers (Sections 4 and 5).

We focus on Bitcoin, Monero and Zcash as Bitcoin is the most widely used currency and still wrongfully assumed by some people to be anonymous [40]. Monero and Zcash are of interest as they are the largest anonymous currencies. Moreover, all three currencies are the ones most studied by academic research and are the main drivers in darknet marketplaces [23]. In this work, we describe the proposed attacks as they are discussed in the cited papers. Thus, it is possible that the attacks are no longer applicable (see Table 3). Nevertheless, the assumptions used in these attacks are still of interest as the attacks were applicable at some point and assumptions are often reused in other attacks.

# 2 Preliminaries

In this section we explain how Bitcoin, Monero and Zcash work and introduce basic terminology as used throughout the paper.

## 2.1 Bitcoin (BTC)

We focus our introduction on Bitcoin [46] as Bitcoin helps to explain how the other cryptocurrencies work. Bitcoin is a decentralized transaction ledger that is maintained in a peer-to-peer network. The transactions are organized in blocks, which is why the ledger is also referred to as blockchain. Using a consensus mechanism, the network agrees on which blocks, i. e. transactions, should extend the ledger. The network nodes participating in this consensus mechanism are called *miners*. The consensus leader, which is the miner that suggests the next block, is rewarded for its participation with a block reward. A block reward consists of newly generated units of the cryptocurrency and transaction fees.

**Mining Pools** In Bitcoin, Monero and Zcash miners usually group their consensus work in so-called *mining pools* to reduce the variance of their payouts. If a miner successfully suggests a block, the block reward is claimed by a mining pool address rather than by one of the miners. The pool leader(s) distribute(s) the individual rewards according to certain rules agreed between the pool and the miners. In general, the transaction that is used to claim the block reward is called *coinbase transaction*.

**Transaction** A transaction $tx$ consists of a list of inputs and outputs. An output usually states an amount of Bitcoin and the hash $h_{\mathsf{pk}}$ of a public key $\mathsf{pk}$, which is also referred to as address. To spend this output, it is required to provide a public key $\mathsf{pk}'$ whose hash equals $h_{\mathsf{pk}}$ and a signature that verifies under $\mathsf{pk}'$. We also refer to outputs as coins and distinguish between unspent and spent coins. An input is a reference to an output of another transaction $tx'$ which is uniquely described by the hash of that other transaction $tx'_{hash}$ and the position of the output in the transaction's list of outputs $out_{pos}$. Usually, transactions have several in- and outputs. The input amount of transaction $tx$ is the sum of the amounts in the referenced outputs and is always consumed entirely. Thus, transaction $tx$ might have a so-called *change* output. A change output pays back to the sender(s) the difference between its input amounts and the amount that the recipient(s) should receive.

**Transaction Privacy** Research showed early on that Bitcoin does not provide anonymity because it is possible to link addresses [1, 42, 51, 52, 57]. Numerous countermeasures were subsequently developed [41, 53, 54], and new cryptocurrencies emerged that feature privacy by design [18, 44, 72]. The most important of those cryptocurrencies are Monero [44] and Zcash [72], which will be discussed in Section 2.3 and Section 2.2 respectively. The countermeasure that is significant in the context of this work is *CoinJoin* [41]. The central element of CoinJoin is a CoinJoin transaction whose inputs and outputs belong to multiple entities by design. The goal of that design is to break address linkability.

**Peer-to-Peer Network** The blockchain of a cryptocurrency is usually maintained in a peer-to-peer (P2P) network. A P2P network is a network without a central server as is the case in a client-server architecture. All cryptocurrencies studied throughout this work are *permissionless*, meaning that anyone can join and participate in the network at any time. In the following, we explain the basic functioning of a P2P network using Bitcoin as an example. Since the networks of Monero and Zcash function in a similar way, we do not present

them separately and only refer to the differences, if necessary.

There are several different nodes participating in the Bitcoin network. *Full nodes* hold the entire blockchain and verify all the data. In contrast, there are also *light clients* which are nodes that only hold few data and therefore rely on communication with full nodes for verification. All nodes exchange messages via TCP. Every 24 hours, or when initially joining the network, each node broadcasts its own IP address to its peers using an *addr* message. The peers will relay this message to some of their peers. Messages concerning transactions or blocks are propagated differently. First, an *inv* message is sent to all peers. The peers that actually want the full data request it via a *getdata* message. Propagation works by the receiving peers then broadcasting to their peers and so on where no peer requests data it already has [7].

## 2.2 Zcash (ZEC)

While Zcash is commonly considered to be anonymous, this is only partly true. In fact, Zcash takes a two-part approach, where coins are either part of the *unshielded* or the *shielded* pool. Zcash and Bitcoin have in common that coins have to be spent entirely which results in the creation of new coins to retrieve the change. The unshielded pool behaves like Bitcoin, while the shielded pool hides senders, recipients and the transferred amount. The coin of the sender is hidden within the whole set of coins ever created in the shielded pool. A zkSNARK [6] is used to prove knowledge of the secret key of this coin without revealing which one as well as proving the coin has not been spent before. The sender also hides the recipient by not putting the recipient's public key directly into the transaction (and therefore on the blockchain) but by using the key to encrypt the information needed to spend the generated coins later.

Zcash calls recipient keys inside the unshielded pool $t$-addresses and inside the shielded pool $z$-addresses. This leads to four types of transactions, which are $t$-to-$t$, $z$-to-$z$, $t$-to-$z$ and $z$-to-$t$ transaction. Based upon several attacks (e. g. [4, 34]), especially transactions between the two different pools seem to be vulnerable to attacks, as illustrated later in more detail.

## 2.3 Monero (XMR)

The overall structure of Monero is similar to the shielded pool in Zcash, i. e. there is no "non-anonymous" part in Monero. In contrast to Zcash, not all coins within the shielded pool are used as input, but rather some kind of decoy selection from the shielded pool takes place. These decoys are other coins which are included into the transaction as dummy inputs to hide the actual input of the spender. Decoy selection was a promising target for attacks in the past (e. g. [38, 45, 71]).

Another crucial difference is the choice of the underlying proof system. While Zcash uses zk-SNARKs and therefore requires a trusted setup, Monero uses zero-knowledge proofs without trusted setup, however, at the cost of larger transactions. Besides that, the two currencies differ in the specific use of recipients' public keys. While in Zcash the same key is used, but only a part of it is re-randomized, in Monero the key is published in an altered way, i. e. the public key and therefore the secret key are re-randomized in every transaction but still accessible by the recipient.

## 2.4 Attacks

In the context of privacy, we define an attack to be any attempt to gain additional knowledge about at least one transaction. This knowledge can refer to the sender(s), recipient(s) or amount(s) of the transaction(s). Attacks often use *heuristics*, i. e. methods that are not guaranteed to be optimal, but nevertheless lead to results in a reasonable time. Heuristics and consequently attacks often rely on assumptions which decide on the meaningfulness of the results, as further discussed in Section 3.

The goal of an attack is deanonymization and/or re-identification. We use the term "deanonymization" in line with Kelly et al., who define anonymity as unidentifiability and unlinkability [35]. Consequently, we refer to the following two methods as deanonymization: 1. Clustering: Different addresses, keys or transactions are linked together/clustered and assumed to be controlled by the same entity. 2. Identification: Identifying the actual spender/recipient in an anonymity set. In contrast, re-identification refers to identifying the entity that controls an address.

In terms of Bitcoin, deanonymization is usually done by address clustering, while attribution tagging might directly or indirectly allow to identify the entity which is controlling the addresses (re-identification). Address clustering is linking several addresses that belong to the same entity. Attribution tagging is tagging address clusters with attribution information that is either personally identifiable information or can be used to retrieve such information. An example for the latter are clusters tagged with exchange information. In that case, law enforcement agencies might retrieve personally identifiable information by requesting it from the exchange.

We focus on *passive* attacks, as elaborated in Section 4. In general, passive attacks are characterized by leaving the data in a system untouched. In terms of blockchain attacks, a passive attack is one that analyses the blockchain data without altering it. In terms of network attacks, passive refers to participation in the network, but without communicating beyond what is required. This means that such attacks only listen and all requests from other participants are answered in accordance with the protocol. Thus, all requests are answered as the standard software of the respective currency would do, without altering any data.

# 3 Legal Relevance

Investigations involving cryptocurrency forensics typically start with a non-blockchain-related event, such as the seizure of a darknet marketplace or a child-pornography platform. In many cases to identify potential offenders, payments conducted via the seized market or platform are traced. As cryptocurrencies are criminals' default payment method on the darknet [23], tracing payments requires cryptocurrency forensics.

The most prominent example where cryptocurrency forensics were decisive for the success of the investigations is the seizure of Wall Street Market, one of the largest darknet marketplaces [19]. Crucial to the investigation was an analysis of the Bitcoin blockchain performed by the US Postal Service [63]. In the analysis, the investigators utilized proprietary software of an undisclosed company and argued that this software had been found to be reliable through numerous unrelated investigations [63]. This statement indicates that the investigators employed the cryptocurrency forensics software as a black box. The problem with the black box usage of software is that the specific methods utilized cannot be observed and, therefore, the quality of the results remains unclear. There are several proprietary cryptocurrency forensics tools, such as Chainalysis [14], Elliptic [21], and CipherTrace [16]. US law enforcement agencies pay millions of dollars every year to utilize those tools [28], which is why it can be assumed that the forensic methods employed by those tools are relevant in practice. As law enforcement does not publicly reveal its methods, in order to prevent criminals from developing and employing anti-forensic measures, publicly available information must be relied upon to determine which techniques closed-source proprietary tools utilize. Besides that, open-source cryptocurrency forensics tools such as BlockSci [33] or GraphSense [30] can be analyzed in order to establish which methods they employ. WalletExplorer [64] was a tool that was capable of address clustering and attribution tagging in Bitcoin. The address clustering of WalletExporer was based on the so-called multi-input heuristic, which states that all addresses in the inputs of a transaction are controlled by the same person [1, 42, 51, 52]. For attribution tagging, WalletExplorer's former operator registered and interacted with several services, such as exchanges. Through the interaction, he was able to cluster the addresses of the services and tag the clusters with the name of the service. According to the former operator, who works at Chainalysis today, Chainalysis offers the same service as WalletExplorer but is far more advanced [64]. The open-source cryptocurrency forensics tool BlockSci performs mainly address clustering and thereby enforces the multi-input heuristic by design [9]. While GraphSense extends BlockSci, inter alia, with attribution tagging, this tool also focuses on the multi-input heuristic for address clustering [30]. Against this background, it can be assumed that at least the multi-input heuristic is crucial for cryptocurrency investigations in practice.

## 3.1 Uncertain Assumptions

The widely employed multi-input heuristic is based on the assumption that all inputs to a transaction are controlled by the same person. This assumption concerns user behaviour as it excludes behaviour where numerous persons contribute inputs to a transaction. However, it is not only possible for multiple persons to contribute inputs to a single transaction, but also desirable as in the case of so-called CoinJoin [41] transactions. In a CoinJoin transaction, addresses in the in- and outputs are controlled by multiple persons by design. This design is intended to prevent address clustering. In specific terms, this means that the multi-input heuristic applied to a CoinJoin transaction yields a false positive as it assumes that the corresponding addresses are controlled by a single person. Avoiding such false positives would only be possible if CoinJoin transactions could be clearly distinguished from other transactions. However, the detection of CoinJoin transactions is also based upon methods of which their reliability is not known. Apart from CoinJoin transactions, it is in general difficult to evaluate the reliability of assumptions that concern user behaviour as user behaviour is subject to change. Evaluating the reliability would require ground truth data about user behaviour at the time the transaction in question was issued. Such ground truth data, however, is usually unavailable or extremely difficult to obtain [25]. Conse-

quently, some uncertainty remains in the multi-input assumption which can neither be assessed nor quantified due to the lack of ground truth data. Uncertainty in assumptions can cause numerous legal issues as discussed in the following.

## 3.2 Legal Issue Underlying Wall Street Market Investigation

The most practically relevant legal issue caused by uncertain assumptions is illustrated by the example of the Wall Street Market (WSM) investigation. The blockchain analysis performed by the US Postal Service can be summarized as identifying wallets, detecting payments between wallets, "de-mixing" and associating wallets with darknet marketplaces. Results of the analysis were ultimately used to request personal data about the sender of a specific transaction from a Bitcoin Payment Processing Company (BPPC). This specific transaction was believed to have originated from a wallet that was used to pay one of the marketplace's administrators. The obtained personal data allowed to associate the wallet (addresses) with a natural person (entity).

In the following, we analyze the investigation with regard to the required degrees of suspicion and possible effects of uncertain assumptions. The legal basis for the request to the BPPC cannot be ascertained from the criminal complaint. If the request was conducted by means of a warrant under the Fourth Amendment, probable cause would have been the required degree of suspicion. The probable cause requirement would not be necessary in the case of a subpoena under the third-party doctrine. However, any subsequent investigative measures, such as electronic surveillance or searches of premises, would require probable cause on the basis of the Fourth Amendment, which must at least extend to the linkage of the wallet in question to criminal activity. The third party doctrine therefore does not preclude the following explanations. The legally relevant question is whether the results of the analysis were sufficient to establish the required degree of suspicion. It is not sufficient that one of the suspected WSM administrators later confessed, because the suspicion must exist at the time when the personal data was requested. The same applies to the argument that the proprietary software employed in the analysis has always been reliable so far. It can be assumed that in the WSM blockchain analysis uncertain assumptions played a role. This follows from the fact that wallets were identified, which typically involves address clustering and thus at least the multi-input heuristic. However, if the heuristic was applied to a CoinJoin transaction, for example, addresses that have nothing to do with the WSM would be incorrectly associated with a WSM wallet. The criminal complaint does not indicate that CoinJoin transactions were excluded and thus false positives were prevented. As a result, the prosecution's argumentation would be broken. Namely, the questionable transaction could not be believed to have originated from a wallet that was used to pay one of the marketplace's administrators. If subsequent investigative measures, such as a search, were to be based solely on the result of the blockchain analysis, it would be questionable whether the necessary degree of suspicion could be established and thus whether the measure was lawful. The blockchain analysis of the Wall Street Market investigation and the impact of uncertain assumptions is discussed in detail in Section 5.5. A detailed explanation of the significance of degrees of suspicion in investigations and why the legal problem extends in principle to all jurisdictions is set out in Appendix A.

## 3.3 Towards a Solution

Insufficient exclusionary or admission rules as well as rash trust in IT expert witnesses even in the highest courts (see Appendix A) demonstrate that there is a need for action especially against the background of uncertain assumptions employed by cryptocurrency forensic methods. Even though the common law and the continental European systems differ greatly in some respects, they have in common that legal decision-makers need to have a precise understanding of the reliability of forensic methods in order to be able to reach a proper judgment. The first step concerning an uncertain assumption is to create awareness of the problem by making the assumption and its uncertainty transparent in research papers. As a result, an expert witness can present assumptions with their uncertainties in criminal proceedings. Only in this way can decision-makers take uncertain assumptions into account and not run the risk of blindly following an expert witness. Another advantage of transparency is that the defence or prosecution can challenge the evidential weight of circumstantial evidence. Likewise, in pre-trial stages, investigators can consider assumptions and their uncertainties when determining whether a required degree of suspicion can be established.

Transparency can be achieved by stating the nature of the assumption following the taxonomy presented in Section 4. In Appendix B we outline how our taxonomy could be used in practice. The taxonomy enables argumentation regarding the general uncertainty because,

| *type* | *assumption* | *usage* |
|---|---|---|
| User Behaviour | Multi-Input | [1, 2, 24, 29, 32, 34, 42, 47, 48, 51, 52, 57] |
| | Change-Address | [1, 34, 42, 47, 57] |
| | Cluster-Intersection | [20, 26] |
| | No-Proxy | [36] |
| | Miner-Payout | [3, 34] |
| | Value-Input-Output | [4, 34, 50] |
| | Fingerprinting | [4] |
| Protocol | Response Time | [61] |
| | Wallet Communication | [61] |
| Computational Hardness | No Double-Spending | [38, 45, 69, 71] |
| Statistical | Unique Entry Nodes and No-Collision | [5] |
| | Multi-Output | [38] |
| | Newest-Account | [38, 45] |

**Table 1.** Overview of assumptions and their usage grouped by their type in accordance with our taxonomy.

for example, well-established computational-hardness assumptions are not as uncertain as those on user behaviour. While, at first glance, relying on a taxonomy seems to be imprecise, with respect to legal decision-making it is not. Using a taxonomy to classify assumptions results in a normative statement which naturally fits the legal decision-making process. Thus, such a taxonomy can be the basis of a common comprehensible language between expert witnesses and legal decision-makers and also be a first step towards a standard for the interpretation of any findings, as proposed in the literature [15].

# 4 Taxonomy of Assumptions

Different assumptions which underly heuristics targeting (anonymous) cryptocurrencies have often been neglected in academic discourse. Consequently, no system has yet evolved to address different types of assumptions, their applicability and their quality. To fill this gap, we propose a taxonomy for classifying the assumptions which consists of four classes. The first of these is *user behaviour*, which relies on e. g. patterns that users follow. The second one is the class of *protocol assumptions*. We understand the term protocol broadly and refer to its meaning in the context of networks, implementations, etc. The third class are *computational hardness assumptions* and the last one *statistical assumptions*, which uses statistical arguments.

We classify all assumptions according to our taxonomy and explain which heuristics/attacks are based upon them. Thereby, we focus on assumptions which are the basis of passive attacks. The reason for this is, firstly, that passive attacks make up the majority of the proposed attacks. Second, because active attacks often rely on protocol specifics which would be so extensive in their presentation that we can only give a high-level description and third, because the motivation for a taxonomy is based upon the increasing use of attacks in criminal investigations. As law enforcement agencies by their very nature operate in the aftermath of crimes, they will predominantly use passive attacks. This is also indicated by the fact that the US spends millions on commercial analysis software [28] that arguably perform mostly passive analysis [12].

Most of the time, heuristics or attacks are based on several assumptions, which is why we locate them at the main assumption. We understand the main assumption as the one whose uncertainty has the greatest impact on the probative value of results obtained by the heuristics/attacks, as discussed in Section 3. By the uncertainty of an assumption, we refer to the probability that the assumption is wrong. If this probability is 0%, the assumption is absolutely reliable. Likewise, if the uncertainty of an assumption is 100%, this states that the assumption is wrong and therefore diminishes the probative value. Depending on the exact definition of an assumption, its uncertainty is also its false positive rate, as illustrated by the following example. Let the assumption be that in a Monero transaction, the key with the smallest hash value always refers to the spender. Then the uncertainty is the probability that this assumption is wrong and therefore the key with the smallest hash is not the spender, i. e. the false positive rate.

An overview of all assumptions and their usage in the sense of our taxonomy is depicted in Table 3. Note that whenever it was appropriate, we named the assumption as the corresponding heuristics/attacks. If they were not named, we provide a suitable name.

## 4.1 User Behaviour Assumptions

User behaviour assumptions are based upon common behavioural patterns of the users of a cryptocurrency. An example could be that there is a payout transaction from a mining pool every day at around 8pm. In part, such behavioural patterns are derived from the standard implementation or known applications of the cryptocurrency in question. It is often assumed that these patterns can be transferred to many other users while it is unclear in reality how precise this assumption is. User behaviour changes over time as, for example, new ap-

plications evolve. Kus Khalilov and Levi also categorize attacks based upon user behaviour but without focusing on the underlying assumptions [39].

#### 4.1.1 Multi-Input Assumption

The multi-input assumption assumes that all inputs to a transaction are controlled by the same entity [1, 42, 51, 52]. This directly leads to the multi-input heuristic which uses this assumption to cluster input addresses. In combination with other clustering heuristics, this enables the tracking of payment flows through the blockchain. Address clustering heuristics are part of deanonymization attacks. Re-identifying the entity behind an address cluster requires additional steps (see Section 2.4).

For CoinJoin transactions, the assumption is false as such transactions combine the inputs of multiple entities by design. Consequently, applying the multi-input heuristic to CoinJoin transactions would lead to false positives which is problematic (see Section 3.1).

The multi-input heuristic is used quite often in Bitcoin and also Zcash analyses [1, 2, 24, 29, 32, 34, 42, 47, 48, 51, 52, 57], although the discussion of how reasonable the assumption is differs greatly. There are papers that completely forgo any discussion of whether the assumption is reasonable. This is done by either directly referencing the Bitcoin whitepaper [51] or by saying that the assumption is safe to make [57]. Other papers recognize the possibility of false-positive results and therefore take greater argumentative effort. Ron and Shamir asked several members of the Bitcoin community who confirmed that overestimations of common ownership are very unlikely [52]. Androulaki et al. argue that Bitcoin client software does not support that different users participating in a single transaction [1]. The authors see the possibility of CoinJoin-like transactions, however they argue that these are unlikely to become the most common transactions in the network. Meiklejohn et al. argue in a similar way by saying that the multi-input heuristic exploits inherent properties of the Bitcoin protocol. Therefore it is unlikely that several entities spend together in one transaction as they would need to reveal their secret keys to each other [42]. In contrast, Koshy, Koshy, and McDaniel explicitly removed all multi-input transactions from their analysis as they wanted to be sure that each transaction was only controlled by a single entity. According to the authors, related work would not acknowledge that a multi-input transaction might be controlled by several entities [36]. For Zcash, Kappos et al. argue that the assumption is used a lot in Bitcoin and suggest it might be even better in Zcash as they are not aware of any protocols such as CoinJoin which explicitly contradict the assumption [34].

#### 4.1.2 Change-Address Assumption

Generally, the multi-input heuristic does indeed only allow tracking of the payment flow of an entity throughout the blockchain in combination with other heuristics. The reason is that the multi-input heuristic is only considering the inputs of a transaction. In order to create address clusters covering several transactions, the outputs of a transaction also need to be taken into account.[1] This is exactly what change-address heuristics [1, 34, 42, 47, 57] are doing. As Bitcoin requires the input values of a transaction to be spend completely, there could be change addresses paying back the remainder to the (potential single) entity that created the transaction. The most basic form of a change-address heuristic works as follows. For every transaction with two output addresses $a_O$ and $a_N$, if $a_N$ never appeared before but $a_O$ did, then $a_N$ is considered the change address [1, 57]. In combination with the multi-input heuristic, this results in an address cluster consisting of the change address and the input addresses of a transaction. This heuristic is based upon the assumption that no transaction spends to two different users [1, 57].

With the rise in gambling sites and mining pools, the assumption no longer holds [42]. Mining pools usually have huge payout transactions rewarding their participants with shares of the block reward. This is why Meiklejohn et al. proposed several refinements [42]. The refined version states that some output address of a non-coinbase transaction (see Section 2.1) is the change address if it is the only address in the outputs that appeared for the first time and there is no "self-change address", meaning no output address that appeared in the inputs [42]. According to Meiklejohn et al. this heuristic is not robust as it is based on the "idiom of use" where the change address is created internally by the bitcoin client and never reused. The authors acknowledge that the heuristic might need to be discarded if usage patterns change. In fact, they found false positives by comparing their cluster results to tags associated with the respective addresses which they obtained by employing off-chain information [42]. This lead to further refinements making the heuristic far more conservative than its basic form. Nevertheless, even after the refinements had been published, the non-refined basic version of the heuristic has been used and also stated

1 An exception is the reuse of addresses which is however considered bad practice and therefore prevented by most wallets.

to be conservative [57]. Finally, the change-address assumption in Bitcoin can be seen as two-fold. First, there is the implicit assumption that there is any change at all, and, second, that transactions are issued using a Bitcoin client which generates a fresh change address for every new transaction.

The first part of the assumption, namely that there is any change at all, is also found in Zcash. In Zcash there were so-called vJoinSplit transactions which allow to have up to two $t$-inputs and $t$-outputs (as well as $z$-inputs and $z$-outputs). The $t$-input-output heuristic states that in a vJoinSplit transaction the $t$-input(s) and a $t$-output belong to the same entity if there is exactly one $t$-output [34]. The intuition behind this heuristic is that the $t$-output is probably the change output when only some of the input amount is moved to the shielded pool. Thus, the heuristic is based on the implicit assumption that there is any change at all. However, Kappos et al. did not use the heuristic in their anonymity analysis of Zcash as they assumed there might be too many false positive in case a transaction just spends to an address in the shielded and one in the unshielded pool [34]. In other words, they did not use the heuristic because they considered the assumption that there is any change at all to be too unreliable.

#### 4.1.3 Cluster-Intersection Assumption

The cluster-intersection attack [26] tries to link address clusters by intersecting the anonymity sets of CoinJoin transactions. The attack exploits additional knowledge about outputs from different CoinJoin transactions being controlled by the same entity. We illustrate the attack with the following example based upon [26]. Let Alice be in control of addresses $A_1$ and $A_2$. Assume that those addresses are linkable by the multi-input and/or change-address heuristic, resulting in address cluster $C_{pre}$ where *pre* means *pre*-mixing. Now Alice uses CoinJoin to break this linkability. $A_1$ will be input to CoinJoin transaction $ctx_1$ and $A_2$ will be input to CoinJoin transaction $ctx_2$. The addresses $A_{1*}$ and $A_{2*}$ in the outputs of $ctx_1$ and $ctx_2$ respectively, should no longer be linkable as there are multiple other entities participating in CoinJoin besides Alice. Furthermore, assume that Alice pays a merchant using $A_{1*}$ and at some point in the future pays the same merchant using $A_{2*}$. Now the merchant learns that $A_{1*}$ and $A_{2*}$ belong together, i. e. belong to address cluster $C_{post}$, where *post* means *post*-mixing. It is further possible for the merchant to determine the anonymity sets of $ctx_1$ and $ctx_2$ by applying the cluster-intersection attack. Both anonymity sets contain the address cluster $C_{pre}$ as Alice participated in both CoinJoin transactions. By intersecting the anonymity sets, which also contain several address clusters controlled by different entities, the merchant might learn that $C_{pre}$ is linkable to $C_{post}$.

The cluster-intersection assumption assumes that an entity uses a single wallet where all addresses prior to some mixing procedure are linkable, i. e. those addresses can be clustered into a single address cluster ($C_{pre}$). Thus, the assumption is at least as uncertain as the most uncertain assumption used in address clustering (to create $C_{pre}$). As it is unclear which assumption that is, we explicitly state the cluster-intersection assumption. We deliberately list the assumption under user behaviour as we already examined the two uncertain address-clustering assumptions (multi-input and change in the previous sections). Besides that, the part of the assumption stating that a user uses a single wallet refers to user behaviour. There is some additional uncertainty beyond the most uncertain assumption used in address clustering because addresses of the wallet need to be linkable. In the above example, $A_1$ and $A_2$ might not have been linkable before mixing. If another entity Eve also participated in $ctx_1$ and $ctx_2$ with the two linkable addresses $A_{1_e}$ and $A_{2_e}$, the attack might link $C_{post}$ to those addresses. This would clearly be a false positive link. Goldfeder et al. showed the general applicability of the attack in Bitcoin on simulated data. The authors further acknowledge that, in reality, a users wallet may not be linkable into a single address cluster [26]. Deuber and Schröder applied the attack on real transactional data of the cryptocurrency Dash [18]. To cope with the uncertainty in the assumption, they, inter alia, added a mechanism to reject obvious false positives [20].

#### 4.1.4 No-Proxy Assumption

The no-proxy assumption states that no proxies have been used. The assumption is used in an attack by Koshy, Koshy, and McDaniel. The authors link Bitcoin addresses and IP addresses by exploiting how transactions are propagated in Bitcoin's P2P network [36]. In contrast to address-clustering attacks, linking clusters to IP addresses is a re-identification attack (see Section 2.4). To link Bitcoin with IP addresses, Koshy, Koshy, and McDaniel build a custom Bitcoin node that connects to all peers. Now the custom Bitcoin node can record the entire transaction propagation history, i. e. the times when a transaction has been relayed and the IP addresses of the relaying peers. The authors identified relaying patterns, the simplest and most common one be-

ing that a transaction is relayed by several peers, but only once per peer. By exploiting relaying patterns, it is possible to link the IP address of the relaying node to the Bitcoin addresses in the transaction. The no-proxy assumption is crucial, as the attack might result in false positives if users are relaying transactions through a proxy or use TOR [36]. Besides that, the attack also assumes that there are no false positives due to slow internet connections [36].

#### 4.1.5 Miner-Payout Assumption

Generally, every transaction requires a fee, which motivates mining pools to pay all their miners in a single transaction. This behaviour leads to transactions with over 100 outputs that sometimes occur regularly at fixed times. The miner-payout assumptions states that a transaction with over 100 outputs is issued by a mining pool. In Zcash, Biryukov and Feher studied this behaviour using addresses and won block information published by mining pools on their websites [3]. This was done for mining pools that receive the reward on $t$-addresses as well as for those that receive it on $z$-addresses. While the authors suggest the miner-payout assumption is reasonable, they also admit that it is difficult to find pools with only a small proportion of the overall mining power. Additionally to this problem, mining pools can use $t$-addresses or $z$-addresses, where $z$-addresses make linking harder. Kappos et al. assume as well that transactions with many outputs are issued by mining pools, but give less specific numbers [34]. In any case, the miner-payout assumption only helps clustering addresses from mining pools, but most likely will not deanonymize single users.

#### 4.1.6 Value-Input-Output Assumption

The unique structure of Zcash requires transactions between the shielded and unshielded pool, which reveal the value of transactions. This reveal is utilized in the value-input-output assumption which states that if a $z$-to-$t$ transaction appears after a $t$-to-$z$ transaction containing the same unique value, they are linked. It is argued that a $t$-to-$z$ transaction and a subsequent $z$-to-$t$ transaction which both contain the same unique value, are very unlikely to happen by chance [4, 34, 50]. Only Biryukov, Feher, and Vitto try to provide a false-positive rate by checking how long the uniqueness of a value is sustained in the blockchain, while other direct evidence seems to be hard to obtain at all [4]. The idea was extended to also account for a number of transactions within the shielded pool, which would decrease the value of the $z$-to-$t$ transaction by that number multiplied by the transaction fee [4, 50]. Kappos et al. only mention the extension but do not apply it [34]. They proposed a special variant of the value-input-output assumption caused by transactions done by the founders, i. e. entities who obtain a share of the mining reward as inventors of Zcash. The corresponding $t$-addresses are specified in the protocol itself.

Kappos et al. observed that the values send from these $t$-addresses to the shielded pool are quite unique. They used the value-input-output assumption to link withdrawal transactions from the shielded pool containing these specific values to the founders. Thus, besides address clustering, this linkage also allows reidentification as the founders are known.

In general, however, the attacks using the value-input-output assumption are address-clustering attacks and additional knowledge is required for reidentification.

#### 4.1.7 Fingerprinting Assumption

The fingerprinting assumption can be seen as a variant of the value-input-output assumption in Zcash. However, it is not about the actual value of a $t$-to-$z$ transaction but rather about the last few digits of the value, i. e. from $10^{-2}$ to $10^{-8}$ ZEC[2]. Thereby, Biryukov, Feher, and Vitto say a value has a unique fingerprint if either the last four digits, i. e. $10^{-5}$ to $10^{-8}$, are not round and unique or at least 5 of the last 7 digits, i. e. $10^{-2}$ to $10^{-8}$ build a unique pattern [4]. The fingerprinting assumption assumes that if a $t$-to-$z$ transaction and a $z$-to-$t$ transaction share a unique fingerprint, they are linked by a sequence of transactions. Uniqueness is only considered for a certain time range and the authors assume that fingerprints occur either intentionally by crafted coins or are the product of mining rewards split by mining pools [4]. For the second case, the authors provide a model to estimate the quality of the assumption. Attacks based on the fingerprinting assumption are address-clustering assumptions.

## 4.2 Protocol Assumptions

Protocol assumptions can be seen as distinguished from the three other ones as they do not depend on behaviour or theory behind cryptocurrencies but on practical means. If there are flaws in the protocol itself, then

**2** These values lie beyond the transaction fee and have basically no economic meaning.

exploiting them does not require any additional "protocol" assumption as they simply exist.

One part of protocol assumptions are assumptions depending on the network topology, e. g. a node directly connected to a client will propose a transaction. Another part are assumptions regarding implementations. A known problem of implementations are different execution paths in the code depending on the validity of the input. While this problem has been well known for several decades, it still appears from time to time.

#### 4.2.1 Response-Time Assumption

Tramèr, Boneh, and Paterson noticed that depending on whether a transaction is destined for a certain node in Zcash, this node responds slower to network requests sent to it after receiving this particular transaction [61]. The response-time assumption states that it takes longer for a node in Zcash using the common wallet software to respond after receiving a transaction destined for this node. The first step when receiving a transaction in Zcash is to check whether decryption works, which is only the case if the node is the actual recipient. If the decryption succeeds, an additional Pedersen commitment [49] check is done to check for the well-formedness of the message. Answering further network requests is delayed until this additional check is completed. Thus, the time can be measured and results in the attack described in [61].

#### 4.2.2 Wallet-Communication Assumption

The three major wallet implementations for the Monero client have in common that by default they connect to a remote node that is responsible for network communication. The wallet-communication assumption states that there are different communication patterns between a remote P2P node and a wallet depending on whether a previously received transaction belonged to the wallet or not [61]. The patterns are due to default strategies of the wallets for requesting transactions. Upon each request, the wallet requests a list of hashes of transactions unconfirmed so far. After that, the wallet requests the bodies of transactions which it has either not processed so far or where it is the payee. At time of publishing of [61], the transaction rate of Monero was so low that the arrival of new transactions unlikely occurred between two wallet refreshes. An adversary can observe which transaction bodies are requested several times and thus belong to the requesting wallet.

### 4.3 Computational Hardness Assumption

Computational hardness assumptions are assumptions about the impossibility of solving computational problems efficiently, i. e. in polynomial time. For most of the currently used assumptions in cryptography and related fields of computer science and mathematics it is not known how to prove such hardness. Nevertheless, some of theses assumptions like the Discrete Logarithm assumption enjoy almost complete trust. Thus, if a heuristic is based on such an assumption, the results have a huge probative value. On the other side, if the assumptions turn out to be wrong, it would have tremendous impact, not only on the probative value, but also on the security of cryptographic primitives and protocols.

#### 4.3.1 No Double-Spending Assumption

While the prevention of double spending is not a cryptographic primitive itself, it directly depends upon cryptographic assumptions such as the discrete logarithm assumption. Attacks based upon graph analysis [38, 45, 69, 71] all implicitly use this assumption as they assume that spenders of different transactions are distinct. The attacks consider each ring as a set with exactly one true signer and search for unions of such sets so that the number of sets in these unions equals the number of accounts. If double spending is impossible, then each account within the union already spent. If the account is part of another set outside the union it can not be the signer and thus reduces this set's anonymity. The extreme case is a union of size one, i. e. a transaction with only the actual signer as ring, which is called a *zero-mixin* transaction. The attacks can be combined with a *black marble attack* where the attacker owns several coins in the system. If these coins are included as decoys in Monero, the attacker can rule out all of their own coins as spender [66, 68, 70].

Another attack vector relying solely on the no double-spending assumption occurs with hard forks in Monero. Wijaya et al. observed the issue of key reuse, where the same key image appears in transactions in two different forks of Monero [67]. The intersection of the two rings for such transactions necessarily contains the true spender and therefore is at least partly deanonymized.

### 4.4 Statistical Assumptions

Statistical assumptions are assumptions about "how likely" an event is. Such an event can either be caused by randomness used in the cryptocurrency on purpose

or caused by outside effects. Examples of the two causes are probabilities for certain accounts being part of a ring in Monero (Section 4.4.3) and the set of entry nodes for a Bitcoin client (Section 4.4.1) respectively. Statistical assumptions always include some error probability, which can, however, be dealt with accordingly, as it can be computed.

### 4.4.1 Unique Entry-Nodes and No-Collision Assumption

Biryukov, Khovratovich, and Pustogarov propose a re-identification attack that links Bitcoin addresses to IP addresses, similar to the attack proposed by Koshy, Koshy, and McDaniel discussed in Section 4.1.4. The attack distinguishes two types of nodes in Bitcoin's P2P network, namely servers and clients. The goal of an attacker is to learn which transactions a specific client issued. Clients only establish at most eight outgoing connections, whereas servers allow for both outgoing and incoming connections. Thus, an attacker cannot directly connect to a targeted client. The basic idea of the attack as proposed by Biryukov, Khovratovich, and Pustogarov [5] is that the attacker connects to as many servers as possible and analyses the transactions relayed by these servers. The attack utilizes the fact that a server that is directly connected to the targeted client will learn about the client's transaction earlier and therefore relay it earlier. The servers to which a client connects directly (via one of its eight outgoing connections) are called *entry nodes*[3]. The unique entry-nodes assumption assumes that a client's entry nodes are unique. For any transaction, the attacker checks whether the first few propagating servers belong to the targeted client's entry node set, which the attacker learned in a previous step. If the first propagating servers match the targeted client's entry nodes, the attacker can infer that the transaction originated from that client. The matching relies on the assumption that there are no collisions, i. e. among the first ten servers that propagate a client's transaction, there is no subset of three servers that accidentally belong to some other client's entry note set. This assumption essentially excludes false positives because if something went wrong with the attack, the attack does not point to a client that did not issue the transaction. Biryukov, Khovratovich, and Pustogarov argue that the probability of a collision is negligible [5].

**3** Biryukov, Khovratovich, and Pustogarov show how to learn the entry node set of a client utilizing address propagation [5].

### 4.4.2 Multi-Output Assumption

In Monero, decoy members of the input ring to a transaction are sampled randomly according to a gamma distribution with fixed parameters [43]. Due to this distribution, the likelihood is very low for two outputs of the same transaction to appear together in different input rings of another transaction. This is captured by the multi-output assumption which says that if two outputs of the same transaction appear together in different input rings of another transaction, these two outputs belong to the same entity and are the actual spenders of the other transaction. Kumar et al. use this assumption in their heuristic of the same name to deanonymize some senders but are aware of the possibility of false positives [38].

### 4.4.3 Newest-Account Assumption

The newest-account assumption states that the newest account in an input ring of a transaction in Monero is the actual spender [38, 45]. This assumption was proposed by Kumar et al. and Möser et al. for older versions of Monero. In the literature, the heuristic using the newest-account assumption is sometimes referred to as "guess-newest heuristic" (e. g. [45]) to emphasise the guess an adversary would make to deduce the actual spender. It is shown that the heuristic delivers reasonable results by comparing the results with ground-truth data gained from the so-called *zero-mixin* attack (see Section 4.3.1) [38, 45]. The current version of Monero uses a gamma distribution [43], which gives preference to newer accounts and therefore the false-positive rate should be increased. Furthermore, the ground-truth data can no longer be generated as before as Monero demands a minimum ring size which prevents the *zero-mixin* attack.

## 4.5 Practical Relevance

The assumptions differ in their practical relevance. The overview in Table 2 indicates how relevant each assumption might be. We consider all assumptions where the weaknesses exploited by the corresponding attacks have been fixed as having low relevance. The reason is their limited effect on current and future blockchain activity. As there are attacks requiring little effort, we consider assumptions employed by attacks that require significant effort, such as relay-pattern attacks which need active engagement in the Bitcoin network, to have low practical relevance as well. We determined medium relevance as follows. For the Zcash assumptions, Chainalysis stated that they are aware of the corresponding

attacks [13], thus they might be actually used. In the case of Monero, Internal Revenue Service entered into two contracts with Chainalysis and Integra FEC for $625,000 to develop tracing methods for Monero [56], so there are also potential use-cases for the corresponding attacks and assumptions they rely upon. Finally, the multi-input and change-address assumption are highly relevant as it is very likely that they are actually relied upon by law enforcement as discussed for the multi-input assumption in Section 3.

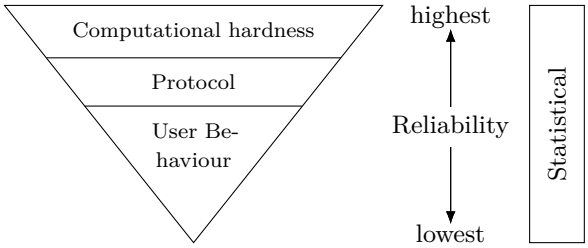


**Fig. 1.** General reliability of assumptions by category from highest to lowest

# 5 Arguing Reliability

Stating assumptions in terms of our taxonomy can just be a first step towards more informed legal decision-making. A further requirement is a general understanding of how reliable an assumption category is and how reliable categories are compared to each other. To this purpose we will discuss the general reliability of the four categories, which leads to a natural relation and partial order between them. Note that this comparison can only be a guideline of how to deal with different categories of assumptions in general, as concrete assumptions might be studied and understood better or worse. Furthermore, it is possible that the reliability of a well-studied assumption deviates from other assumptions in its category.

An overview of our result is shown in Figure 1. We identify well-established computational hardness assumptions to be the most reliable ones, as they have been studied extensively over a longer time period and independent of cryptocurrencies. Protocol assumptions might also be very reliable but are probably less reliable than computational hardness assumptions. The reason is that they heavily depend on the protocol which, for example, can be changed by the integration of new features or because parts that where shown to be vulnerable to attacks were removed or fixed. On the other hand, if several protocols, e. g. wallet implementations, are allowed in the same environment, to make attacks using a protocol assumption reliable in practice it has to be known which protocol is used. The last assumption in this order are user behaviour assumptions, which appear to be the most unreliable ones in general, as user behaviour is subject to change and hard to assess. Statistical assumptions are not part of this order as, in contrast to the other three, their reliability can be computed due to the definition of statistical assumptions (see Section 4.4).

While it might seem desirable to order the assumptions within a category, this is neither possible nor useful for the application of our taxonomy. To see why it is not possible, let there, for example, be two user behaviour assumptions. Both of them might rely on the use of some special (but different) wallet but otherwise rely on similar ideas. If we do not know anything about the actual use of the wallets, there is no reasonable way to rank one assumption as "better" than the other. It is also not useful because our taxonomy aims to work for *every* assumption within one of the categories. Therefore, even if a kind of ordering within the considered category could be established, it cannot easily be extended to new/other assumptions. For this reason, rather than providing orderings within the categories, we point out factors which influence the individual quality of assumptions. Thereby, these factors might both improve or downgrade the quality. We will discuss, by way of example, what practical consequences the factors can have, guided by whether investigative measures are lawful regarding whether the degree of suspicion necessary for the investigative measures could be established. However, we point out that our taxonomy should only be understood as an argumentation tool, which is why we cannot reach a conclusive assessment. The reason for this is that the assessment of whether the necessary degree of suspicion has been reached must be decided by legal practitioners on a case-by-case basis. The need for a normative assessment on a case-by-case basis results from the facts of the case, which can be highly complex and diverse. In addition, circumstances beyond blockchain transactions potentially affect the interpretation of results obtained from blockchain analyses.

## 5.1 User Behaviour Assumptions

The most unreliable assumptions are in general user behaviour assumptions. There are two main reasons for this. First, they are subject to change as user behaviour changes. Second, assessing their reliability would require ground-truth data on user behaviour at (a) specific point(s) in time which depends on the use case.

| *(Main) assumption(s)* | Affected | *Type* | *Exploited weakness fixed* | *Relevance* |
|---|---|---|---|---|
| Multi-Input | BTC, ZEC | User Behaviour | NO | HIGH |
| Change-Address | BTC, ZEC | User Behaviour | NO | HIGH |
| No-Proxy | BTC | User Behaviour | NO | LOW |
| Unique Entry-Nodes / No-Collision | BTC | Statistical | NO | LOW |
| Miner-Payout | ZEC | User Behaviour | NO | LOW |
| Value-Input-Output | ZEC | User Behaviour | NO | MEDIUM |
| Fingerprinting | ZEC | User Behaviour | NO | MEDIUM |
| Response-Time | ZEC | Protocol | YES | LOW |
| Wallet-Communication | XMR | Protocol | YES | LOW |
| No Double-Spending | XMR | Computational Hardness | PARTLY | MEDIUM |
| Multi-Output | XMR | Statistical | NO | MEDIUM |
| Newest-Account | XMR | Statistical | YES | LOW |

**Table 2.** Overview over the assumptions with indication of their practical relevance and whether the weaknesses exploited by the corresponding attacks have been fixed. A high relevance means that the attack is very likely actually employed by law enforcement; medium means that it is not known whether the attack is used but we see potential use-cases; low means that we do not see a potential use-case because the exploited weakness has been fixed or the attack would require too much effort. For a general overview on reliability of the assumptions see Section 5.

The changing nature of user behaviour assumptions can be best seen by the very different treatment of the multi-input assumption (see Section 4.1.1) and change-address heuristic (see Section 4.1.2). Although in the past it could be argued that the multi-input assumption was reasonable, this can no longer be done against the background of false positives in case heuristics are applied to CoinJoin transactions. In the case of change-address heuristics, one assumption is always that change exists at all. On top of that, the open-source Bitcoin analysis software BlockSci implements ten different change-address heuristics [33] which depend on the client(s) used. This makes the heuristics heavily dependent upon the assumptions about the usage of the corresponding clients.

Obtaining the requisite ground-truth data for user behaviour in cryptocurrencies is hard [25]. The existing work of user studies [37] or the development of user mental models [40] does not solve the problem of missing ground-truth data as they try to answer different research questions. They do not focus on obtaining any ground-truth data to assess the reliability of user behaviour assumptions in the context of cryptocurrency anonymity. Thus, their results only permit the drawing of marginal and very limited conclusions about user behaviour in the context of this work. Besides that, the participant population in user studies might not reflect the population that is actually targeted by the attacks surveyed here. It is particularly difficult to find representative participant populations that can be trusted and provide the necessary information to establish a ground truth. For example, criminals might behave very differently from users who are willing to participate in user studies, however criminals are usually the ones targeted by law enforcement investigations. The conclusion is that it is crucial to treat user behaviour assumptions with care by making them explicit and discussing whether they are reasonable every time anew.

**Factors:** 1) Ground truth: For the quality of user behaviour assumptions, ground truth is a crucial factor. We denote with this term insights into actual user behaviour, which have to be gained by other means than using the assumption it is used for. 2) False positive detection: Parallel to ground truth, the detection of false positives is important, especially if no ground truth data is available. This can be seen in the example of the multi-input assumption in Bitcoin. While there is often no ground truth, a perfect detection of CoinJoin transactions would remove all false positives produced by them and therefore improve the quality of attacks using the multi-input assumption. 3) Additional information gained from other sources: If a specific user is targeted it might be possible to improve the quality by some additional information gained from, for example, a forum post where a Bitcoin address is posted together with a (user)name. Another example would be some information about whether users exchanged secret keys offline. 4) Protocol-induced behaviour: It describes user behaviour that is due to properties of the protocol. An example of this are change outputs. Change is something that usually occurs naturally as the input amount of a transaction is consumed entirely. A characteristic of protocol-induced behaviour is that it needs active engagement to deviate from, for example, the decision not to generate change. Thus, assumptions based on protocol-induced behaviour can in principle be considered

more reliable than those based on non-protocol-induced behaviour.

**Practical Consequences:** As long as there is no ground-truth data, the findings of attacks based on user behaviour assumptions might only establish the degree of suspicion which is necessary for less intense investigative measures such as requesting personal data form third parties. Without ground truth, other factors are needed for more intensive measures, such as that the user behaviour was protocol induced. A prerequisite for establishing a degree of suspicion at all is that a false positive control has been performed, which will be explained using the example of the multi-input heuristic. If it is known that the multi-input heuristic was applied to CoinJoin transactions when identifying a wallet, no degree of suspicion for any investigative measures might be established solely on the basis of the obtained address clusters, as false positives will necessarily occur in this case. Consequently, in order to establish any degree of suspicion at all, it must be ruled out that the multi-input heuristic was applied to CoinJoin transactions according to the current state of the art. The CoinJoin detection of the open-source cryptocurrency forensic tool BlockSci [33] is tailored to specific CoinJoin transactions [8]. As a consequence, CoinJoin transactions of recent services such as Wasabi [65] or Samourai [55] will in general not be detected which might result in false positives when applying BlockSci's multi-input heuristic. As a consequence, BlockSci must be extended by methods to detect Wasabi and Samourai CoinJoin transactions such as the ones proposed by Stockinger et al. [58]. Even with these extensions, it can not be ruled out that CoinJoin transactions from lesser known or custom services remain undetected. While this might be unsatisfying, it still satisfies that false positives have been excluded as far as possible according to the current state of the art. Thus, to establish a degree of suspicion at all solely based on some blockchain analyses, a minimum requirement must be to detect false positives according to the current state of the art. If there is no false positive detection, there must be other blue (additional) indications to establish the necessary degree of suspicion, even for less intensive investigative measures.

## 5.2 Protocol Assumptions

Protocol assumptions themselves can be seen as very reliable as random behaviour within a protocol is usually very limited. Thus, if a specific protocol, for example, the implementation of communication between a wallet and a full node, shows specific behaviour, this behaviour can always be found. While protocol assumptions seem very reliable, it must not be forgotten that, in practice, they always come with the drawback that they might become outdated and are only useful for data specifically produced within the correct time frame. Likewise, if a currency allows for several protocols to be used, for example, when an update to a newer encryption is done over several months, there will always be an accompanying user behaviour assumption. This assumption is about the concrete use of the protocol and has to be taken into account when applying an attack. We note that these types of user behaviour assumptions are at least sometimes of a more reliable nature as there might be concrete identifiers of which protocol was used. For these reasons, we declare protocol assumptions to be slightly more unreliable than computational hardness assumptions, however still more reliable than plain user behaviour assumptions.

**Factors:** 1) Spread of technology: If, for example, the assumption relies on the special behaviour of wallets, it is necessary to know about the spread of these wallets, i. e. if a wallet is "the" standard wallet, the assumption might be very reliable, but for some rarely used wallets it is not. 2) Delays: In particular for assumptions about timing patterns, delays within the network have to be taken into account as they strongly influence the patterns observed. 3) Packet loss: Similar to timing patterns, for communication patterns it is important to consider how many packets were lost in communication, as this might, for example, increase the number of communication rounds between node and wallet.

**Practical Consequences:** Protocol assumptions have presumably not played a role in blockchain forensics so far, as they have also played a rather subordinate role in research and the corresponding attacks have been fixed by software updates (see Table 3). Consequently, very little can be said about the practical consequences. Nevertheless, the following considerations are intended to assist in the event that protocol assumptions become relevant to investigations in the future. If analyses are based on implementation details of a certain wallet, then the circulation of the wallet determines the established degree of suspicion. If the circulation is not known and other popular wallets exist that stand up to these analyses, then only less intense investigative measures might be employed. On the other hand, if the cryptocurrency in question is less widespread and only one wallet exists, then the corresponding analyses might be sufficient even for intensive investigative measures.

## 5.3 Computational Hardness Assumptions

Computational hardness assumptions are the most reliable ones as they are quite stable, i. e. not subject to change as user behaviour or protocol assumptions are. Besides that, they have usually been well established and thus studied for a long time. Computational hardness assumptions do not require any ground-truth data. Finding ground-truth data would mean ascertaining specific examples where the assumption does not hold and immediately breaking the security of possibly the entire system and many other systems which use the same parameters, groups, etc. If an attack is solely based on a well-established computational hardness assumption, the findings might have high probative value and thus establish the required degree of suspicion even for intense investigative measures.

**Factors:** Acceptance within the research community: To assess the reliability of a hardness assumption, for example, the discrete logarithm assumption, it has to be taken into account how well established the assumption is within the community. Thereby, the assumption can be considered accepted if it has been studied intensively without proving it wrong and/or the assumption is utilized in many different protocols. In contrast, new assumptions that are "invented" for a specific protocol should be handled with great caution as Goldwasser and Kalai [27] already pointed out.

**Practical Consequences:** If an analysis solely relies on well-established hardness assumptions, the degree of suspicion required for any investigative measures might be established. On the other hand, if the assumption is relatively new and cannot be said to be "accepted" by the community, then only less intensive measures such as requesting personal data from third parties might be lawful. However, such new assumptions might not be sufficient for pre-trial detention.

## 5.4 Statistical Assumptions

Statistical assumptions differ in two dimensions from the aforementioned three. First, by the very nature of statistical assumptions, we can (in theory) asses the reliability exactly, as probabilities can be computed. Second, the variance in reliability is very high. This might sound counter-intuitive, but the fact that we can compute something concrete solely means that we can compute how "good" or "bad" an assumption is, which does not strengthen the assumption itself. For example, the newest suggestion in Monero for sampling rings favours recent accounts over older ones by using a gamma distribution [43]. Doing some computations with conditional probabilities for a concrete ring provides the exact probability for the newest-account assumption (see Section 4.4.3) to be correct, but with high probability will tell us that the probability for correctness is very low. This makes it impossible to put statistical assumptions into a relation concerning reliability with the other three assumptions.

**Factors:** 1) Correct use of protocol: Statistical assumptions can only be made if the protocol specifies some probability distribution or some behaviour. Therefore, they have to rely on the correct use of the protocol. 2) Likelihood of preconditions: In addition to the correct use of protocol, the likelihood of preconditions needs to be taken into account, i. e. that some events happen at all and if they happen how likely it is that they are the desired events. The multi-output assumption in Monero states that if two different outputs from the same transaction $tx_1$ occur in two different rings of the same transaction $tx_2$ (which we call double link for now), then they are likely to be the real spenders of that transaction as solely being part of the decoy set is very unlikely. While it is true that it is very unlikely that they are part of the decoy set, the conclusion that they must be the real spender is problematic. The probability that a double link can be observed is the sum of the probabilities that it happens by chance and that someone does it on purpose. The latter probability not only depends on the behaviour of the owner of these two outputs but also on the probability that a transaction contains two outputs actually belonging to the same owner. Thus, the probability that a double link happens on purpose might still be smaller than that it happening by chance. Additionally, according to the decoy selection algorithm, such a double link is more likely to happen for a newer $tx_1$ by chance. These two are thus preconditions that have to be taken into account when evaluating the quality of that statistical assumption.

**Practical Consequences:** The multi-output assumption for Monero states that it is very unlikely that two outputs of the same transaction appear as inputs in another transaction by chance. The assumption further states that the appearance of two such outputs would imply that they belong to the same entity. In probabilities, this would mean the following. Let $MO$ be the event that a transaction spends to two different addresses but the same entity. Let $RMI$ be the event that two outputs of the same transaction appear as inputs of another transaction by chance and $PMI$ the probability that they are used intentionally in the same transaction. Note that $RMI$ does not exclude the case that the two outputs actually belong to

the same entity. Let $MI$ be the event that we can observe two outputs from the same transaction in the input ring of another transaction. The assumption states that $\Pr[RMI|MI] < \Pr[PMI|MI]$. If the assumption is wrong, it cannot establish a link between the two addresses. As a consequence, no investigative measures can be based on such a link. This means that in the end, the assessment of the probabilities determines whether the degree of suspicion could be established or not.

### 5.5 Consequences Using Wall Street Market as an Example

We show how our taxonomy can be applied to real criminal cases, using the Wall Street Market (WSM) investigations as an example, and more specifically the example of one of its alleged administrators called "Frost" [63]. In the case of Frost, blockchain analyses were used, at the end of which personal data was requested from a Bitcoin payment processing company (BPPC). Subsequent investigative measures made possible by obtaining the personal data, such as a search of Frost's premises, would require a degree of suspicion and we will discuss what reasons there are to believe whether it was established or not. First, we summarize the analysis that preceded the request. There were mainly four wallets reported in the analysis, wallets $W1$, $W2$, $W4$ and $W5$ (numbering of the wallets based on the one used in the criminal complaint [63]). Each wallet was detected by the US Postal Service (USPS) using proprietary software [63, p. 20, footnote 2]. Wallets $W1$, $W4$ and $W5$ were found to be origin of payments to various services via BPPC. Prior to the payments, the corresponding Bitcoins were supposedly mixed via a commercial mixing service. However, the USPS stated that they reversed the mixing ("de-mix"). The request to the BPPC for personal data on the payments was finally conducted because the wallets that funded wallets $W1$, $W4$ and $W5$ were associated with WSM. One of these funding wallets was wallet $W2$. In other words, four analytical steps were involved in the analysis: *identifying wallets*, *detecting payments between wallets*, *de-mixing* and the *association of wallets* with darknet marketplaces. All these steps are based on user behaviour assumptions.

Even though it is unclear how exactly the proprietary software utilized by the USPS works, we can at least assume that it employs the most common technique, namely address clustering based on the multi-input heuristic. As there is no meaningful ground truth data for the multi-input heuristic so far, it is not possible to argue with ground truth as a factor. Furthermore, we do not know whether any form of false positive detection has taken place, for example, whether it was excluded that the heuristic was applied to CoinJoin transactions. On the other hand, at least for wallet $W2$ there might be additional information that confirms the address clustering, which was obtained in the course of a seizure of another darknet marketplace [63, p. 28 f.]. This marketplace had already been under independent investigation before. For the other wallets, no such additional information is known. Therefore, it would be possible that there were only some addresses in wallets $W1$, $W4$ and $W5$ that could be linked to WSM. The addresses that could not be linked to WSM could belong to unsuspicious third parties. As a result, the request to the BPPC could reveal personal data of an unsuspicious third party. Only based on the results of the blockchain analysis, the premises of this party may not be searched. The reason is that in the case of a search, the suspicion must be particularized against the person being searched. In summary, the reliability of clustering determines the individualization of suspicion, with individualization being a prerequisite for the lawfulness of certain investigative measures.

## 6 Conclusion

We demonstrated that cryptocurrency forensic methods based on uncertain assumptions might cause legal issues. One example is the question whether further investigative measures are lawful that are based solely on findings from cryptocurrency analysis using uncertain assumptions. To address these issues, we proposed a taxonomy in which we categorized common assumptions underlying deanonymization attacks found in research papers. We elaborated that in general assumptions based on user behaviour are the least reliable, while at the same time they are amongst those with the highest potential to be practically relevant. The user behaviour assumptions include the multi-input and the change heuristic, which we believe are currently the most relevant in practice, as there is strong indication that both are already relied upon by law enforcement agencies. As the reliability of forensic methods must always be evaluated on a case-by-case basis, we complement our taxonomy with factors that serve to argue reliability on that basis. In the case of the multi-input heuristic, one important factor is the detection of false positives introduced by CoinJoin transactions that by design break the heuristic.

## Acknowledgments

We thank Catherine Nowland, Kevin Pike, all anonymous reviewers of this work and our shepherd, Fabio Massacci, for their corrective feedback and support. Work was supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) as part of the Research and Training Group 2475 "Cybercrime and Forensic Computing" (grant number 393541319/GRK2475/1-2019).

## References

[1] E. Androulaki, G. Karame, M. Roeschlin, T. Scherer, and S. Capkun. "Evaluating User Privacy in Bitcoin." In: *FC 2013*. Ed. by A.-R. Sadeghi. Vol. 7859. LNCS. Springer, Heidelberg, Apr. 2013, pp. 34–51. DOI: 10.1007/978-3-642-39884-1_4.

[2] A. Biryukov and D. Feher. *Deanonymization of Hidden Transactions in Zcash*.

[3] A. Biryukov and D. Feher. "Privacy and Linkability of Mining in Zcash." In: *2019 IEEE Conference on Communications and Network Security*. 2019, pp. 118–123.

[4] A. Biryukov, D. Feher, and G. Vitto. "Privacy Aspects and Subliminal Channels in Zcash." In: *ACM CCS 2019*. Ed. by L. Cavallaro, J. Kinder, X. Wang, and J. Katz. ACM Press, Nov. 2019, pp. 1795–1811. DOI: 10.1145/3319535.3345663.

[5] A. Biryukov, D. Khovratovich, and I. Pustogarov. "Deanonymisation of Clients in Bitcoin P2P Network." In: *ACM CCS 2014*. Ed. by G.-J. Ahn, M. Yung, and N. Li. ACM Press, Nov. 2014, pp. 15–29. DOI: 10.1145/2660267.2660379.

[6] N. Bitansky, A. Chiesa, Y. Ishai, R. Ostrovsky, and O. Paneth. "Succinct Non-interactive Arguments via Linear Interactive Proofs." In: *TCC 2013*. Ed. by A. Sahai. Vol. 7785. LNCS. Springer, Heidelberg, Mar. 2013, pp. 315–333. DOI: 10.1007/978-3-642-36594-2_18.

[7] *Bitcoin Network*. URL: https://en.bitcoin.it/wiki/Network (visited on 01/15/2021).

[8] *BlockSci Issue #174 - coinjoin heuristic bug?* URL: https://github.com/citp/BlockSci/issues/174 (visited on 02/14/2022).

[9] *BlockSci Source Code*. URL: https://github.com/citp/BlockSci/blob/master/src/cluster/cluster_manager.cpp (visited on 07/22/2021).

[10] Bundesgerichtshof (BGH) (Federal Court of Justice), BGH: Illegales Bitcoinschürfen, Neue Zeitschrift für Strafrecht 2018, 401 (m. Anm. Safferling).

[11] Bundesverfassungsgericht (BVerfG) (Federal Constitutional Court), 2 BvR 766/03, Jan. 23, 2004.

[12] *Chainalysis Data*. URL: https://www.chainalysis.com/chainalysis-data/ (visited on 12/07/2020).

[13] Chainalysis Team. *Introducing Investigation and Compliance Support for Dash and Zcash*. June 8, 2020. URL: https://blog.chainalysis.com/reports/introducing-chainalysis-investigation-compliance-support-dash-zcash/ (visited on 02/12/2022).

[14] *Chainalysis*. URL: https://www.chainalysis.com/ (visited on 12/07/2020).

[15] C. Champod and J. Vuille. "Scientific Evidence in Europe – Admissibility, Evaluation and Equality of Arms." In: *International Commentary on Evidence* 9.1 (2011). DOI: doi:10.2202/1554-4567.1123. URL: https://doi.org/10.2202/1554-4567.1123.

[16] *CipherTrace*. URL: https://ciphertrace.com/ (visited on 12/07/2020).

[17] M. Conti, E. S. Kumar, C. Lal, and S. Ruj. "A survey on security and privacy issues of bitcoin." In: *IEEE Communications Surveys & Tutorials* 20.4 (2018), pp. 3416–3452.

[18] *Dash*. URL: https://www.dash.org/ (visited on 06/16/2020).

[19] Department of Justice - Office of Public Affairs. *Three Germans Who Allegedly Operated Dark Web Marketplace with Over 1 Million Users Face U.S. Narcotics and Money Laundering Charges*. Department of Justice, 2019. URL: https://www.justice.gov/opa/pr/three-germans-who-allegedly-operated-dark-web-marketplace-over-1-million-users-face-us (visited on 12/07/2020).

[20] D. Deuber and D. Schröder. "CoinJoin in the Wild." In: *Computer Security – ESORICS 2021*. Ed. by E. Bertino, H. Shulman, and M. Waidner. Cham: Springer International Publishing, 2021, pp. 461–480. ISBN: 978-3-030-88428-4.

[21] *Elliptic*. URL: https://www.elliptic.co/ (visited on 12/07/2020).

[22] European Cybercrime Center (EC3). *Internet Organised Crime Threat Assessment 2019*. Europol. URL: https://www.europol.europa.eu/activities-services / main - reports / internet - organised - crime - threat - assessment - iocta - 2019 (visited on 02/19/2021).

[23] European Cybercrime Center (EC3). *Internet Organised Crime Threat Assessment 2020*. Europol. URL: https://www.europol.europa.eu/activities-

services / main - reports / internet - organised - crime - threat - assessment - iocta - 2020 (visited on 02/19/2021).

[24] M. Fleder, M. S. Kester, and S. Pillai. "Bitcoin transaction graph analysis." In: *arXiv preprint arXiv:1502.01657* (2015).

[25] M. Fröwis, T. Gottschalk, B. Haslhofer, C. Rückert, and P. Pesch. "Safeguarding the Evidential Value of Forensic Cryptocurrency Investigations." In: *arXiv e-prints*, arXiv:1906.12221 (2019), arXiv:1906.12221. arXiv: 1906 . 12221 [cs.CY].

[26] S. Goldfeder, H. Kalodner, D. Reisman, and A. Narayanan. "When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies." In: *Proceedings on Privacy Enhancing Technologies* 2018.4 (2018), pp. 179 –199. URL: https://content.sciendo.com/view/journals/popets/2018/4/article-p179.xml.

[27] S. Goldwasser and Y. T. Kalai. "Cryptographic Assumptions: A Position Paper." In: *TCC 2016-A, Part I*. Ed. by E. Kushilevitz and T. Malkin. Vol. 9562. LNCS. Springer, Heidelberg, Jan. 2016, pp. 505–522. DOI: 10.1007/978-3-662-49096-9_21.

[28] C. Haan. *Blockchain Analysis Spending by US Government Agencies Has Tripled in 2018*. URL: https://www.crowdfundinsider.com/2018/09/139486-blockchain-analysis-spending-by-us-government-agencies-has-tripled-in-2018/ (visited on 12/07/2020).

[29] M. Harrigan and C. Fretter. "The unreasonable effectiveness of address clustering." In: *2016 Intl IEEE Conferences on Ubiquitous Intelligence & Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld)*. IEEE. 2016, pp. 368–373.

[30] B. Haslhofer, R. Stütz, M. Romiti, and R. King. "GraphSense: A General-Purpose Cryptoasset Analytics Platform." In: *Arxiv pre-print* (2021). URL: https://arxiv.org/abs/2102.13613.

[31] *How German and US authorities took down the owners of darknet drug emporium Wall Street Market*. URL: https://techcrunch.com/2019/05/03/how-german-and-us-authorities-took-down-the-owners-of-darknet-drug-emporium-wall-street-market/ (visited on 09/17/2020).

[32] M. Jourdan, S. Blandin, L. Wynter, and P. Deshpande. "Characterizing entities in the bitcoin blockchain." In: *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE. 2018, pp. 55–62.

[33] H. A. Kalodner, M. Möser, K. Lee, S. Goldfeder, M. Plattner, A. Chator, and A. Narayanan. "BlockSci: Design and applications of a blockchain analysis platform." In: *USENIX Security 2020*. Ed. by S. Capkun and F. Roesner. USENIX Association, Aug. 2020, pp. 2721–2738.

[34] G. Kappos, H. Yousaf, M. Maller, and S. Meiklejohn. "An Empirical Analysis of Anonymity in Zcash." In: *USENIX Security 2018*. Ed. by W. Enck and A. P. Felt. USENIX Association, Aug. 2018, pp. 463–477.

[35] D. Kelly, R. Raines, R. Baldwin, M. Grimaila, and B. Mullins. *Exploring Extant and Emerging Issues in Anonymous Networks: A Taxonomy and Survey of Protocols and Metrics*. 2012. DOI: 10.1109/SURV.2011.042011.00080.

[36] P. Koshy, D. Koshy, and P. McDaniel. "An Analysis of Anonymity in Bitcoin Using P2P Network Traffic." In: *FC 2014*. Ed. by N. Christin and R. Safavi-Naini. Vol. 8437. LNCS. Springer, Heidelberg, Mar. 2014, pp. 469–485. DOI: 10.1007/978-3-662-45472-5_30.

[37] K. Krombholz, A. Judmayer, M. Gusenbauer, and E. R. Weippl. "The Other Side of the Coin: User Experiences with Bitcoin Security and Privacy." In: *FC 2016*. Ed. by J. Grossklags and B. Preneel. Vol. 9603. LNCS. Springer, Heidelberg, Feb. 2016, pp. 555–580.

[38] A. Kumar, C. Fischer, S. Tople, and P. Saxena. "A Traceability Analysis of Monero's Blockchain." In: *ESORICS 2017, Part II*. Ed. by S. N. Foley, D. Gollmann, and E. Snekkenes. Vol. 10493. LNCS. Springer, Heidelberg, Sept. 2017, pp. 153–173. DOI: 10.1007/978-3-319-66399-9_9.

[39] M. C. Kus Khalilov and A. Levi. "A Survey on Anonymity and Privacy in Bitcoin-Like Digital Cash Systems." In: *IEEE Communications Surveys Tutorials* 20.3 (2018), pp. 2543–2585. DOI: 10.1109/COMST.2018.2818623.

[40] A. Mai, K. Pfeffer, M. Gusenbauer, E. Weippl, and K. Krombholz. "User Mental Models of Cryptocurrency Systems-A Grounded Theory Approach." In: *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*. 2020, pp. 341–358.

[41] G. Maxwell. *CoinJoin: Bitcoin privacy for the real world*. 2013. URL: https://bitcointalk.org/index.php?topic=279249 (visited on 08/12/2019).

[42] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage. "A fistful of bitcoins: characterizing payments among men with no names." In: *Proceedings of the 2013 conference on Internet measurement conference*. ACM. 2013, pp. 127–140.

[43] *Monero source code commit using gamma distribution*. URL: https://github.com/monero-project/monero/commit/34d4b798d44250f64fdcac61439a86afa8607c3b (visited on 02/26/2021).

[44] *Monero*. URL: https://www.getmonero.org/ (visited on 08/12/2019).

[45] M. Möser, K. Soska, E. Heilman, K. Lee, H. Heffan, S. Srivastava, K. Hogan, J. Hennessey, A. Miller, A. Narayanan, and N. Christin. "An Empirical Analysis of Traceability in the Monero Blockchain." In: *PoPETs* 2018.3 (July 2018), pp. 143–163. DOI: 10.1515/popets-2018-0025.

[46] S. Nakamoto. *Bitcoin: A peer-to-peer electronic cash system*. 2008.

[47] T. Neudecker and H. Hartenstein. "Could Network Information Facilitate Address Clustering in Bitcoin?" In: *FC 2017 Workshops*. Ed. by M. Brenner, K. Rohloff, J. Bonneau, A. Miller, P. Y. A. Ryan, V. Teague, A. Bracciali, M. Sala, F. Pintore, and M. Jakobsson. Vol. 10323. LNCS. Springer, Heidelberg, Apr. 2017, pp. 155–169.

[48] M. Ober, S. Katzenbeisser, and K. Hamacher. "Structure and anonymity of the bitcoin transaction graph." In: *Future internet* 5.2 (2013), pp. 237–250.

[49] T. P. Pedersen. "Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing." In: *CRYPTO'91*. Ed. by J. Feigenbaum. Vol. 576. LNCS. Springer, Heidelberg, Aug. 1992, pp. 129–140. DOI: 10.1007/3-540-46766-1_9.

[50] J. Quesnelle. "On the linkability of Zcash transactions." In: *CoRR* abs/1712.01210 (2017). arXiv: 1712.01210. URL: http://arxiv.org/abs/1712.01210.

[51] F. Reid and M. Harrigan. "An analysis of anonymity in the bitcoin system." In: *Security and privacy in social networks*. Springer, 2013, pp. 197–223.

[52] D. Ron and A. Shamir. "Quantitative Analysis of the Full Bitcoin Transaction Graph." In: *FC 2013*. Ed. by A.-R. Sadeghi. Vol. 7859. LNCS. Springer, Heidelberg, Apr. 2013, pp. 6–24. DOI: 10.1007/978-3-642-39884-1_2.

[53] T. Ruffing, P. Moreno-Sanchez, and A. Kate. "CoinShuffle: Practical Decentralized Coin Mixing for Bitcoin." In: *ESORICS 2014, Part II*. Ed. by M. Kutylowski and J. Vaidya. Vol. 8713. LNCS. Springer, Heidelberg, Sept. 2014, pp. 345–364. DOI: 10.1007/978-3-319-11212-1_20.

[54] T. Ruffing, P. Moreno-Sanchez, and A. Kate. "P2P Mixing and Unlinkable Bitcoin Transactions." In: *NDSS 2017*. The Internet Society, 2017.

[55] *Samourai wallet*. URL: https://samouraiwallet.com (visited on 11/29/2021).

[56] I. R. Service. *Pilot IRS Cryptocurrency Tracing Award Notice*. URL: https://sam.gov/opp/5ab94eae1a8d422e88945b64181c6018/view (visited on 02/12/2022).

[57] M. Spagnuolo, F. Maggi, and S. Zanero. "BitIodine: Extracting Intelligence from the Bitcoin Network." In: *FC 2014*. Ed. by N. Christin and R. Safavi-Naini. Vol. 8437. LNCS. Springer, Heidelberg, Mar. 2014, pp. 457–468. DOI: 10.1007/978-3-662-45472-5_29.

[58] J. Stockinger, B. Haslhofer, P. Moreno-Sanchez, and M. Maffei. *Pinpointing and Measuring Wasabi and Samourai CoinJoins in the Bitcoin Ecosystem*. 2021. arXiv: 2109.10229 [cs.CR].

[59] A. E. Taslitz. "WHAT IS PROBABLE CAUSE, AND WHY SHOULD WE CARE?: THE COSTS, BENEFITS, AND MEANING OF INDIVIDUALIZED SUSPICION." In: *Law and Contemporary Problems* 73.3 (2010), pp. 145–210. ISSN: 00239186. URL: http://www.jstor.org/stable/25766403.

[60] S. C. Thaman. *Comparative Criminal Procedure: A Casebook Approach*. 2nd ed. Carolina Academic Press Comparative Law Series, 2008.

[61] F. Tramèr, D. Boneh, and K. Paterson. "Remote Side-Channel Attacks on Anonymous Transactions." In: *USENIX Security 2020*. Ed. by S. Capkun and F. Roesner. USENIX Association, Aug. 2020, pp. 2739–2756.

[62] J. I. Turner. "Limits on the search for truth in criminal procedure: a comparative view." In: *Comparative Criminal Procedure*. Cheltenham, UK: Edward Elgar Publishing, 2016.

[63] United States Discrict Court. *Criminal Complaint - United States of America v. Tibo Lousee, Klaus-Martin Frost, and Jonathan Kalla - Case No. 19MJ1843*. 2019. URL: https://www.justice.gov/

opa/press-release/file/1159706/download (visited on 12/07/2020).

[64] *WalletExplorer*. URL: https://www.walletexplorer.com/info (visited on 07/22/2021).

[65] *Wasabi wallet*. URL: https://wasabiwallet.io (visited on 11/29/2021).

[66] D. A. Wijaya, J. Liu, R. Steinfeld, D. Liu, and T. H. Yuen. "Anonymity Reduction Attacks to Monero." In: *Information Security and Cryptology*. Ed. by F. Guo, X. Huang, and M. Yung. Cham: Springer International Publishing, 2019.

[67] D. A. Wijaya, J. K. Liu, R. Steinfeld, D. Liu, and J. Yu. "On The Unforkability of Monero." In: *ASIACCS 19*. Ed. by S. D. Galbraith, G. Russello, W. Susilo, D. Gollmann, E. Kirda, and Z. Liang. ACM Press, July 2019, pp. 621–632. DOI: 10.1145/3321705.3329823.

[68] D. Wijaya, J. Liu, R. Steinfeld, and D. Liu. "Monero Ring Attack: Recreating Zero Mixin Transaction Effect." In: Aug. 2018, pp. 1196–1201. DOI: 10.1109/TrustCom/BigDataSE.2018.00165.

[69] J. Yu, M. H. A. Au, and P. Esteves-Verissimo. "Re-Thinking Untraceability in the CryptoNote-Style Blockchain." In: *2019 IEEE 32nd Computer Security Foundations Symposium (CSF)*. 2019, pp. 94–9413.

[70] J. Yu, M. H. A. Au, and P. J. E. Veríssimo. "Re-Thinking Untraceability in the CryptoNote-Style Blockchain." In: *CSF 2019Computer Security Foundations Symposium*. Ed. by S. Delaune and L. Jia. IEEE Computer Society Press, 2019, pp. 94–107. DOI: 10.1109/CSF.2019.00014.

[71] Z. Yu, M. H. Au, J. Yu, R. Yang, Q. Xu, and W. F. Lau. "New Empirical Traceability Analysis of CryptoNote-Style Blockchains." In: *FC 2019*. Ed. by I. Goldberg and T. Moore. Vol. 11598. LNCS. Springer, Heidelberg, Feb. 2019, pp. 133–149. DOI: 10.1007/978-3-030-32101-7_9.

[72] *Zcash*. URL: https://z.cash/ (visited on 08/12/2019).

# A Generalizability and Scope of the Legal Issues

Uncertain assumptions underlying forensic methods might raise legal issues at any stage of the criminal proceedings and independent of the jurisdiction. In the pre-trial stages, the uncertainty could be too high to establish the required degree of suspicion for investigative measures. During the actual trial, the uncertainty could either hinder the scientific evidence from being admitted or decrease its probative value, which affects its assessment. These issues are not specific to a particular jurisdiction but apply generally in both the common law and the continental European system as elaborated in the following.

**Pre-Trial Stages** The most important element in the pre-trial stages are investigations in order to obtain evidence. As investigative measures, such as searches or even arrests, interfere with the fundamental and human rights of the persons concerned, they require a legal basis. This basis defines the prerequisites that must be satisfied before investigative measures can be conducted. One of those requirements is always some degree of suspicion [11, 60]. The effects that uncertain assumptions exert on this degree will be illustrated using the example of a search of a suspect's premises,[4] referring to England and the US for the common law system and to Germany for the continental European system. In the US, the Fourth Amendment stipulates that the requisite degree of suspicion for searches and seizures is "probable cause", while in England it is "reasonable grounds" pursuant to Section 8 (1) Police and Criminal Evidence Act 1984 (PACE). In Germany, the required degree of suspicion is "sufficient factual indications" as enshrined in Section 102 German Code of Criminal Procedure (*Strafprozessordnung*; StPO) in conjunction with Section 152 (2) StPO. These degrees of suspicion have in common that they must be sufficiently individualized [59]. This means that the mere suspicion that a criminal offense was committed is not sufficient; rather, the suspicion must be individualized with respect to the person whose premises are to be searched. This individualization is precisely where uncertain assumptions become pertinent. Cryptocurrency addresses associated with a crime could be assigned to a person by means of address clustering and attribution tagging. If that person's premises are to be searched, the uncertainty in the assumptions determines how much the suspicion against that person is individualized. With maximum uncertainty in the assumption, no individualization is possible, which is why a search in such a case would be unlawful. As a consequence, any evidence obtained in the course of an unlawful search might be rendered inadmissible under exclusionary rules that exist not only in common law but also in most civil law jurisdictions [62]. In general, despite the different prerequisites regarding

**4** For the sake of clarity, the search of non-suspects' premises is ignored at this point, although similar considerations apply.

| *Attack/Heuristic* | Used | *(Main) assumption(s)* | Affected | *Type* | *Fixed* |
|---|---|---|---|---|---|
| Multi-Input | [1, 2, 24, 29, 32, 34, 42, 47, 48, 51, 52, 57] | Multi-Input | BTC, ZEC | User Behaviour | NO: would require to prevent multiple inputs |
| Change-Address | [1, 34, 42, 47, 57] | Change-Address | BTC, ZEC | User Behaviour | NO: would require to change the transaction structure |
| Cluster-Intersection | [26] | depends on the actually used address-clustering heuristics | BTC | User Behaviour | NO: would at least require that the restriction of many mixing services to a fixed value be removed [20] |
| Relay-Pattern Attacks | [5, 36] | No-Proxy / Unique Entry-Nodes / No-Collision | BTC | User Behaviour/Statistical | NO: would require to change Bitcoin's P2P network |
| Mining Pool Heuristics | [3, 34] | Miner-Payout | ZEC | User Behaviour | NO: would require different user behaviour |
| Value-Input-Output | [4, 34, 50] | Value-Input-Output | ZEC | User Behaviour | Partly fixed: No founder's reward anymore by design |
| Fingerprinting | [4] | Fingerprinting | ZEC | User Behaviour | NO: would require to hide the value |
| Response-Time | [61] | Response-Time | ZEC | Protocol | YES: Software update |
| Wallet-Communication | [61] | Wallet-Communication | XMR | Protocol | YES: Software update |
| Zero-Mixin | [38, 45, 69, 71] | No Double-Spending | XMR | Computational Hardness | YES: Protocol update |
| Closed-Set | [69, 71] | No Double-Spending | XMR | Computational Hardness | NO: would require to change user behaviour or signature generation |
| Multi-Output | [38] | Multi-Output | XMR | Statistical | NO: would require different user behaviour, unclear if fix needed |
| Newest-Account | [38, 45] | Newest-Account | XMR | Statistical | YES: Monero suggests a different distribution |

**Table 3.** Overview over attacks with corresponding assumptions indicating if they have been fixed

the admissibility of evidence, in all jurisdictions there is at least a real danger of exclusion if there was no suspicion.

**Actual Trial** In the pre-trial stages, the question was whether further investigative measures could be based on the findings of cryptocurrency forensics methods and what the consequences would be if this question were to be answered in the negative. As further evidence is often found in the course of such investigations, the findings themselves play a subordinate role, if any, in the actual trial. However, it might also occur that the findings of cryptocurrency forensics methods become the direct subject of main proceedings as scientific evidence. This may be the case if they are directly employed to prove a certain element of a crime, for example, in cases of money laundering or terrorist financing, or if no further evidence was obtained during the pre-trial stages. In this case, the questions of the admissibility of such evidence and its assessment arise. US law distinguishes strictly between the admission[5] and the assessment[6] of expert evidence and provides precise rules for the former. Continental European legal systems, however, follow the principle of freedom of evidence, which means that, in general, there are no rules of admissibility and any issues that may arise in this context are addressed in the assessment of the evidence [15]. A major problem in jurisdictions where the question of admissibility is neglected is that there is often a great deal of trust placed in the expert witness who provides the scientific evidence [15]. The problem became evident in a recent decision of the German Federal Court of Justice (*Bundesgerichtshof*; BGH) where an IT expert witness not

**5** Admission refers to whether the evidence is allowed for consideration.

**6** Assessment refers to the evaluation of the evidence's probative value.

only did not have to explain his methodology but also did not have to justify the conclusions drawn in his expert testimony [10].

## B Utilizing the Taxonomy

Our survey and taxonomy are meant to provide support in the following three situations, which are at the same time the major points of this paper.

First, research in the area of cryptocurrency attacks should treat underlying assumptions with great care and whenever possible argue how reliable those assumptions are. This could be done by explicitly stating the employed assumptions and classifying them according to our taxonomy. Besides that, factors (see Section 5) that might play a relevant role in the practical application of the attacks should be discussed. The cautious handling of assumptions in research is a basic prerequisite for the following points.

Second, expert witnesses presenting any findings based on assumptions should address their reliability such that legal decision-makers can draw informed conclusions. To do this, the experts must explicitly state the assumptions and argue the reliability thereof in the light of the factors that played a role in the individual case. Thus, our taxonomy can be the basis of a common comprehensible language between expert witnesses and legal decision-makers and also be a first step towards future standardization.

Finally, law enforcement agencies should question whether any results of attacks based on uncertain assumptions really establish the necessary degree of suspicion for more intensive investigative measures. For this purpose, analysis software must not be used as a black box or pressure must be built up on the companies developing such software so that they disclose the methods employed and, if necessary, argue their reliability.

Even though substantial differences between legal systems exist, our taxonomy targets their common basis, which is the necessity to reason. Only if the points listed are complied with, can legal decision-makers take into account the uncertainty in the assumptions and reach decisions that are in accordance with the law.