

# DEFENSE EXPERT REPORT

---

*United States v. Roman Sterlingov*

21-CR00399 (RDM)

Jonelle Still

Director of Investigations and Intelligence

Ciphertrace, by Mastercard

August 7, 2023

ciphertrace

by 

Background.....	3
Critical Factors that Impact and Contradict the Government’s Conclusions.....	3
Executive Summary.....	7
A Brief History of CoinJoins and Mixers.....	11
<b>CoinShuffle - 2014</b> .....	<b>13</b>
<b>CoinShuffle++ - 2016</b> .....	<b>13</b>
<b>TumbleBit - 2016</b> .....	<b>13</b>
<b>ValueShuffle - 2017</b> .....	<b>13</b>
<b>CoinJoinXT - 2018</b> .....	<b>14</b>
<b>SNICKER - 2019</b> .....	<b>14</b>
<b>Wormhole - 2020</b> .....	<b>14</b>
<b>WabiSabi - 2020</b> .....	<b>15</b>
<b>Non-Exclusive List of Wallets that Implement CoinJoins</b> .....	<b>15</b>
CoinJoin Examples .....	16
<b>Example of a 3x4 Mt. Gox Coinjoin Transaction from 2015:</b> .....	<b>16</b>
<b>Example of a 2x4 CoinJoin Transaction from 2014:</b> .....	<b>17</b>
Payjoins .....	18
Things that Break Clustering Heuristics.....	18
<b>Why it Matters</b> .....	<b>18</b>
Evaluation of Sarah Meiklejohn’s Papers: How to Peel a Million & A Fistful of Bitcoins .....	18
Evaluation of the Application of Heuristic 2 to Chainalysis’ Model .....	20
Summary of Literature on False Discovery Rates .....	20
Difference Between a Wallet and a Cluster.....	23
Best Practices in Address Attribution & Data Integrity: Collection, Storage and Access .....	23
Evaluation of Chainalysis Expert Reports .....	24
Evaluation of Chainalysis Bitcoin Fog Attribution.....	27
<b>Summary of Chainalysis bitcoin_fog_market_addrs_cospend.csv Data</b> .....	<b>27</b>
<b>Ciphertrace Sentry API Pulls from Clusters</b> .....	<b>30</b>

Comparison of Ciphertrace and Chainalysis Dark Market Attributions .....	34
Evaluation of the Scholl Report.....	34
What is Needed to Determine Control of Bitcoin Fog .....	39
Other Domain Name Registrations of Bitcoin Fog.....	39
Notes on Publicly Available Tools.....	39
Conclusion .....	39

## Background

Roman Sterlingov's defense counsel reached out to Ciphertrace in July 2023 and requested case support regarding cryptocurrency portions in U.S. v. Sterlingov. Ciphertrace accepted the case pro bono. Ciphertrace understands that a majority if not all of Mr. Sterlingov's accounts have been frozen and he does not therefore have access to them. Ciphertrace understands that the Government has presented Chainalysis data for the on-chain cryptocurrency portion of this investigation and unless otherwise noted, I have used this data to interrogate the Government's claims. Ciphertrace understands that the Government utilized at least two proprietary (commercial) tools to trace cryptocurrency, one from Chainalysis and one from TRM, however, Ciphertrace does not have access to these tools as it is not a customer of either company.

My opinions are based upon my analysis of Chainalysis data, Expert Reports and Declarations, and the exchange production provided in the Discovery unless otherwise noted. I, Jonelle Still, reserve the right to amend my opinion as I go through the current evidence in this case and as I am presented with new evidence. Given the large volume of data and the large volume of undisclosed information by the Government and Chainalysis, this analysis is ongoing and this expert report will be supplemented as necessary.

## Critical Factors that Impact and Contradict the Government's Conclusions

For the reasons discussed below, Chainalysis' attributions are unverifiable and should not be used in a Court of law. These data have never been verified externally nor independently, have not been audited, utilize novel algorithms, are based upon experimental research, and, as expert witness Elizabeth Bisbee, from Chainalysis, testified at the Daubert Hearings, there are no known error rates, false positive rates, false negative rates, or any scientifically peer-reviewed inquiry validating the accuracy of Chainalysis' data application of its models. Therefore, I cannot verify the vast majority of Chainalysis' attribution as presented by the Government.

Three primary factors which limited this expert report are as follows: the integrity of Chainalysis' data, access to the pertinent data, and time.

Because of the lack of Chainalysis' data integrity, arising from the absence of any statistical analysis, model validation or external audits, to effectively review the scope of Chainalysis' work here necessitates the review over six million cryptocurrency addresses and over 20 million cryptocurrency transactions. This is on top of the multiple expert reports from the Government and Chainalysis, and their declarations. At essence, the lack of Chainalysis' data integrity, the lack of any scientific validity, the lack of any statistical analysis as to error rates, and the lack of model validation, all unnecessarily increase the scope of work needed to thoroughly interrogate the data to root out all of the errors. This work should have already been done by the FBI and Chainalysis but it was not. Because of this failure, it is impossible to rule out other suspects as operators

and administrators of Bitcoin Fog. A primary component of the Government investigation is its singular focus on Mr. Sterlingov while ignoring almost all other Bitcoin Fog transactions.

The Government's investigation is massively incomplete. The singularity of their focus on making attributions to Mr. Sterlingov ignores an entire universe of transactions that the Government fails to discuss in their expert reports.

For a case this size Ciphertrace will typically spend a minimum of four months utilizing a team of at least three analysts. My review of Chainalysis' analysis is limited by the fact that Chainalysis has not done the proper validation, authentication, or auditing on their data and models necessary for the Government to justify their conclusions. However, what can be analyzed reveals errors, omissions, and a lack of methodological rigor calling the Government's conclusions into serious doubt.

The Government's discovery does not contain the Bitcoin Fog ledger, nor does Ciphertrace have access to it to determine all cryptocurrency addresses, transactions, dates and amounts used in the lifespan of the privacy service. It is my understanding that the Government does not have possession of the Bitcoin Fog servers and has not yet found this ledger nor the server(s) on which it was stored, nor the private keys; therefore, the Government cannot verify its conclusions. This makes any analysis on Bitcoin Fog's inner workings and its owners speculative.

The Government provided no access to Roman Sterlingov's seized wallets, only screen shots. These screenshots require manually entering the data from the screen shots line by line, a time consuming and tedious process. Moreover, the production of screenshots is not a forensically appropriate method of presenting evidence to the Defense for analysis. I cannot confirm there are no errors in the screen shots of the government's rebuilding of Mr. Sterlingov's wallets. The naming conventions are such that it appears the investigators were unsure how to properly name each wallet. Due to the number of errors identified in the Government's experts' reports, I am concerned there are errors in the imported data. The only way to confirm this is to have access to the wallets seed phrase (private key) or rebuild it in "watch only" mode via a Master Extended Public key which would allow me to rebuild the wallet and confirm all addresses, amounts, dates/times, and transactions. This extended public key does not allow me to change or modify the wallets or transactions, nor does it allow me to spend any bitcoin. The Defense requested this from the Government, the Government stated that they were unable to provide the master extended public key. They further stated that they would provide a list of addresses, unfortunately a list of addresses is not the complete record necessary for verifying the wallets and their addresses, amounts, dates/times, and transactions.

The Government has provided no source for Chainalysis' attribution for Bitcoin Fog or all dark markets mentioned in the Government's Expert Reports. The Defense requested this information but has not received it. The Government responded by asking for the relevancy and discoverability of this data. The requested data speaks directly to the integrity of Chainalysis' data and audit trail. The Department of Justice maintains a

number of internal policies and guidance regarding information dissemination and data integrity.<sup>1</sup> The Department of Justice's own guidelines discuss the importance of data integrity:

Investigations and prosecutions based in whole or in part upon forensic science must be based upon sound science - from the crime scene to the courtroom to post-conviction reviews, and each step along the way.<sup>2</sup>

The Defense requested access to Chainalysis Reactor. The Government stated that they could not provide the Defense with access to Chainalysis Reactor. I am therefore unable to rebuild the cryptocurrency traces exactly as presented in the Expert Reports; meaning I cannot use the same proprietary tools the Government used during the course of its cryptocurrency investigation.

I also had no access to TRM attribution data and source, date, or the time of collection for Bitcoin Fog as it was identified in FBI Analyst Luke Scholl's Expert Report ("Scholl Report"). It is my understanding that the Defense has requested this data, and that the Government responded that they did not have it.

In a letter dated July 26, 2023, the Defense requested from the Government a summary of Chainalysis' and TRM's data integrity as follows:

7. We further request all of Chainalysis' Bitcoin Fog attributions; the source and date of all collections of data related to the attributions on a per address basis; the number of clustering errors and how they were resolved and the date/timestamp when those errors and resolutions occurred on a per address basis; the type and quality of each Bitcoin Fog address attribution as they relate to the heuristics described in the Bisbee report on a per address basis, specifying whether they were direct, via clustering, change addresses, or through some other heuristic. Included in this request is a breakdown of each cluster which should contain: Cluster identifier, Number of addresses in cluster, for each address in a cluster whether it identified by heuristic 1, 2, or 3 as described in the Bisbee report, Source Date/Timestamp of data collected, number of clustering errors over time with each identified cluster....

---

<sup>1</sup> See e.g. <https://www.justice.gov/information-quality>; [https://www.whitehouse.gov/wp-content/uploads/2022/01/01-22-Protecting\\_the\\_Integrity\\_of\\_Government\\_Science.pdf](https://www.whitehouse.gov/wp-content/uploads/2022/01/01-22-Protecting_the_Integrity_of_Government_Science.pdf); <https://www.justice.gov/sites/default/files/open/legacy/2013/07/29/doj-scientific-integrity-policy.pdf>.

<sup>2</sup> See U.S. Department of Justice Scientific and Research Integrity Policy, (available at <https://www.justice.gov/sites/default/files/open/legacy/2013/07/29/doj-scientific-integrity-policy.pdf>).

17. We requested information on Mr. Scholl's alleged confirmation of his tracing using TRM's blockchain tracing software. You stated that you had nothing to provide on this front as Mr. Scholl merely used the TRM software and it generated no records.

The Government told the Defense that there was nothing to provide in relation to Mr. Scholl's use of the TRM software because it generated no records. However, if Mr. Scholl utilized TRM's Application Programming Interface<sup>3</sup> ("API"), the API generates log data and TRM would have a record of those API calls.

The Government has failed to provide the necessary data to evaluate their attributions. In order to verify the Government's attributions, it is necessary to assess every attribution point. This means assessing all the records and cryptocurrency addresses referenced in the Government's expert reports, and their attribution data which was largely obtained via Chainalysis; the source and date of all collections of data related to the attributions on a per address basis; the number of clustering errors and how they were resolved and the date/timestamp when those errors and resolutions occurred on a per address basis; the type and quality of all address attributions as they relate to the heuristics described in the Government's expert reports on a per address basis which may differ depending on which heuristics were used; knowing whether they were direct attribution, via clustering, change addresses, or through some other heuristic. Neither the Government nor Chainalysis have provided this information. Nor is it apparent from the Government's expert reports, and expert testimony at the *Daubert* hearings, that they have done this review.

In order to verify the Government's attribution properly I need a breakdown of each cluster on a per address basis which should contain: cluster identifier; number of cryptocurrency addresses in each cluster; whether, for each address in a cluster, it was identified by heuristic 1, 2, or 3 as described in the Government's expert reports; the source date/timestamp of data collected; the number of clustering errors over time within each identified cluster (date/timestamp added). Basically, I need an audit trail of every single cryptocurrency address from the time it was collected to the present. This is a basic requirement necessary to validate any data.

Attributions can change over time, just like heuristics. Ms. Bisbee, in her expert report and declaration, claims that Chainalysis Reactor software is deterministic. However, cluster errors can arise when an address has more than one attribution. For instance, an address can have two owners, like in a "nested exchange," such as Chatex or Suex, where the service operates inside of the exchange. The wallets offered by nested services operate in such a way that ownership may be attributed to the exchange, the service, or the user. Furthermore, exchanges, like Coinbase, Binance, Kraken and Mt. Gox, maintain such a massive amount of liquidity that entire services operate inside of

---

<sup>3</sup> An Application Programming Interface is a set of defined rules that enable different applications to communicate with each other. API allows users to pull large amounts of data and automate processes without going through the front-end of user interfaces.

them. The concepts of ownership and control are plastic in the sense that they are subject to the vagaries of how they are defined. There is intense debate in the Cryptocurrency Anti-Money Laundering (“AML”) community about the definition of these core concepts. Mr. Scholl references this debate when he discusses the very definitions he uses in his report.

The issue is critically relevant when it comes to the Mt. Gox data. Before Mt Gox collapsed, Mt. Gox keys were hacked, stolen and sold. This makes attributions related to Mt. Gox data challenging because it’s difficult, if not impossible, to make accurate ownership attributions. Additionally, the exchange Kraken stepped in to support Mt. Gox following Mt. Gox’s collapse and is rumored to hold some of the keys. During its operation, Mt. Gox also allowed users to import their own private keys which created a massive issue in determining correct clustering and ownership. Deterministic algorithms are problematic in that they’ll always attribute those keys to Mt. Gox, even though they are no longer under Mt. Gox’s control.

In Ms. Bisbee’s report she stated that the Bitcoin Fog address types changed over time. She also stated that Chainalysis’ software is deterministic, and that Chainalysis uses Heuristic 2 (behavioral) which includes cryptocurrency address types, and the raw data associated with those addresses. This raises serious questions regarding the claim that Chainalysis Reactor is deterministic. In order for Chainalysis Reactor to be deterministic, the outputs should not change. But if that’s true, then there is a serious problem with Chainalysis’ attribution process because services and users can change the address types they use over time. Ownership can change over time as well, as demonstrated by the Mt. Gox private keys being hacked, stolen and sold. This is a glaring inconsistency in Ms. Bisbee’s expert report and is exactly why audit trails are necessary.

## Executive Summary

The Government’s claims that Roman Sterlingov operated/administered Bitcoin Fog cannot be verified using on-chain data because:

- **High Number of Errors:** Numerous factual errors and improper assumptions in the Government’s Expert Reports.
- **Data Discrepancy:** There exists discrepancy rates between Ciphertrace and Chainalysis attribution data upwards of 60%. Ciphertrace utilizes Heuristic 1 (multi-input clustering) for attributions other than direct attribution. Ciphertrace does not utilize Heuristic 2 (behavioral) because it is inaccurate, error-prone, and over inclusive.
- **Data Standards for Court:** The Chainalysis attribution data should not be used in court for this case nor any other case: it has not been audited, the model has not been validated, nor has the collection trail been identified.
  - Upwards of 64% error rate for Heuristic 2;
  - Over 527,000 Bitcoin Fog addresses clustered by Chainalysis using Heuristic 2;



- This means that the Government's reliance on Chainalysis' count of dark market interactions with Bitcoin Fog is likely misplaced and flawed.
- **Lack of Data Integrity:** The Government relies on data for which there is no independent model validation. There are no error rates cited for each cryptocurrency address or clustering heuristic, as Ms. Bisbee testified at the *Daubert* Hearings, and reiterated in her Declaration. Chainalysis did not / has not collected the data. I estimate there are hundreds of millions of data points that are unverified. Chainalysis does not appear to have verified any of their datasets via independent audits, and only makes vague claims referencing unnamed data scientists in regards to any internal validation.
  - Other cases relying upon Chainalysis data may warrant re-examination in light of these revelations;
  - Recently, a federal judge dismissed the testimony of a witness who relied upon data which had not been verified within a two-year timeframe as it violated DOJ guidelines.<sup>4</sup>
- **CoinJoins break heuristics:** Privacy services are intentionally designed to break clustering and other heuristics blockchain analytics companies rely upon for the creation of their attribution database. This is a bugbear for companies like Chainalysis whose customers rely on their data.
- **Lack of Blockchain Analytics Oversight:** Currently, there does not exist a standards body which oversees blockchain analytics companies, their models, data collection or attribution. As an emerging field, definitions, practices, and data collection are not standardized.
- **Chainalysis 'fixes' information and timeline for IRS:** The discovery produced by the Government contains a spreadsheet authored by IRS-CI Devon Beckett, last updated on August 8, 2016. In it, he appears to refer to Chainalysis manipulating the data in this case because it did not fit in to the Government's preconceived notions. The spreadsheet states:

? If BCF is truly not up an [sic] until this date, then timeline appears to fit well excluding the custom onion generation. Chainalysis seems to think there [sic] transactions before this date and should be releasing a "fix" which provide more accurate display of information/timeline

- I understand this to mean that the traces were not aligning in an advantageous way, so Chainalysis offered to adjust their algorithm or data manually to create a more favorable trace for the IRS.

---

<sup>4</sup> See e.g. Jude tosses testimony on 'unverified' data from Penguin Random House executive, COURTHOUSE NEWS SERVICE, Emily Zantow (Aug. 17, 2022) (available at: <https://www.courthousenews.com/judge-tosses-testimony-on-unverified-data-from-penguin-random-house-executive/>).

- **IRS Consensus on Withdrawal Patterns:** Mr. Sterlingov's withdrawal pattern from Bitcoin Fog is entirely consistent with user withdrawals; this is corroborated by the Search Warrant Affidavit signed by IRS-CI Special Agent Leo Rovensky<sup>5</sup> "These withdrawals occurred sporadically and in the same manner as a regular user."<sup>6</sup> The affidavit goes on to make a leap of logic stating that the likely reason Mr. Sterlingov's withdrawals match other user withdrawals is that he was trying to obfuscate his ownership. However, we know that Mr. Sterlingov deposited his funds into exchanges that required him to upload a copy of his government ID and take a selfie. The conclusion made by IRS-CI Special Agent Leo Rovensky that Mr. Sterlingov is the administrator of Bitcoin Fog is unfounded and illogical.
- **IP Address is likely a VPN or Proxy Server:** On page 18 of Mr. Rovensky's Warrant Affidavit, IP address 212.117.160.123 is identified as the address that accesses Liberty Reserve and Mt Gox accounts.<sup>7</sup> The Government uses this IP address to attribute ownership and control of the Mt. Gox accounts #2 and #3 to Mr. Sterlingov. However, this IP address appears to be a VPN, or a proxy server. That is, any number of entities or persons from anywhere in the world could be using this IP address at any one time. Ciphertrace collects IP addresses via our own node operation and links them to bitcoin addresses. The large number of bitcoin address clusters, cryptocurrency services, and exchanges that are linked or traceable to this IP address identified in the Warrant strongly point to a VPN or proxy server. Additionally, some of the bitcoin addresses linked to this IP address are also linked to multiple IP addresses. Critically, none of the bitcoin addresses linked to the 212.117.160.123 IP address are listed in the Government's findings. None of the bitcoin addresses linked to 212.117.160.123 IP address ever interacted with Bitcoin Fog. None of the bitcoin addresses linked to 212.117.160.123 IP address were ever under Mr. Sterlingov's control in his Mycelium wallet.
- **Assumption Errors:** No address attributed to Mr. Sterlingov in the Scholl Report sent funds on Oct. 27, 2011, to Bitcoin Fog via Wallet 2, as Mr. Scholl stated in his report.<sup>8</sup> Mr. Scholl stated in his report that the deposit occurred prior to the announcement of Bitcoin Fog on the Bitcoin Talk forum. However, this appears to not be the case and warrants further examination. This is but one of the many assumptions that are pervasive in the Scholl Report and this case.
- **The Tesla:** Mr. Sterlingov's purchases and behaviors do not match that of a crypto bro who made millions. In Mr. Rovensky's Warrant Affidavit, he claims that a majority of cryptocurrency in Mr. Sterlingov's accounts are back traceable to Bitcoin Fog and alleges that Mr. Sterlingov used funds coming from the privacy service to buy gift cards, goods and services, and other cryptocurrencies.<sup>9</sup>

---

<sup>5</sup> Search Warrant Application by IRS-CI Leo Rovensky, Dkt. 22-SC-2023 (Aug. 9, 2022).

<sup>6</sup> *Id.* at 20.

<sup>7</sup> *Id.* at 18.

<sup>8</sup> See TxID 6586649970D8FE8A8DB1DACF17665AE6BADE88E090FA73904EFB05F49DCC379A.

<sup>9</sup> Search Warrant Application by IRS-CI Leo Rovensky, Dkt. 22-SC-2023, p. 21 (Aug. 9, 2022).

However, the Superseding Indictment does not include a wire fraud charge. Therefore, the fiat sources were not determined to be illicit. A Tesla purchase is not unheard of, but is out of the ordinary for a person who allegedly made millions in cryptocurrency. The phrase, “wen lambo,” describes the crypto investor’s goal of purchasing a Lamborghini as a sign of wealth and status.<sup>10</sup> There are countless examples of crypto investors and traders using their gains to purchase ‘lambos,’ and these cars make appearances at cryptocurrency-centric events like Bitcoin Miami and ETH Denver.



An example of crypto-bro meme culture

- **The Government Ignored Other Leads:** Other withdrawal patterns for unknown addresses may better align with administrator payouts (size, frequency, amounts). The IRS sent subpoenas to Binance, and requested they not alert the account holders. At least 3 account holders were identified who withdrew from Bitcoin Fog. However, the Government generally ignores the universe of Bitcoin Fog withdrawals and almost exclusively focuses on Mr. Sterlingov. There are millions of Bitcoin Fog transactions.
- **There are no wire fraud charges:** In cryptocurrency cases, a wire fraud charge is typically included as the cryptocurrency (purportedly from an illicit source) is swapped for fiat currency and used to purchase luxury goods, services and in some cases, gold.<sup>11</sup>
- **Other Privacy Services:** Mr. Sterlingov’s use of other privacy services like Wasabi CoinJoin, Bitmixer, and Mt Gox CoinJoins demonstrate he was a user of privacy services, not that he controlled them.

<sup>10</sup> See When Lambo? How Lamborghini became the status brand of the crypto boom, DIGIDAY, Shareen Pathak (May 24, 2018) (available at: <https://digiday.com/marketing/lambo-lamborghini-became-status-brand-crypto-boom/>).

<sup>11</sup> See Buried gold, burning trash: US couple admits to hiding hacked crypto, REUTERS, Luc Cohen (Aug. 3, 2023).

- **Questions of Bitcoin Fog Ownership:** The privacy service Bitcoin Fog utilized multiple different deposit and withdrawal patterns in its lifetime. This suggests that the service may have had multiple owners or changed hands.
- **Not a true mixer?:** Analysis shows that it is possible Bitcoin Fog was utilizing user deposits for withdrawals. This could explain why users could withdraw funds after a month or two, and not regularly. Mr. Sterlingov typically made several withdrawals on the same day once a month or once every other month into his Mycelium wallet. This pattern matches other user withdrawals as investigated by the IRS, and complaints made on the Bitcoin Talk forum that Bitcoin Fog was a scam.
- **Advertising for Privacy is not Illegal:** Bitcoin Fog advertised itself as a privacy service. Other mixers and privacy services have sometimes advertised themselves on the Clearnet and websites accessible via TOR that they will help hide illicit proceeds. Bitcoin Fog did not.
- **Questioning the Narrative:** If Mr. Sterlingov was running a mixer and taking payouts, why was he buying bitcoin at Local Bitcoins and other exchanges?
- **Differences in Graphical Tracing Tools Leads to Errors:** Chainalysis Reactor employs single-entity clustering; that means that for each transaction on a graph, the entire entity will appear to spend funds in that interaction, even if only one address is assigned to that entity transacted. A root address is assigned to the entity which may or may not be the correct address that transacted.<sup>12</sup> This single entity clustering leads to many tracing errors.
- **Ciphertrace Inspector employs separated clusters:** This means that our clusters are truest to the activity that occurred on chain. Our graphs show the true addresses that participated in a transaction, individually. Law enforcement and other customers of Chainalysis have approached Ciphertrace on this topic and have expressed frustration related to the errors they experience using Chainalysis Reactor. The Scholl Report exhibits these types of errors.<sup>13</sup>

## A Brief History of CoinJoins and Mixers

It is often said that exchanges make the best mixers. This is due to the shared practices between the two. Exchanges accomplish what is termed ‘off-chain’ transactions, which means that cryptocurrency and values of cryptocurrency are moved around without being sent to the mempool (pool of transactions waiting to be validated and added to the next block) and validated. When a retail customer logs into their exchange account, they will see a deposit address and amount of cryptocurrency associated with that address or account. However, that number is essentially an IOU. Exchanges may use user withdrawals for other purposes such as withdrawal requests and trading. In the infamous case of exchange FTX, it is alleged that Sam Bankman-Fried used user deposits to trade at his institutional trading firm Alameda Research.

---

<sup>12</sup> See Data Credibility in Cryptocurrency Investigations, CIPHERTRACE (2021) (attached as Ex. A).

<sup>13</sup> See e.g. Scholl Report at 22-23, 46.

CoinJoins and mixers have a long history in Bitcoin. In *Section 10 Privacy* of his Satoshi Nakamoto's Bitcoin whitepaper that invents Bitcoin and the blockchain, Mr. Nakamoto states:

The traditional banking model achieves a level of privacy by limiting access to information to the parties involved and the trusted third party. The necessity to announce all transactions publicly precludes this method, but privacy can still be maintained by breaking the flow of information in another place: by keeping public keys anonymous. The public can see that someone is sending an amount to someone else, but without information linking the transaction to anyone. This is similar to the level of information released by stock exchanges, where the time and size of individual trades, the "tape", is made public, but without telling who the parties were. As an additional firewall, a new key pair should be used for each transaction to keep them from being linked to a common owner. Some linking is still unavoidable with multi-input transactions, which necessarily reveal that their inputs were owned by the same owner. The risk is that if the owner of a key is revealed, linking could reveal other transactions that belonged to the same owner.<sup>14</sup>

Early adopters of Bitcoin understood the value of privacy and adapted their on-chain transactions to preserve their privacy as outlined by Satoshi, to break the potential linking accomplished via multi-input transactions.<sup>15</sup>

The first main discussions around CoinJoin techniques were from Gregory Maxwell in 2013 iterating off of David Chaum's approach to privacy from the 1980s and 1990s.<sup>16</sup> CoinJoin operations were taking place already on the blockchain prior to Mr. Maxwell's 2013 post. Mr. Maxwell posted to the Bitcoin Dev Forum in August 2013 regarding:

[A] transaction style Bitcoin users can use to dramatically improve their privacy which I've been calling CoinJoin. It involves no changes to the Bitcoin protocol and has already seen some very limited use spanning back a couple of years now but it seems to not be widely understood.<sup>17</sup>

CoinJoin transactions started circulating more and more through forums as people noted that inputs of a CoinJoin-based Bitcoin transactions should be separately signed

---

<sup>14</sup> See Bitcoin: A Peer-to-Peer Electronic Cash System, Satoshi Nakamoto (2008) (attached as Ex. B).

<sup>15</sup> Bitcoin Q + A, BITCOIN.GUIDE, Gregory Maxwell (2013) (available at [https://bitcoiner.guide/qna/CoinJoin/#:~:text=CoinJoin%20\(sometimes%20called%20mixing\)%20is,belong%20to%20the%20same%20entity](https://bitcoiner.guide/qna/CoinJoin/#:~:text=CoinJoin%20(sometimes%20called%20mixing)%20is,belong%20to%20the%20same%20entity)).

<sup>16</sup> Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms, TECHNICAL NOTE PROGRAMMING 21 TECHNIQUES AND DATA STRUCTURES 2, David L. Chaum (Feb. 1981) (available at: <https://dl.acm.org/doi/pdf/10.1145/358549.358563>).

<sup>17</sup> See <https://bitcointalk.org/index.php?topic=279249.0>.

with associated signatures, thus the users could jointly create one transaction with their inputs. In that manner, they would break the “common ownership” heuristic (Heuristic 1), and they can hide the relation of inputs to outputs. Outside of just CoinJoin itself, there were protocols that built off of it such as CoinShuffle, CoinShuffle++, ValueShuffle, CoinJoinXT, and others.

#### CoinShuffle - 2014

CoinShuffle focused on a small communication overhead via the Dissent protocol. This protocol allows every participant to generate a fresh ephemeral encryption/decryption key pair and broadcast the public encryption key to the network. Note that this is not via the Bitcoin network. Every participant then generates a fresh Bitcoin address, designates their output address in the mixing transaction and then becomes part of the shuffling. The shuffling is the movement of these freshly generated output addresses in an oblivious manner. Each shuffling becomes another shuffle by the next participant. Each participant can individually verify that their output address is in the list of outputs. If so, the participant signs the transaction with their signing key and broadcasts the signature. Upon receiving the signatures from all participants, the transaction can be fully-signed and broadcast to the Bitcoin network.

#### CoinShuffle++ - 2016

CoinShuffle++ builds off of CoinShuffle, however it is only utilized via Decred CoinJoin transactions. The process is to obfuscate ownership of DCR (Decred) coins, where the output addresses are anonymized via some type of mixnet. The mixnet used for this is DiceMix. DiceMix claims to allow pseudonymous users to issue transactions which would be unlinkable and fully compatible with various blockchain-based systems. Through the use of an additional mixnet, CoinShuffle++ can make the outputs indistinguishable by allowing each output “mix” to have a fixed denomination. This is similar to future protocols such as CoinMixer.

#### TumbleBit - 2016

TumbleBit allows parties to make fast, anonymous, off-chain payments through an untrusted service known as a “tumbler”. This “tumbler” follows the principles set out by David Chaum’s eCash principles while allowing the payments and mixing to be done offline through TumbleBit itself.

#### ValueShuffle - 2017

ValueShuffle builds off of the predecessors in CoinJoin technology, however this time the developers focus on hiding the funds. This would be one of the first instances of trying to build out privacy enhancements to Bitcoin itself. The way of hiding transaction amounts was created much earlier than CoinJoin transactions and is known as Confidential Transactions. Confidential Transactions (CT) is a cryptographic principle



which allows one to make the value amounts in a given transaction be hidden, i.e. encrypted by one which makes it possible for the parties participating in the transaction to view the amounts. As this builds off of CoinShuffle++, the use of DiceMix will be utilized. Outside of the ValueShuffle whitepaper, there is not a lot of credible information on whether this was actually used. As it integrates CT, this would require a full consensus upgrade to Bitcoin which did not happen.

#### CoinJoinXT - 2018

CoinJoinXT was presented by Adam Gibson in 2018 which focused on both improving the privacy of Bitcoin, while also focusing on breaking transactional graph analysis. At the time of the presentation, Mr. Gibson focused on how SegWit enables pre-signing of not just individuals but chains of transactions.<sup>18</sup> Thus, he aims to create co-agreed upon contractual agreements within the Bitcoin scripts to require transferring ownership of coins into a shared controlled area - i.e. the CoinJoinXT interface itself. It then allows a refund policy which can occur through multiple steps. A lot of the work around this was considered PoC (Proof-of-Concept), however over time one can see how it correlates to JoinMarket.

#### SNICKER - 2019

Adam Gibson proposed another alternative to CoinJoins building off of his previous work in 2019 entitled "SNICKER." SNICKER stands for *Simple Non-Interactive CoinJoin with Keys for Encryption Reused*. It's a method for allowing various wallets to create CoinJoin transactions non-interactively through a multi-step process. The first step would be that User A determines a UTXO for which they know the owner's public key. User A selected those UTXO's whose value is less than the amount controlled by their wallet and creates a proposed CoinJoin between that UTXO and their own wallet's UTXO. This transaction creates three outputs: CoinJoin Output for User A, CoinJoin output of the owner of the selected UTXO (User B), and the change output to User A. From this, User A then creates a shared secret via ECDH (Elliptic Curve Diffie-Hellman) and allows User B to derive User A's public key. The full proposal is available on GitHub.<sup>19</sup>

#### Wormhole - 2020

Wormhole was proposed in 2020 by Max Hillebrand on the Bitcoin-Dev mailing list.<sup>20</sup> Wormhole was initially proposed by the Wasabi Wallet team, yet Mr. Hillebrand took the proposal and built upon it and requested feedback. The protocol sends payments as

---

<sup>18</sup> AdamISZ GitHub Repository (available at:

<https://gist.github.com/AdamISZ/a5b3fcdd8de4575dbb8e5fba8a9bd88c>).

<sup>19</sup> AdamISZ GitHub Repository, SNICKER\_BIP\_draft.mediawiki (available at:

<https://gist.github.com/AdamISZ/2c13fb5819bd469ca318156e2cf25d79>).

<sup>20</sup> Wormhole: Sending and receiving bitcoin anonymously, LINUX FOUNDATION, Max Hillebrand (Jan. 15, 2020) (available at: <https://lists.linuxfoundation.org/pipermail/bitcoin-dev/2020-January/017585.html>).

part of a Chaumian CoinJoin but also prevents the spender from learning of the received Bitcoin address. Similar to TumbleBit, it provides a trustless payment service that issues multiple rounds of communication.

#### WabiSabi - 2020

Yuval Kogman posted to the Bitcoin-Dev mailing list research into CoinJoins about a new protocol called WabiSabi.<sup>21</sup> The protocol extends the existing Wasabi Wallet protocol with an adapted technique of Confidential Transactions (CT) that was mentioned above. Through this the client can create a commitment to arbitrary outputs and amounts, without ever revealing the amounts, and is still able to prove that each amount is individually within a specified range, thus collectively summing the outputs to a specified value. The protocol is very different than the existing Wasabi Wallet implementation as of 2020 and replaces Blind Signatures with keyed-verification anonymous credentials.

#### Non-Exclusive List of Wallets that Implement CoinJoins

- Wasabi Wallet;
- Samurai Wallet;
- Sparrow Wallet;
- JoinMarket.
- And more

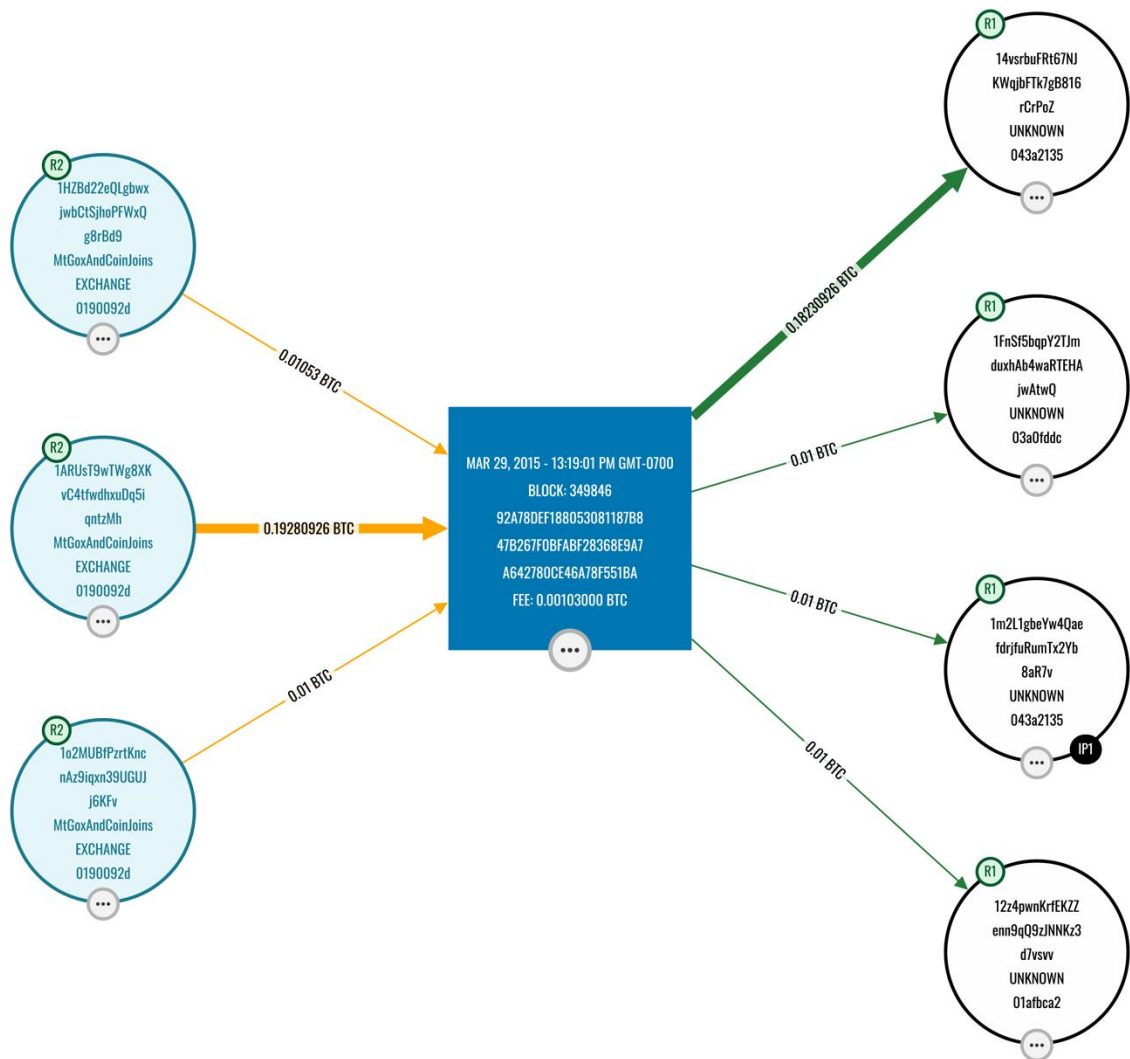
---

<sup>21</sup> WabiSabi: a building block for coordinated CoinJoins, LINUX FOUNDATION , Yuval Kogman (Jan. 11, 2020).



## CoinJoin Examples

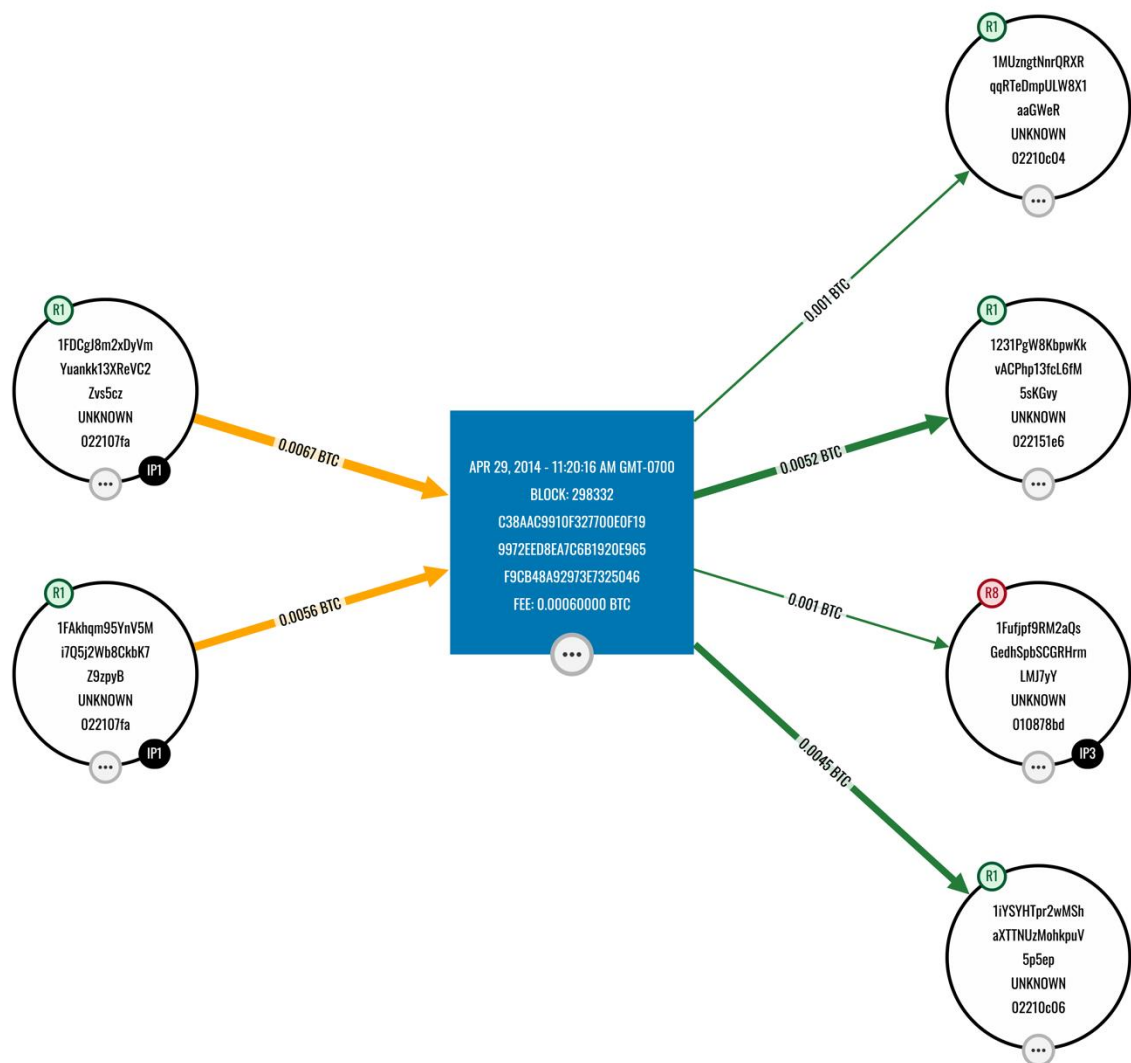
Example of a 3x4 Mt. Gox Coinjoin Transaction from 2015:<sup>22</sup>



\*Note that the attribution states “Mt Gox and Coin Joins”. Mt Gox, in addition to operating an exchange, also offered CoinJoin services.

<sup>22</sup> See e.g. 3x4 CoinJoin transaction from 2015 (available at: <https://btc.bitaps.com/92a78def188053081187b847b267f0bfabf28368e9a7a642780ce46a78f551ba>).

Example of a 2x4 CoinJoin Transaction from 2014:<sup>23</sup>



<sup>23</sup> See e.g. 2x4 CoinJoin transaction from 2014 (available at: <https://bitaps.com/c38aac9910f327700e0f199972eed8ea7c6b1920e965f9cb48a92973e7325046>).

## Payjoins

Payjoins or Pay to End Point (P2EP) are a special type of CoinJoin in which one participant pays another and the transaction is indistinguishable from a regular bitcoin transaction. The amount paid from one participant to the other cannot be determined. These Payjoins may be accomplished with or without special software.

CoinJoins provide users with a level of entropy, or privacy, for example via Chaumian blinding or Schnorr signatures.<sup>24</sup> Mixers can be custodial (Blender.io) or non-custodial (Tornado.Cash) – both services are sanctioned by OFAC.<sup>25</sup> It is important to note here that Bitcoin Fog is not, and has never been, sanctioned.

## Things that Break Clustering Heuristics

- Mixers
- CoinJoins
- Payjoins
- WabiSabi
- Layer 2 solutions (L2) – Omni Layer, Counterparty, Lightning Network
- Cross-chain atomic swaps
- Built-in privacy in wallet software such as Samurai Wallet, Electrum, Blue Wallet, JoinMarket, Sparrow, Wasabi Wallet etc.<sup>26</sup>

## Why it Matters

CoinJoins break Heuristic 1, which all blockchain tracing companies use to try to deanonymize cryptocurrency transactions on a blockchain.

## Evaluation of Sarah Meiklejohn's Papers: How to Peel a Million & A Fistful of Bitcoins

Some of the key assumptions and limitations in Sarah Meiklejohn's, et al. research papers include:

- Chainalysis provided Sarah Meiklejohn with data related to this case. In *How to Peel a Million* Sarah Meiklejohn, et al. lays out a chapter in which she attempts to validate Chainalysis' findings using a pair of proposed algorithmic tracing models. The two new models produced contradictory results.
- *How to Peel a Million* assumes the validity of the data provided to them by Chainalysis. No effort was made by Sarah Meiklejohn, et al. to independently verify the dataset. We

---

<sup>24</sup> See e.g. Schnorr Identification and Signatures, STANFORD UNIVERSITY, David Mandell Freeman (Oct. 29, 2011) (available at: <https://web.stanford.edu/class/cs259c/lectures/schnorr.pdf>).

<sup>25</sup> See U.S. Treasury Issues Forth-Ever Sanctions on a Virtual Currency Mixer, Targets DPRK Cyber Threats, U.S. DEP'T OF THE TREASURY (May 6, 2022) (available at: <https://home.treasury.gov/news/press-releases/jy0768>).

<sup>26</sup> See e.g. Samurai Wallet (available at: <https://samuraiwallet.com/features>).

cannot verify the dataset, and we know that Chainalysis has not performed any independent audits on their data collection process. Therefore, we cannot reproduce the work.

- The dataset was hand curated by Chainalysis. This suggests a concern that the dataset was inaccurate.
- Without any independent verification of the accuracy of the provided dataset, there is no way to assess the accuracy of the conclusions drawn from that dataset.
- In 2011, Bitcoin addresses all had the same features.
- All the transactions discussed in Section 7.3.1 (Bitcoin Fog) of *How to Peel a Million* had the same address features. This is why the algorithms used in the paper produced contradictory results.
- *How to Peel a Million* acknowledges that the models used are ineffective for peel-chains with the same address types.
- As *How to Peel a Million* states:

We then followed the funds from the Mt. Gox withdrawal forwards, using FOLLOWFWD, to see if we would reach the deposit to Bitcoin Fog. Both FINDNEXT and FINDNEXT2 failed after only one hop, however, as the two outputs in TX2 had the same address features and were spent in transactions with the same features. Our algorithms were thus unable to isolate the change output. These outputs were both furthermore fresh, meaning it was their first appearance in the blockchain, so the other change heuristics described in Section 7.2 also would have been unable to follow the transaction forwards.<sup>27</sup>

- In *A fistful of Bitcoins*, Ms. Meiklejohn's previous research into peel chain attribution via algorithms computed a False Discovery Rate ("FDR") of 51.64%.<sup>28</sup>
- The research demonstrates that for 80% new attribution there will be a 50% FDR. This is not a reliable model.

Heuristic	Expsn	FDR
findNext	147.43	0.62
findNext2	124.46	0.02
Androulaki et al. [2]	93.03	64.19
Meiklejohn et al. [31]	79.94	51.64
Goldfeder et al. [14]	73.7	48.7
Ermilov et al. [10]	28.6	12.7

29

<sup>27</sup> S. Meiklejohn et al., *How to Peel a Million: Validating and Expanding Bitcoin Clusters*, Sec. 7.3.1 (May 2022).

<sup>28</sup> 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*, PROCEEDINGS OF THE INTERNET MEASUREMENT CONFERENCE - IMC '13, number 6, pages 127–140, 2013.

<sup>29</sup> See *Id.*

- The FindNext2 Heuristic was applied to the hand curated dataset that Chainalysis provided for *How to peel a million*. This heuristic failed to validate the Mt. Gox trace to Bitcoin Fog that the Government attributes to Mr. Sterlingov.
- As Ms. Bisbee testified at the *Daubert* Hearings, Chainalysis has not conducted any statistical analysis of their false positive or error rates.

## Evaluation of the Application of Heuristic 2 to Chainalysis' Model

Chainalysis is vague with regards to Heuristic 2. The three heuristics identified by Ms. Bisbee in her expert report are not static. Many different assumptions can be applied under the category of Heuristic 2. Which subcategories of assumptions that are applied impact the accuracy of Heuristic 2.

Chainalysis has not revealed which sub-categories of Heuristic 2 they apply in their model in relation to Mr. Sterlingov. An external model validation, which Chainalysis has not done, could confirm the accuracy of the results. In my expert opinion the application of Heuristic 2 by Chainalysis is reckless.

Ciphertrace identified 527,731 addresses that did not cluster via Chainalysis Heuristic 1. Ciphertrace also uses Heuristic 1 multi-input clustering as the primary heuristic for non-direct attribution. Ciphertrace does not utilize Heuristic 2 as described by Chainalysis because it is often unreliable and not a true representation of the flow of funds on chain. The high prevalence of errors in Heuristic 2 are described by Ms. Meiklejohn to be between 12.7% and 64%. It is surprising that, knowing this, Chainalysis chose to apply the over inclusive Heuristic 2, despite its claims that it takes a conservative approach.

Chainalysis and Ciphertrace have a shared 397,255 addresses clustered via multi-input clustering (Heuristic 1). The 527,731 outstanding addresses were clustered under Chainalysis' undefined Heuristic 2 model. We have calculated a discrepancy rate of roughly 64% which is due to Chainalysis' use of the over inclusive Heuristic 2. The over-inclusivity of Heuristic 2 leads to dramatically high rates of false positives and implicates innocent cryptocurrency users.

In this case, the over inclusivity of Chainalysis' flawed heuristics leads to false positives. Chainalysis' over inclusive clustering methods leave it to be determined whether they have attributed addresses to Bitcoin Fog that never had anything to do with Bitcoin Fog. Without significant time to go through each of the 527,731 addresses by hand, we cannot account for the extremely large discrepancy except to point to the research and note that the difference lies in the application and error rates of Heuristic 2.

## Summary of Literature on False Discovery Rates

The Government's Supplemental Notice of Intent to Present Expert Testimony states:

Additionally, attached as Exhibit 2 is a research paper authored by several notable blockchain academics, including Sarah

Meiklejohn. See Ex. 2 (George Kappos et al., *How to Peel a Million: Validating and Expanding Bitcoin Clusters* (2022), at 2, <https://arxiv.org/abs/2205.13882>). The paper was previously cited in the government's Opposition to Defendant's Omnibus Motions in Limine. ECF No. 73, note 4. The paper's focus is on the proposal of a new peel chain heuristic, but, **relevant to the Court's inquiry, the researchers used information provided by Chainalysis as "ground truth" data, indicating a high confidence among the academic community in the reliability of the information.** Ex. 2 at 1-2. The paper's "Related Work" section includes discussion of and citation to further academic research in blockchain analysis. Id. at 2. (emphasis added).

Section 4 of *How to peel a million* states that the 'ground truth data' consisted of 60 hand curated clusters:

To start, we were given 241 Bitcoin addresses and 20,016 Bitcoin transactions by Chainalysis, a company that provides blockchain data and analysis to businesses and government agencies. The addresses represented true positive clusters, in the sense that Chainalysis had manually verified that all the addresses in the same co-spend cluster as this address really did belong to the same service (typically by confirming directly with the service). The transactions were all CoinJoins and thus represented false positive clusters, meaning all of the addresses in the resulting co-spend cluster would not actually belong to the same service. Each address formed a distinct cluster, and there was no overlap between the addresses in the true positive (TP) clusters and the ones used as inputs in the false positive (FP) transactions. This ground-truth dataset was necessary for evaluating our heuristics, and would not have been possible to get at this scale without working with Chainalysis or directly with the services themselves. None of the clusters represented individual users, and we had no additional information about the entities represented by the clusters (e.g., the name of the service).<sup>30</sup>

These hand curated data points were selected specifically to eliminate the possibility of false positives. The fact that hand curated data was used rather than raw data speaks to the fact that raw data is not reliable enough to validate heuristics.

---

<sup>30</sup> S. Meiklejohn et al., *How to Peel a Million: Validating and Expanding Bitcoin Clusters*, Sec. 4, p. 3 (May 2022).

Section 7.3.1 of *How to peel a million* attempts to find a link between a withdrawal from Mt Gox and a deposit to BitcoinFog. Both forward tracing heuristics, FINDNEXT and FINDNEXT2, failed after the first transaction hop. They were able to produce a trace using a backward tracing heuristic, but they say this heuristic produces more errors.

The academic cited in *How to peel a million* lists a wide-range of error rates for various heuristics ranging from 12.7% to 64.19%. Notably, the heuristics with the highest claimed accuracy rate, FINDNEXT and FINDNEXT2, were the heuristics that failed to find a link between the Mt. Gox transactions and Bitcoin Fog.

- **Androulaki et al. [2]** identify the change output in a transaction tx if (1) the transaction has exactly two outputs, and (2) it has the only fresh address in tx.outputs, meaning output.addr is the only one appearing for the first time in the blockchain.
  - o This method produced a **64.19%** False Discovery Rate using the hand curated set
  - o [2] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun. Evaluating user privacy in Bitcoin. In International Conference on Financial Cryptography and Data Security, volume 7859 LNCS, pages 34–51, 2013.
- **Meiklejohn et al. [31]** identify the change output in a transaction tx if (1) it has the only fresh address in tx.outputs; (2) tx is not a coin generation; and (3) there is no selfchange address in tx.outputs, meaning no address used as both an input and an output.
  - o This method produced a **51.64%** FDR
  - o [31] 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 Internet Measurement Conference - IMC '13, number 6, pages 127–140, 2013.
- **Goldfeder et al. [14]** use the same conditions as the one by Meiklejohn et al. but additionally require that (4) the transaction tx is not a CoinJoin.
  - o This method produced a **48.7%** FDR
  - o [14] S. Goldfeder, H. Kalodner, D. Reisman, and A. Narayanan. When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies. arXiv preprint arXiv:1708.04748, 2017.
- **Ermilov et al. [10]** were the first to consider not only the behavior of the outputs and their addresses but also the value they received. They identify the change output in a transaction tx if (1) the transaction has exactly two outputs; (2) the transaction does not have two inputs; (3) there is no self-change address; (4) the output has the only fresh address in tx.outputs; and (5) the output's value is significant to at least the fourth decimal place
  - o This method produced a **12.7%** FDR
  - o [10] D. Ermilov, M. Panov, and Y. Yanovich. Automatic Bitcoin address clustering. In Proceedings of the 16th IEEE International Conference on Machine Learning and Applications (ICMLA 2017), pages 461–466, 2018.



The two new heuristics cited in *How to peel a million*, that failed to find a link between the Mt. Gox and Bitcoin Fog have alleged FDR rates as follows:

- **FINDNEXT - 0.62% FDR**
- **FINDNEXT2 - 0.02% FDR**

This new heuristic is purportedly accurate because it is more discriminating.

## Difference Between a Wallet and a Cluster

A wallet “contains” all the addresses derived from a private key. This includes addresses on multiple chains if they are from the same private key. It is specific to only one private key. It is only possible to know the addresses in the wallet if the key to the wallet is possessed. A cluster is all the addresses which are related to each other via co-spending. A cluster can include addresses from multiple wallets if the wallets have been used to construct a “CoinJoin” transaction. CoinJoin is used here to denote any transaction involving multiple private keys (wallets) regardless of who controls them. Wallet addresses that have not co-spent are not in the same cluster.

This effects Heuristic 1 (multi-input clustering) because: in an explorer there may be multiple clusters associated with a wallet, especially in wallets that are managed to minimize co-spending; in a public block explorer there may be multiple wallets in a cluster. The multi-input clustering heuristic assumes that the keys to the inputs of a transaction are controlled by the same party. This allows for multiple keys being used. This also assumes the logistical difficulty, but not impossibility, of multiple parties coordinating to sign a given transaction.

This impacts Heuristic 2 because: the major issue with Meiklejohn’s research is this heuristic can have a False Positive Rate between 12.7% and 64.19% depending on the methodology used. This was measured by running the various heuristics on a hand curated data set; these errors would then be compounded by successive runs of Heuristic 1 and 2 on the new clusters interactions.

## Best Practices in Address Attribution & Data Integrity: Collection, Storage and Access

The Best Practices in Cryptocurrency Address Attribution are meant to serve as a guide and are non-exhaustive. They generally follow the data integrity definition and roadmap outlined by Harvard Business School (HBS). HBS defines data integrity as “the accuracy,



completeness, and quality of data as it's maintained over time and across formats. Preserving the integrity of your company's data is a constant process.”<sup>31</sup>

## Evaluation of Chainalysis Expert Reports

There are number of errors, omissions, and inconsistencies in Ms. Bisbee's Chainalysis Report. A non-exhaustive list of the errors follows:

### Page 8, Table 2<sup>32</sup>

- This table is missing a number of Bitcoin script types such as: P2PK (different than P2PKH), P2MS, P2WSH, or P2TR.
- The table states that P2SH is “*a SegWit address that begins with 3*”. That is false. Once SegWit became active on August 23rd, 2017, it could utilize SegWit to create a nested P2WSH address that is nested in a BIP-16 enabled P2SH address format. However, prior to August 23rd, 2017, all Bitcoin addresses that have the prefix of “3” are pure BIP-16 P2SH addresses, not SegWit.
- The table uses the wrong unit of data when referring to *compressed* and *uncompressed* keys. The table uses bits instead of bytes. This is a critical error. A *bit* is the smallest unit of data that a computer can process and store. Public Key Cryptography, or asymmetric cryptography, utilize the use of private and public key pairs. For Bitcoin, we use ECDSA (Elliptic Curve Digital Signature Algorithm) - with this we generate a 256-*bit* number for a Private Key. 256 *bits* is equivalent to 32 *bytes*. Thus, through ECDSA, an uncompressed public key is 65 *bytes* and a compressed public key is 33 *bytes*. Because of this, the mathematics in footnotes 6 and 7 are incorrect.

### Page 12, Section 1.2

- It does not make sense that Bitcoin Fog used uncompressed keys until 2012 when the report then states “*the P2PKH Compressed addresses changed to using SegWit addresses*”. How can the same addresses be switched from uncompressed to compressed and still output the same human-readable address hash?
- The report states that “*The first known Bitcoin Fog transaction where the change is a P2SH-WPKH Segwit address in block 534129: 9a7e1cdb9f68573eaf64ba4f8908ebf05aee932124 6188c1c746a297e8821ffb*”. Reviewing this transaction, reveals there is no witness data and no addresses participating in this transaction that are SegWit enabled. The following two open source explorers confirm this:

<sup>31</sup> What is Data Integrity and Why Does it Matter, HARVARD BUSINESS SCHOOL, Catherine Cote at 1 (Feb. 4, 2021) (available at: <https://online.hbs.edu/blog/post/what-is-data-integrity>).

<sup>32</sup> The Nov. and Dec. Expert Reports by Ms. Bisbee are substantially the same.

Inputs & Outputs

Details

1BywFVHrLRMVRE2weKVxfUhm4Gi7oz3AdB

0.15975226 BTC

ScriptSig (ASM)

OP\_PUSHBYTES\_71 30440220344f0edf1025cc85d640be1354da3cb0f076938751bf7544d043e198565a2b5f0220782172a5a2a8b0dda52732d767b65008c005b050c4b34f597494b97d3c48a4b01  
OP\_PUSHBYTES\_33 029a7453030a41d5cea8e9a617cf38919aaf2259d7c39664eaea124c1265a5c464

ScriptSig (HEX)

4730440220344f0edf1025cc85d640be1354da3cb0f076938751bf7544d043e198565a2b5f0220782172a5a2a8b0dda52732d767b65008c005b050c4b34f597494b97d3c48a4b0121029a7453030a41d5cea8e9a617cf38919aaf2259d7c39664eaea124c1265a5c464

nSequence

0xffffffff

Previous output script

OP\_DUP  
OP\_HASH160  
OP\_PUSHBYTES\_20 7860f7db737e3a5d693bf56f55dff4a040546932  
OP\_EQUALVERIFY  
OP\_CHECKSIG

Previous output type

P2PKH

35B5bRo2Q4Ro9KB5WNMGzkRFWEuKPLhcKF

0.00489000 BTC

ScriptPubKey (ASM)

OP\_HASH160  
OP\_PUSHBYTES\_20 26371694086d40e1f52f0f515169add26b124cd8  
OP\_EQUAL

ScriptPubKey (HEX)

a91426371694086d40e1f52f0f515169add26b124cd887

Type

P2SH

3DJaGcMpQDRkSVFrdHeGnYKpVpeE19mw

0.15484897 BTC

ScriptPubKey (ASM)

OP\_HASH160  
OP\_PUSHBYTES\_20 7f62f17f27e621dfbdc6612da10a43a23b0d67b  
OP\_EQUAL

ScriptPubKey (HEX)

a9147f62f17f27e621dfbdc6612da10a43a23b0d67b87

Type

P2SH

0.15973897 BTC

mempool.space

Bitcoin transaction

9a7e1cdb9f68573eaf64ba4f8908ebf05aee9321246188c1c746a297e8821ffb

Time

2018-07-28 20:14:13 UTC

4 years 11 months ago

Block

534129 [717]

Type

legacy

Confirmations

266 / 135

Confirmation time

less than minute

Size / base size

221 / 221 bytes

Virtual size / weight

221 / 884

Hash

9a7e1cdb9f68573eaf64ba4f8908ebf05aee9321246188c1c746a297e8821ffb

Version

2

Lock time

534 128

Fee

0.00001329 BTC

Fee rate

6.00 satoshi/vByte

Raw transaction

Schema

1 input

0.15975226 BTC

2 outputs [spent]

0.15973897 BTC

1BywFVHrLRMVRE2weKVxfUhm4Gi7oz3AdB

0.15975226

35B5bRo2Q4Ro9KB5WNMGzkRFWEuKPLhcKF

0.00489000

3DJaGcMpQDRkSVFrdHeGnYKpVpeE19mw

0.15484897

Miner fee:

0.00001329 BTC

Amount:

0.15973897 BTC

Input scripts

0 P2PKH 1BywFVHrLRMVRE2weKVxfUhm4Gi7oz3AdB

SigScript OP\_PUSHBYTES[71] 30440220344f0edf1025cc85d640be1354da3cb0f076938751bf7544d043e198565a2b5f0220782172a5a2a8b0dda52732d767b65008c005b050c4b34f597494b97d3c48a4b01 OP\_PUSHBYTES[33] 029a7453030a41d5cea8e9a617cf38919aaf2259d7c39664eaea124c1265a5c464

Output scripts

0 P2SH OP\_HASH160 OP\_PUSHBYTES[20] 26371694086d40e1f52f0f515169add26b124cd8 OP\_EQUAL

1 P2SH OP\_HASH160 OP\_PUSHBYTES[20] 7f62f17f27e621dfbdc6612da10a43a23b0d67b OP\_EQUAL

bitaps

Page 13, Section 1.2

- What is *RF*? If *RF* means *RBF*, (Replace-by-Fee), then it is important to note that a majority of transactions in the report would never have been able to utilize RBF as it was not present in Bitcoin Core until Bitcoin Core v0.12 which was released November 1st, 2016.<sup>33</sup>

Page 15, Section 2.1

- There are no screenshots to support the identified deposit addresses from pwoah7foa6au2pul[.]onion.

Page 16, Section 2.2

- The report states that the address identified as a co-spend participant, 1NGpmfXeFmKB4csqeUhSqCNXBJtCWua8fr, had “*activity from August 2015 and September 2015*”. When looking at the address, it was not active during this time-period. It only had two transactions, one on April 29th, 2015, and the other on May 4th, 2015.

Page 17, Section 3.1

- There are no screenshots to support the identified deposit addresses from k5zq47j6wd3wdvjq[.]onion.

Page 19, Section 4.1

- There are no screenshots to support the identified deposit addresses from agorahooawayyfoe[.]onion.

Page 20, Section 5.1

- There are no screenshots to support the identified deposit addresses from dkf2lnsctjvoivow[.]onion and o7v3h5ts5tah4yiw[.]onion.

Page 21, Section 6.1

- There are no screenshots to support the identified deposit addresses from abraxasdegupusel[.]onion.

Page 23, Section 7.1

- There are no screenshots to support the identified deposit addresses from pandorajodqp5zrr[.]onion.

Page 24, Section 8.1

- There are no screenshots to support the identified deposit addresses from sheep5u64fi457aw[.]onion.

---

<sup>33</sup> Bitcoin Core v0.12.0, sipa, (Nov. 1, 2016) (available at: <https://github.com/bitcoin/bitcoin/releases/tag/v0.12.0>).

## Page 26, Section 9.1

- There are no screenshots to support the identified deposit addresses from wztyb7vlfcw6l4xd[.]onion.

## Evaluation of Chainalysis Bitcoin Fog Attribution

This evaluation uses Chainalysis provided data in the CSV file produced by the Government titled: "bitcoin\_fog\_market\_addrs\_cospend".

### Summary of Chainalysis bitcoin\_fog\_market\_addrs\_cospend.csv Data

Ciphertrace identified 527,731 addresses that did not cluster via Chainalysis Heuristic 1. Ciphertrace also uses Heuristic 1 Multi-input Clustering as the primary heuristic for non-direct attribution. Ciphertrace does not utilize Heuristic 2 as described by Chainalysis, as it is often unreliable and not a true representation of the flow of funds on chain. The high prevalence of errors in Heuristic 2 are described to be between 12.7% and 64%. Therefore, the discrepancy rate between Ciphertrace and Chainalysis Bitcoin Fog attribution is roughly 67%.

Without significant time to go through each of the 527,731 addresses by hand, we cannot account for the extremely large discrepancy except to point to the research documenting error rates and note the difference lies in the use of Heuristic 2.

The data provided in the csv is formatted as follows:

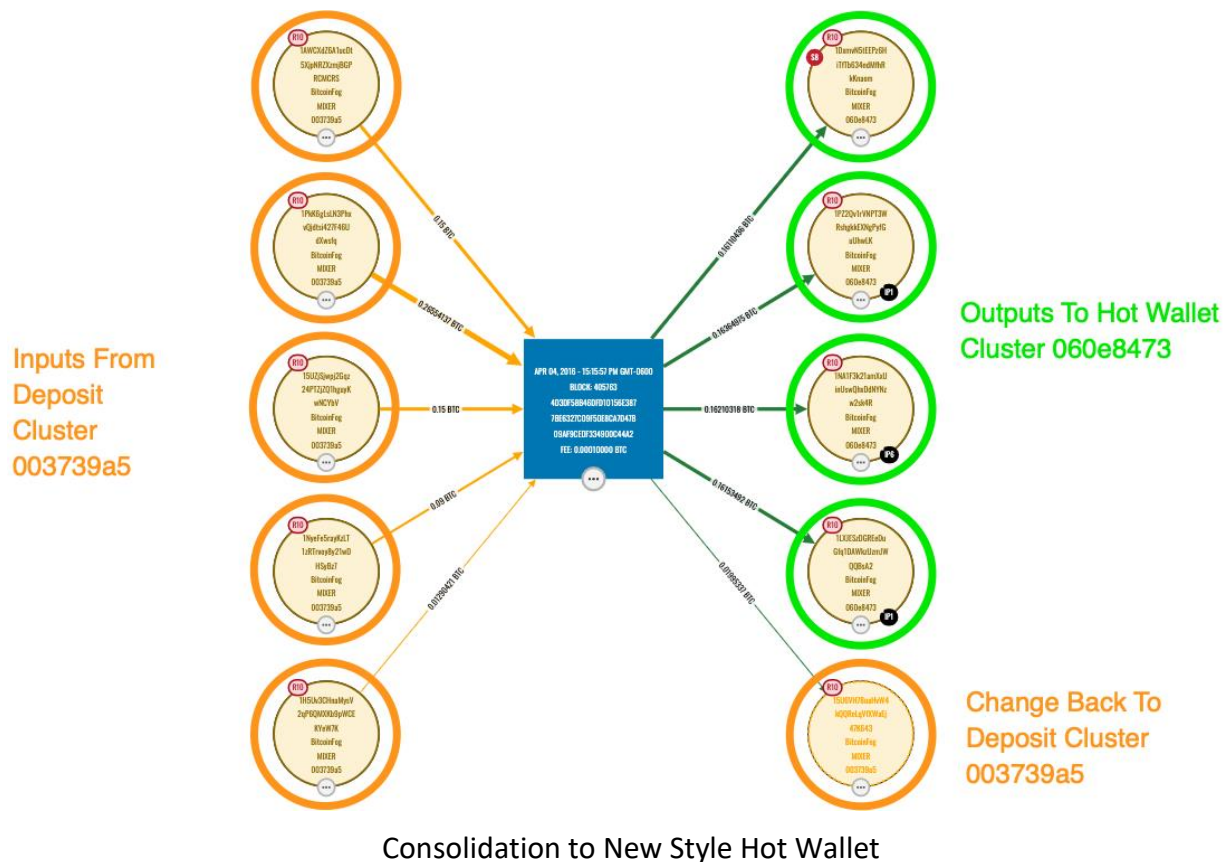
address	asset	aid	entity	bid	bid_last_change	name	category	co_spend_flag	co_spend_root_address
111233LRj5a	0	74780392	2479113	355136	355136	Bitcoin Fog	mixing	1	UNCLUSTERED IN CO-SPEND
11126bGUS	0	141613993	2479113	407872	407872	Bitcoin Fog	mixing	1	UNCLUSTERED IN CO-SPEND
11128U2TcD	0	35885413	2479113	300708	300708	Bitcoin Fog	mixing	0	UNCLUSTERED IN CO-SPEND
11129bx2Xkl	0	61774570	2479113	340098	340098	Bitcoin Fog	mixing	1	17aBK3VYVYvWwSEc4rGaYJ6b3qriB845a

There is one row for each address attributed to BitcoinFog. The addresses are categorized according to their co\_spend\_root\_address, the address they are clustered with. The following Summary of co\_spend\_root\_address comparing Chainalysis' attribution data with Ciphertrace's.

co_spend_root_address	Chainalysis Address Count	Ciphertrace WalletId	Owner	Type	Ciphertrace Cluster Size
UNCLUSTERED IN CO-SPEND	527731				
17aBK3VYVYvWwSEc4rGaYJ6b3qriB845a	244975	003739a5	BitcoinFog	mixer	244975
15CLUub6yaov3yMZtmxPQ4pSeR22PuDSL T	46532	0323355a	BitcoinFog	mixer	46532
1Mzz2hhbrCr26HX6wDgSBRheu8qphDoR9 b	31994	046d1ba1	BitcoinFog	mixer	31994

1EXTHjVRMao2AQsMTRJE5aTAGoYMmxc3 Ji	22043	0538451e	BitcoinFog	mixer	22043
1DxmvN5tEEPz6HiTfTb634ndMfhRkKnaom	50569	060e8473	BitcoinFog	mixer	50569
16FgQXGzSLRtdwuWcN7mcaPUtbJ6JFTkV w	1140	01b92444	BitcoinFog	mixer	1140
17gH1u6VJwhVD9cWR59jfeinLMzag2GZ43	726	0274e40d	Unknown	Unknown	726
1P5pMkN1wr3ozHXwCXtRmv6kJBYLyPPMz c	4	01918b80	Unknown	Unknown	4
1Cwb33nqn4S2uDsXwhNrUNy7FPdiRYhyM 8	16	016a834c	Unknown	Unknown	16
15U5NjgAbKqKyGkWayS648WwJoiCCvGnT G	2	1257968	Unknown	Unknown	2
1JmQN8NvX3XXWwRjW3rEEcKMQd5DU gkH3	6	0099a1fc	Unknown	Unknown	6
1F9kYDpu2Cqwr18ovineZr8Y88NQfW1bzR	2	0042232e	Unknown	Unknown	2
1YZJKaAx2HRWvcbCXDBtQbBZcRU46WJq w	2	351640	BitcoinFog	mixer	2
Total					398011

There are 398,011 addresses which are clustered together, and another 527,731 addresses attributed by other heuristics. Cluster 003739a5 is the customer deposit cluster. This is corroborated by the Chainalysis, FBI, and IRS deposits. Cluster 060e8473 is the withdrawal hot wallet cluster corroborated by the Chainalysis, FBI, and IRS withdrawals.



## CIPHERTRACE Sentry API Pulls from Clusters

Cluster 0323355a

Example

TX

69AD7BAEE97CD809D71A1EA4D72758A0CF5C2D3C56B22D5B078CF1AE888E3F01, Feb 7, 2015.

Inputs

```
"pos": 0,
  "address": "157zj77TD8CT62wPTp7YBEVWtMXNB3iDg6",
  "value": 0.172

"pos": 1,
  "address": "1Pet1bkEtXJSHZXpcEhiJn4Up8cZt1kydM",
  "value": 0.20725958

"pos": 2,
  "address": "14zwCx9nXPkiwPUqnEuKnRynP4D2gVpwwD",
  "value": 0.07785078

"pos": 3,
  "address": "1AyRHdPouNcrEGGECuWKWExdRGP6u74p2q",
  "value": 0.8999

"pos": 4,
  "address": "19C4vVRq9r1igEZjfGvcYkXlohoFKKZhDf",
  "value": 0.3496

"pos": 5,
  "address": "1FWHVxL73sX9SmbJpU9U7Vhfb9Fft2RHJ3",
  "value": 0.07575247
```

Outputs

```
"pos": 0,
  "address": "17ehuXt67dMDvQFV7Mfg9735B3yisXSHzQ",
  "value": 0.18741046
```

```
"pos": 1,
"address": "1A2VpWmdGW2aR8bFVkpW3v82Pe63bys7G7",
"value": 0.03314727

"pos": 2,
"address": "1DnhH18t8A1Km5z6nGLjQLVrzPGS2kAcG",
"value": 0.21808353

"pos": 3,
"address": "1PtoUTyaVZvqXT2wjSuj73e81VNsnZk2p2",
"value": 0.20053006

"pos": 4,
"address": "1LNBfzZgAbsHqbnQrMznRjiwe53VUFh3K4",
"value": 0.17846811

"pos": 5,
"address": "17ahuMui9cTRNHpyZFTgtUpzr2TJ3sjtQr",
"value": 0.18208911

"pos": 6,
"address": "1F6kGH6BRGe9iSNcUJp6egEU9DYFzYcQxi",
"value": 0.16379711

"pos": 7,
"address": "18bULF4dmniBdS4SUkkKB14qooLdoBCBdj",
"value": 0.21211967

"pos": 8,
"address": "1HDo364e5PnMbC4Wb6XADkGKrgSTheV4dG",
"value": 0.20836623

"pos": 9,
```



```
"address": "15CLUub6yaov3yMZtmxPQ4pSeR22PuDSLt",  
"value": 0.19815128
```

All the inputs and output position (pos) 1 are from cluster 003739a5, the deposit cluster. The remaining outputs belong to Cluster 0323355a. A graphical trace is below:



## Comparison of Ciphertrace and Chainalysis Dark Market Attributions

The following lists the discrepancy rates between Ciphertrace and Chainalysis's dark market attributions.

Abraxas - 20%  
Agora – 3.5%  
AlphaBay – 96%  
Bitcoin Fog – 67%  
BlackBank – 16%  
Nucleus – 44%  
Pandora – 21%  
Sheep – 0%  
SilkRoad – 43%  
SilkRoad2.0 – 0%  
WelcomeToVideo – 1%

Ciphertrace attributes these discrepancies to Heuristic 2, and other unnamed heuristics utilized by Chainalysis, which were not explicitly stated in their expert reports.

Chainalysis has not produced the sources, dates and times of collections, clustering errors and error rates on a per address basis but have not received this information.

## Evaluation of the Scholl Report

There are numerous errors, omissions, and inaccuracies in the Scholl Report.

### Page 10

Dates are incorrect. The 2014 date should be 2019.

### Page 34

When stating *“Blockchain analysis of the transactions included in the file indicated that the MYCELIUM WALLET received a deposit of approximately 29 BTC valued at \$280,544 on 6/17/2020. This deposit came directly from the BITCOIN FOG CLUSTER and was the source of funds of all 69 subsequent withdrawals from the MYCELIUM WALLET.”*.

This is a conclusory statement with no verifying data. There is no list of transactions that total 29 BTC. In order to verify this, the master extended public key or the seed phrase are required. Neither of which the Government has provided to the Defense.

### Page 45

There is no data presented justifying the conclusion that address 1LZvkK1QMCCPUoRRsJb7mxqX2tcLUqGuYX is Bitcoin Fog. The conclusory attribution lacks support.

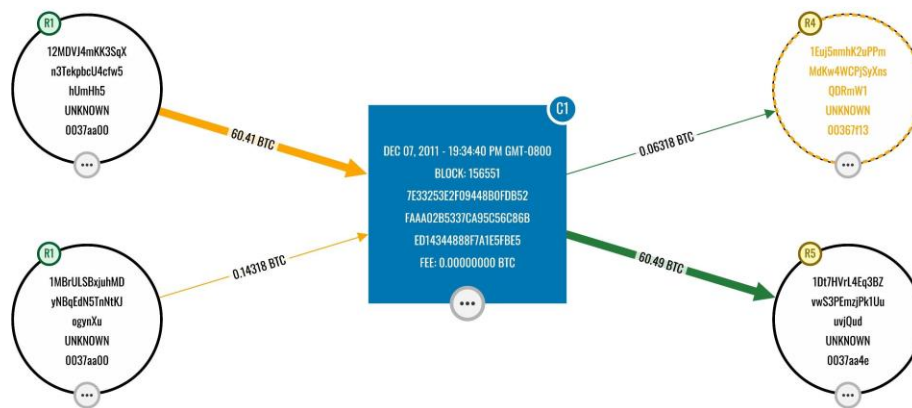
Page 46

There are multiple errors on this page.

The address 12MDVJ4mKK3SqXn3TekpbcU4cfw5hUmHh5 never sends funds to the address identified on this page. Also, the date is incorrect. Furthermore, the user ID is incorrect according to internal Silk Road records provided in the Government's discovery. In the previous transaction, from Mr. Sterlingov's Mt. Gox account to address 1Pfkqm3YsCYnWeA7h14Zmm1j8kiFruFvqA, there is no co-spend, as identified in the Scholl Report. The 1Pfk has no additional input with the 7.9 BTC prior balance as mentioned in the Scholl report. This address only receives funds one time, from Mt. Gox and sends funds one time. The owner of this address is unknown.

Silk Road deposit address 12MDVJ4mKK3SqXn3TekpbcU4cfw5hUmHh5 received funds one time, in a co-spend, whose source is the exchange VirWox.com. No such exchange has been identified or discussed in any of the Government's reports.

Below is tracing demonstrating the complete history for cryptocurrency address 12MDVJ4mKK3SqXn3TekpbcU4cfw5hUmHh5. This contradicts Mr. Scholl's report.



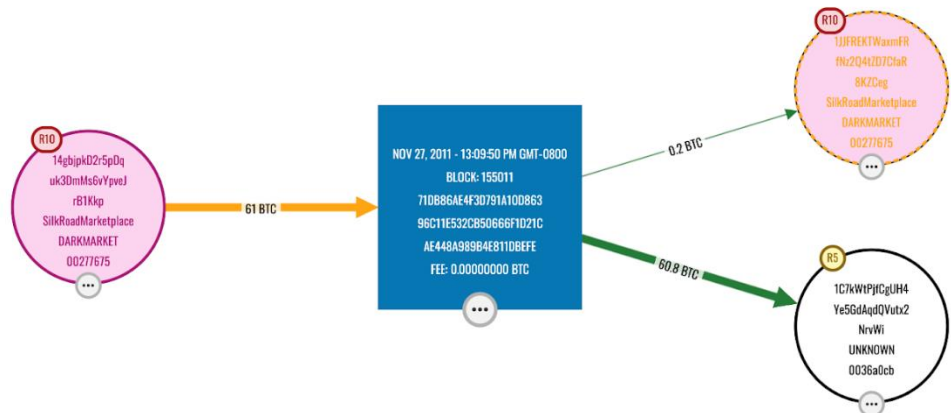
Ciphertrace Inspector

Inputs & Outputs				Details
12MDVJ4mKK3SqXn3TekpbcU4cfw5hUmHh5	60.41000000 BTC	1EuJ5nmhK2uPPmMdKw4WCPjSyXnsQDRmW1	0.06318000 BTC	
1MBrULSBxjuhMDyNBqEdN5TnNtKJogynXu	0.14318000 BTC	1Dt7HVR14Eq3BZvwS3PEmzjPk1UuuVjQud	60.49000000 BTC	
			60.55318000 BTC	

mempool.space

The Scholl Report claims that on 10/27/2011 60.8 BTC transaction occurred. No such transaction occurred on this date. There is a transaction on 11/27/2011 that sends 60.8 BTC to 1C7kWtPjfCgUH4Ye5GdAqdQVutx2NrvWi but it is not from 12MDVJ4mKK3SqXn3TekpbcU4cfw5hUmHh5. The 60.8 BTC comes from 14gbjpkD2r5pDquk3DmMs6vYpveJrB1Kkp. 14gbjpkD2r5pDquk3DmMs6vYpveJrB1Kkp according to the Scholl Report is part of Silk Road. However it is not the actual “spender” of funds nor is there a UTXO that displays what Mr. Scholl’s graph shows. Thus, it is incorrect statement of the flow of funds from a blockchain forensics perspective.

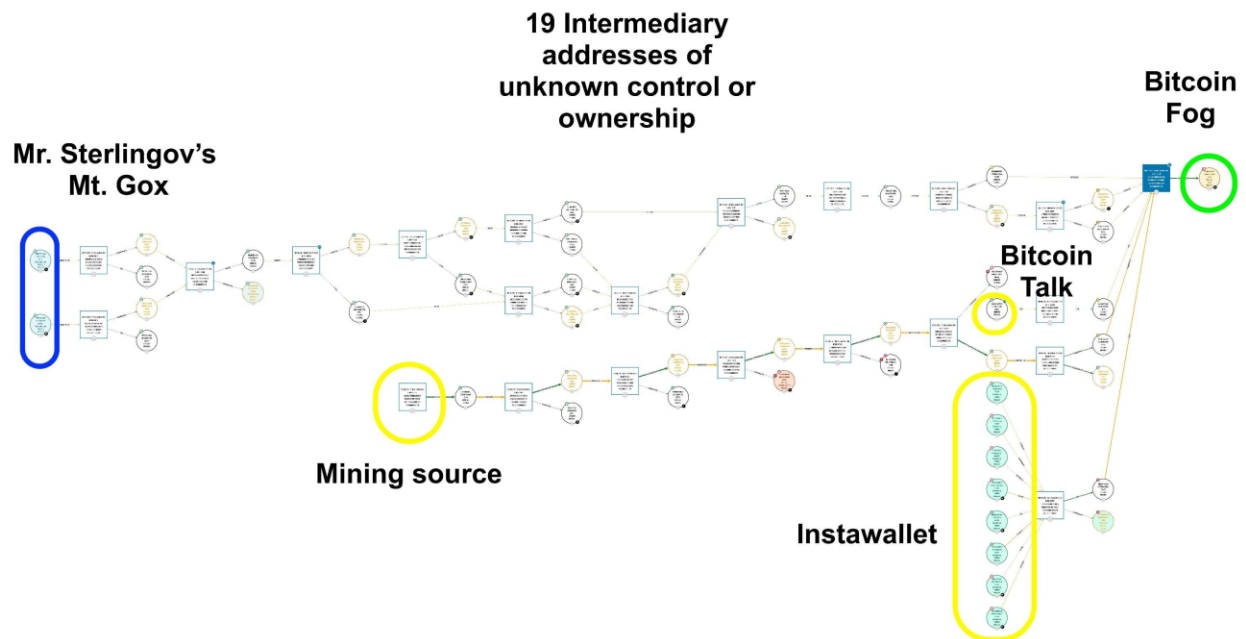
Trace: 71db86ae4f3d791a10d86396c11e532cb50666f1d21cae448a989b4e811dbefe 8 BTC (Addresses 3) SEARCH GRAPH UNDO OPTIONS



14gbjpkD2r5pDquk3DmMs6vYpveJrB1Kkp	61.00000000 BTC	1JFFREKTWaxmFRfNz2Q4t2D7CfaR8K2Ceg	0.20000000 BTC
		1C7kWtPjfCgUH4Ye5GdAqdQVutx2NrvWi	60.80000000 BTC
			61.00000000 BTC

Page 49

Mr. Scholl incorrectly concludes that Mr. Sterlingov deposited funds into Bitcoin Fog on Bitcoin Fog's first day of public operation. Between Mr. Sterlingov's Mt. Gox account withdrawal and his alleged deposit into Bitcoin Fog, there are 19 unattributed, unknown cryptocurrency addresses. The image below documents this fact:



Crucially, none of these 19 intermediary addresses are in Mr. Sterlingov's Mycelium wallet, or any other wallet identified by the Government as being under his control. The conclusion that Mr. Sterlingov deposited funds into Bitcoin Fog on its first day of operation is illogical, speculative, and incorrect.

What appears to have happened is because there are deposits in the same timeframe whose ownership it is impossible to determine, the Government seems to have latched on to Mr. Sterlingov because his account was the only one that could be traced back to a KYC account – his Mt. Gox account. The other deposits to Bitcoin Fog at this time come from an Instawallet/Paymium address (which kept no records), a mining transaction, and BitcoinTalk.

There is no evidence that Mr. Sterlingov controlled any of the 19 intermediary addresses between Bitcoin Fog and his Mt. Gox account. Any attribution of Mr. Sterlingov controlling these 19 intermediary addresses is pure speculation for which there is no corroborating evidence.

#### Page 54

Mr. Scholl makes the following observation on the bottom of page 54 of his expert report:

Tx10 appeared to spend funds from multiple sources and consolidate them at 1YZJKa. According to the Bitcoin blockchain, 1YZJKa address was involved in 66 total Bitcoin transactions, including 33 deposits and 33 withdrawals. These 33 deposits totaled approximately 30,458 BTC and occurred from on or about 11/10/2011 to on or about 3/25/2012. Multiple individual deposit transactions made into 1YZJKa included over 50 input addresses. Based on my training and experience, 1YZJKa was not consistent with a typical user deposit address at a service. 1YZJKa appeared to be an internal address at BITCOIN FOG used to consolidate multiple deposits made by multiple users.

#### Page 55

According to Chainalysis Reactor, 31 of the 33 deposits to 1YZJKa (including the two examples above) came from Bitcoin addresses within the BITCOIN FOG CLUSTER. Only the first two deposits, including Tx10 and one other deposit<sup>12</sup>, were from addresses not attributed by Chainalysis to the BITCOIN FOG CLUSTER.

Bitcoin address 1YZJKa was very likely an internal consolidation address at BITCOIN FOG and not a user deposit address. Therefore, Tx10 was very likely an internal transaction at BITCOIN FOG and addresses in WALLET 2 were part of BITCOIN FOG.

Without the internal ledgers to Bitcoin Fog, this conjecture cannot be verified. Ciphertrace compiled the following chart showing all received transactions for the likely internal consolidation address for Bitcoin Fog: 1YZJKaAx2HRWvcbCXDBtQbBZcRU46WJqw.

## What is Needed to Determine Control of Bitcoin Fog

- Bitcoin Fog Server access – the Government does not have the Bitcoin Fog servers.
- Private keys to all wallets – the Government has not provided ANY private keys to the Defense.
- Evidence as to what parties had control of the private keys to Bitcoin Fog and/or had access to the servers. – the Government has produced no evidence at all regarding these two crucial factors.

## Other Domain Name Registrations of Bitcoin Fog

Since Mr. Sterlingov has been in jail awaiting trial, someone registered the Bitcoin Fog mixing service on an Ethereum-based DNS domain.

Ethereum Name Service (ENS) registered bitcoinfog.eth on Address 0xE7f1f0657128e1eD321B0A849F12457AC5eF608F on the Ethereum Network as an ERC-721 on March 31, 2023 and is due to expire on March 30, 2024 Ref: <https://app.ens.domains/bitcoinfog.eth?tab=more>

<https://etherscan.io/tx/0xeabf670000064081aa5efe9affe72897941033c5594f96f1966b00874954f28e>

It is possible that this domain was registered prior to the ENS implementation of token standard ERC-721. ENS names are non-fungible tokens (NFTs).

## Notes on Publicly Available Tools

Chainalysis Reactor is a proprietary, black-box, forensic surveillance software that is not publicly available. It is very expensive to purchase a license. It should not be confused with publicly available explorers, which are free, and often open source.

## Conclusion


Blockchain forensics should only be used to generate investigatory leads. Standing alone, they are insufficient as a primary source of evidence. What is striking about this case is the conclusions reached without any corroborating evidence for the blockchain forensics.

The blockchain forensics and tracing tools used in this case were misused to erroneously conclude that Mr. Sterlingov was the operator of Bitcoin Fog when no such evidence exists on-chain.

The failures in the blockchain analysis in this case highlight some of the structural problems with this space. To prevent wrongful arrests like this one, and failures in compliance, like with FTX, it is recommended that Chainalysis, and their methodologies of blockchain analysis be independently audited.



Date: August 7, 2023

DocuSigned by:  
  
3928BCE51C15418

Jonelle Still, Ciphertrace