INVESTIGATION · DIGITAL INVESTIGATION

Blockchain Analysis and Cryptocurrency Investigation

Cryptocurrencies were once thought to make payments untraceable, but most public blockchains record every transaction permanently and openly. This article explains how investigators trace cryptocurrency, link addresses to real people and services, and the limits and safeguards of these methods.

ModQu research article · 2,304 words · 23 references (APA 7th edition)

When Bitcoin appeared in 2009, many people assumed it would make payments anonymous and therefore ideal for crime. The reality turned out to be more complicated. Bitcoin and most similar cryptocurrencies record every transaction on a public ledger, the blockchain, that anyone can download and examine. Users are identified not by name but by addresses, long strings of characters, so the system is pseudonymous rather than anonymous. Once an address can be linked to a real person or organisation, its entire transaction history may be exposed.

This permanent public record has made blockchain analysis an important investigative method. Law enforcement agencies, tax authorities, financial regulators, exchanges, insurers and journalists use it to trace stolen funds, identify the operators of illegal markets, enforce sanctions and recover assets. This article explains how cryptocurrency transactions work from an investigator's point of view, the main techniques of blockchain analysis, how analysis is combined with other evidence, and the limits and safeguards that apply.

How the record works

In Bitcoin, the system Nakamoto (2008) described, a transaction takes coins from one or more previous transactions, the "inputs", and assigns them to one or more addresses, the "outputs". To spend coins from an address, the owner must sign the transaction with the private key that corresponds to it. Every transaction is broadcast to the network, collected into blocks and added permanently to the blockchain (Narayanan et al., 2016).

Several features of this design matter to investigators. First, the record is complete: every movement of every coin since the system began is visible. Second, it is time-stamped, so the order and approximate timing of transactions are known. Third, it is permanent, so evidence cannot be deleted by a suspect. Fourth, it records addresses rather than identities, so the central investigative problem is linking addresses to people.

Other cryptocurrencies differ. Ethereum uses accounts rather than coin outputs and supports programmable "smart contracts", which create further records of activity. Some currencies, such as Monero and Zcash, were designed specifically to hide the sender, receiver or amount. Böhme et al. (2015) describe how these technical design choices shape both the uses and the risks of different systems.

Linking addresses: clustering

Because one person may control thousands of addresses, the first analytical task is to group addresses that probably belong to the same owner. This is called clustering, and it relies on heuristics, rules of thumb derived from how wallet software typically behaves.

The most important is the **common-input heuristic**. When a transaction spends coins from several addresses at once, the sender must hold the private keys for all of them, so those addresses probably belong to the same owner. Reid and Harrigan (2013) were among the first to apply this to the Bitcoin transaction record and show how it reduced anonymity.

A second is the **change-address heuristic**. When someone pays an amount smaller than the coins they hold, the unspent remainder is usually sent back to a new address they control, called a change address. Analysts use features such as round payment amounts and newly created addresses to identify which output is the change.

Meiklejohn et al. (2013) combined these heuristics with a simple but powerful technique: they made purchases from and deposits to many known services, such as exchanges, gambling sites and online shops, so that they learned addresses belonging to each. Clustering then extended those known addresses into large groups attributed to each service. They showed that this allowed them to follow flows of money, including stolen funds, to the exchanges where thieves tried to convert them into ordinary currency. Ron and Shamir (2013) analysed the full Bitcoin transaction graph and identified patterns such as long chains of transactions used to move large sums. Harrigan and Fretter (2016) explained why clustering works so well in practice, noting that ordinary wallet behaviour and address reuse quickly merge addresses into large clusters.

Heuristics can be wrong. Some transaction types, such as joint transactions in which several users combine their inputs to obscure ownership, deliberately break the common-input rule. Wallet software varies, and a change-address guess may be mistaken. Analysts therefore treat clusters as probabilities, not facts, and test key links against other evidence.

Following the money

Once addresses are clustered and some are attributed to known entities, investigators trace flows of funds through the transaction graph. A typical trace begins at a known address, such as the one to which a ransomware victim paid, and follows the coins forward through subsequent transactions.

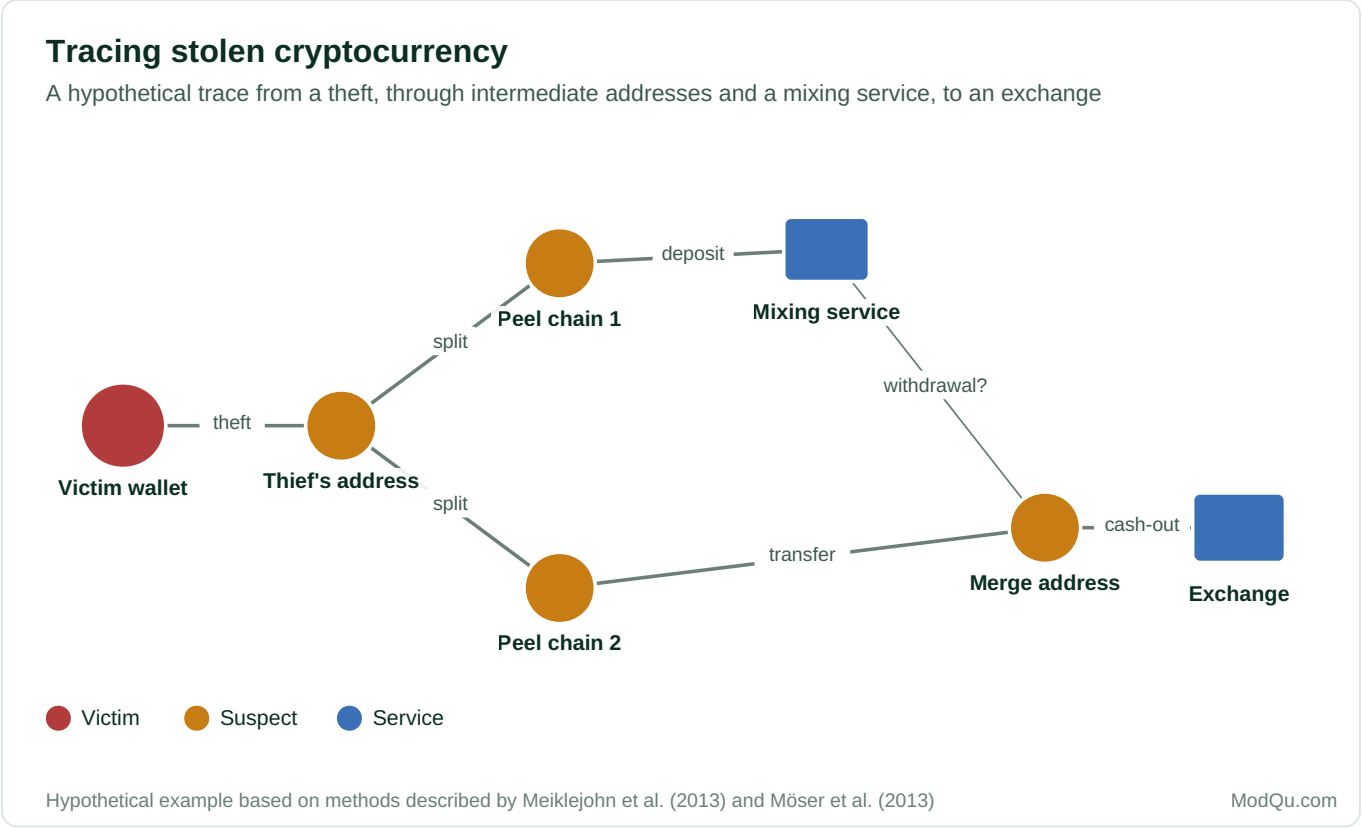


Figure 1. A simplified trace. Funds move from a victim through addresses controlled by a suspect. The link through the mixing service is uncertain, while the cash-out at a regulated exchange gives investigators a point where legal process can identify the account holder. Diagram by ModQu.

Criminals try to break the trail. Common methods include splitting funds across many addresses, sending them through long "peel chains" in which a small amount is removed at each step, converting them into other cryptocurrencies through exchange services, and using mixers, services that pool coins from many users and return different coins to each. Möser et al. (2013) tested several Bitcoin mixing services and found that some provided weak anonymity and that others simply stole the deposits. Tracing through mixers is sometimes possible by matching amounts and timing, but the confidence of such links is lower and must be stated.

The most valuable points in a trace are usually the places where cryptocurrency meets the regulated financial system. Exchanges and other virtual asset service providers in many jurisdictions must verify their customers' identities and keep records. The Financial Action Task Force (FATF), which sets international anti-money laundering standards, has extended its requirements to such providers, including the "travel rule" requiring originator and beneficiary information to accompany transfers (FATF, 2021). In the European Union, this rule applies to crypto-asset transfers under Regulation (EU) 2023/1113 (2023). When a trace reaches an exchange, investigators can use legal process to obtain the account holder's identity, records and other transactions.

Beyond the blockchain

Blockchain analysis rarely identifies a suspect by itself. It produces leads that must be combined with other evidence. Investigators link addresses to identities through exchange records, seized devices containing wallets and keys, online posts in which people publish their addresses, email and server records, and the network data of internet providers. Biryukov et al. (2014) showed that observers of the peer-to-peer network can sometimes link transactions to the internet addresses that broadcast them, and Androulaki et al. (2013) showed in a simulated university setting that behavioural patterns could identify many users even when they followed recommended privacy practices.

Several major cases illustrate this combination. Christin (2013) measured the Silk Road, an online market for illegal drugs that accepted only Bitcoin, describing its scale and structure. The market's operator, Ross Ulbricht, was arrested in 2013 after investigators connected online posts, server evidence and blockchain records, and his conviction was upheld on appeal (*United States v. Ulbricht*, 2017). In 2019, United States authorities announced the takedown of Welcome to Video, a site distributing child sexual abuse material funded by Bitcoin payments; tracing payments to exchanges led to the identification of the site's operator and hundreds of users in many countries (U.S. Department of Justice [DOJ], 2019). In 2022, investigators traced Bitcoin stolen in the 2016 hack of the Bitfinex exchange to accounts controlled by two people in New York and seized coins then worth about 3.6 billion dollars (DOJ, 2022). Greenberg (2022) gives a detailed account of how these and other investigations developed.

How much crime is involved?

Estimates of the share of cryptocurrency activity linked to crime vary greatly depending on the method. Foley et al. (2019) used clustering and network analysis to estimate that around a quarter of Bitcoin users and nearly half of Bitcoin transactions in its early years were associated with illegal activity, a share that declined as the user base grew. Commercial blockchain analysis firms publish annual estimates of illicit volume based on addresses they have attributed to criminal activity; these estimates suggest that illicit transactions now form a small percentage of total volume, though they depend on the firms' own attribution and are revised as more addresses are identified (Chainalysis, 2024). Both kinds of figure are best treated as estimates with substantial uncertainty.

Privacy coins and other obstacles

Cryptocurrencies designed for privacy make tracing much harder. Monero hides senders using "ring signatures" that mix a real input with decoys, and hides amounts and recipients. Möser et al. (2018) found that early versions of Monero were vulnerable to analysis that identified the real input in many transactions, but that later changes greatly reduced this weakness. Zcash offers optional "shielded" transactions; Kappos et al. (2018) found that most users did not use them and that patterns of use allowed much shielded activity to be linked, showing that privacy depends on how people actually use a system, not only on its design.

Other obstacles include decentralised exchanges and cross-chain "bridges" that move value between blockchains without identity checks, and services based in jurisdictions that do not cooperate with foreign investigators. Governments have responded in part with sanctions. In 2022 the United States Treasury sanctioned Tornado Cash, a mixing service built on

smart contracts on Ethereum, stating that it had been used to launder more than seven billion dollars, including funds stolen by a North Korean hacking group (U.S. Department of the Treasury, 2022). The legal basis for sanctioning software of this kind has since been contested in the courts.

Evidence and safeguards

Because blockchain analysis relies on heuristics, its results must be presented carefully. Fröwis et al. (2020) argue that forensic cryptocurrency investigations need documented, reproducible methods so that their findings can be tested in court, and that commercial tools whose methods are not disclosed raise problems for the evidential value of their results. Good practice includes recording which heuristics were used for each link, distinguishing direct links between addresses from probabilistic ones, keeping copies of the blockchain data relied on, and stating the confidence of attributions.

Seizures raise practical problems too. Cryptocurrency can be seized only by gaining control of the private keys, and keys found on devices or written down must be secured immediately, because anyone else who holds them can move the funds. Agencies need secure wallets and procedures for managing seized assets, and must account for large changes in value between seizure and disposal.

There are also privacy concerns. Because the blockchain is public and permanent, analysis can expose the financial lives of ordinary users who have done nothing wrong. Investigators and firms that hold large databases of attributed addresses should apply the same principles of necessity, proportionality and data protection that govern other financial investigations.

Tools and skills

Investigators can examine the blockchain directly using free block explorers, which display any transaction or address, or by running their own copy of the blockchain and querying it. For larger cases, most agencies use commercial analysis platforms that maintain databases of clustered and attributed addresses and draw transaction graphs automatically. These platforms save a great deal of time, but they also concentrate attribution decisions in private companies whose data and methods are only partly visible to the courts (Fröwis et al., 2020).

The skills involved are a mix of the technical and the traditional. Analysts need to understand how different wallets construct transactions, how exchanges handle deposits and withdrawals, and how smart contracts record activity. They also need the habits of a financial investigator: keeping careful records, building timelines, testing alternative explanations for each link, and recognising when a pattern could have an innocent cause. A payment to a gambling site or an exchange is not in itself evidence of wrongdoing, and an address that once received stolen funds may belong to someone who sold goods to the thief without knowing it. Many agencies now train financial investigators in blockchain analysis rather than relying only on specialists, because the cases increasingly overlap.

A short example

The following hypothetical case illustrates the process. A small business is hit by ransomware and pays the equivalent of 80,000 dollars in Bitcoin to the address in the ransom note. The business reports the payment to police, who engage a blockchain analyst.

The analyst follows the payment. The ransom address sends the coins in a single transaction to a new address, which combines them with coins from three other addresses; the common-input heuristic suggests these four addresses have one owner. Two of them received ransom payments reported by other victims, linking the case to a wider campaign. From the combined address, the funds are split along a peel chain. At several steps, small amounts are sent to deposit addresses that the analyst's tool attributes to a large exchange.

Police send a legal request to the exchange, which identifies the account holder, provides identity documents and records of logins from a particular internet address, and freezes the remaining balance. The account holder turns out to be a money mule paid to cash out funds, who identifies the person who recruited them. The analyst's report sets out each link, the heuristic used and its confidence, and notes that one branch of the funds passed through a mixing service and could not be traced reliably.

Conclusion

Public blockchains give investigators a complete, permanent and time-stamped record of transactions, which has made cryptocurrency far more traceable than early users expected. Clustering heuristics group addresses by owner, attribution links clusters to known services, and tracing follows funds to points where they meet the regulated financial system and can be connected to real identities. The methods are powerful but probabilistic, they are weakened by mixers, privacy coins and uncooperative services, and they must be combined with other evidence and documented carefully. Used with stated confidence and appropriate safeguards, blockchain analysis has become one of the most effective tools for following money in the digital economy.

References

Androulaki, E., Karame, G. O., Roeschlin, M., Scherer, T., & Capkun, S. (2013). Evaluating user privacy in Bitcoin. In A.-R. Sadeghi (Ed.), *Financial cryptography and data security* (pp. 34–51). Springer. https://doi.org/10.1007/978-3-642-39884-1_4

Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, *29*(2), 213–238. https://doi.org/10.1257/jep.29.2.213

Biryukov, A., Khovratovich, D., & Pustogarov, I. (2014). Deanonymisation of clients in Bitcoin P2P network. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (pp. 15–29). ACM. https://doi.org/10.1145/2660267.2660379

Chainalysis. (2024). *The 2024 crypto crime report*. Chainalysis.

Christin, N. (2013). Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 213–224). ACM. https://doi.org/10.1145/2488388.2488408

Financial Action Task Force. (2021). *Updated guidance for a risk-based approach: Virtual assets and virtual asset service providers*. FATF.

Foley, S., Karlsen, J. R., & Putniņš, T. J. (2019). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? *The Review of Financial Studies, 32*(5), 1798–1853. https://doi.org/10.1093/rfs/hhz015

Fröwis, M., Gottschalk, T., Haslhofer, B., Rückert, C., & Pesch, P. (2020). Safeguarding the evidential value of forensic cryptocurrency investigations. *Forensic Science International: Digital Investigation, 33*, Article 200902. https://doi.org/10.1016/j.fsidi.2019.200902

Greenberg, A. (2022). *Tracers in the dark: The global hunt for the crime lords of cryptocurrency*. Doubleday.

Harrigan, M., & Fretter, C. (2016). The unreasonable effectiveness of address clustering. In *2016 International IEEE Conferences on Ubiquitous Intelligence & Computing* (pp. 368–373). IEEE. https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0071

Kappos, G., Yousaf, H., Maller, M., & Meiklejohn, S. (2018). An empirical analysis of anonymity in Zcash. In *Proceedings of the 27th USENIX Security Symposium* (pp. 463–477). USENIX Association.

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G. M., & Savage, S. (2013). A fistful of bitcoins: Characterizing payments among men with no names. In *Proceedings of the 2013 Internet Measurement Conference* (pp. 127–140). ACM. https://doi.org/10.1145/2504730.2504747

Möser, M., Böhme, R., & Breuker, D. (2013). An inquiry into money laundering tools in the Bitcoin ecosystem. In *2013 APWG eCrime Researchers Summit* (pp. 1–14). IEEE. https://doi.org/10.1109/eCRS.2013.6805780

Möser, M., Soska, K., Heilman, E., Lee, K., Heffan, H., Srivastava, S., Hogan, K., Hennessey, J., Miller, A., Narayanan, A., & Christin, N. (2018). An empirical analysis of traceability in the Monero blockchain. *Proceedings on Privacy Enhancing Technologies, 2018*(3), 143–163. https://doi.org/10.1515/popets-2018-0025

Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system*.

Narayanan, A., Bonneau, J., Felten, E., Miller, A., & Goldfeder, S. (2016). *Bitcoin and cryptocurrency technologies: A comprehensive introduction*. Princeton University Press.

Regulation (EU) 2023/1113 of the European Parliament and of the Council of 31 May 2023 on information accompanying transfers of funds and certain crypto-assets. (2023). *Official Journal of the European Union, L 150*, 1–39.

Reid, F., & Harrigan, M. (2013). An analysis of anonymity in the Bitcoin system. In Y. Altshuler, Y. Elovici, A. B. Cremers, N. Aharony, & A. Pentland (Eds.), *Security and privacy in social networks* (pp. 197–223). Springer. https://doi.org/10.1007/978-1-4614-4139-7_10

Ron, D., & Shamir, A. (2013). Quantitative analysis of the full Bitcoin transaction graph. In A.-R. Sadeghi (Ed.), *Financial cryptography and data security* (pp. 6–24). Springer. https://doi.org/10.1007/978-3-642-39884-1_2

United States v. Ulbricht, 858 F.3d 71 (2d Cir. 2017).

U.S. Department of Justice. (2022, February 8). *Two arrested for alleged conspiracy to launder $4.5 billion in stolen cryptocurrency* [Press release].

U.S. Department of Justice. (2019, October 16). *South Korean national and hundreds of others charged worldwide in the takedown of the largest darknet child pornography website, which was funded by bitcoin* [Press release].

U.S. Department of the Treasury. (2022, August 8). *U.S. Treasury sanctions notorious virtual currency mixer Tornado Cash* [Press release].