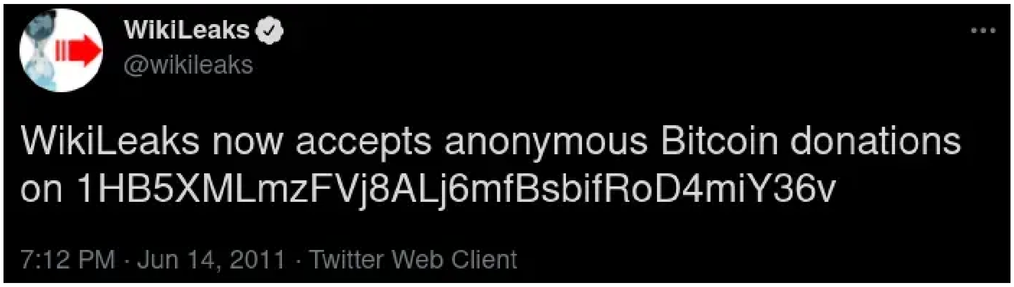
AUG 4, 2021 • 10 MIN READ • UNDERSTANDING BITCOIN PRIVACY

# Understanding Bitcoin Privacy with OXT — Part 1

## Introduction

Awareness of Bitcoin's value proposition for censorship resistant payments took off in 2011 with the launch of Silk Road and the rescue of the otherwise donation starved Wikileaks.

Simultaneously a dangerous and inaccurate narrative surrounding bitcoin transaction privacy took off: "anonymous" payments.


WikiLeaks
@wikileaks

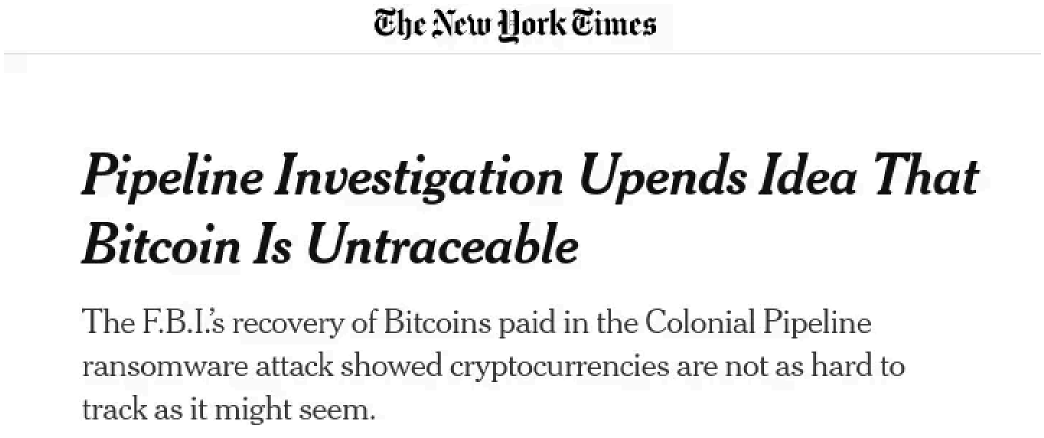
WikiLeaks now accepts anonymous Bitcoin donations on 1HB5XMLmzFVj8ALj6mfBsbifRoD4miY36v

7:12 PM · Jun 14, 2011 · Twitter Web Client

Fig 1.1 — Wikileaks Donation Tweet

This perception was never true and largely due to a lack of understanding of bitcoin's underlying functionality. Further, almost no one was watching the blockchain in 2011. But the blockchain is forever.

Today law enforcement's access to chain analytic tools has aided in many high profile take-downs. Chain surveillance firms clamor for first publication of blog posts on the latest "crime" using bitcoin for payment. The mainstays of old media are running headlines on bitcoin's traceability.



The New York Times

*Pipeline Investigation Upends Idea That Bitcoin Is Untraceable*

The F.B.I.'s recovery of Bitcoins paid in the Colonial Pipeline ransomware attack showed cryptocurrencies are not as hard to track as it might seem.

Fig 1.2 — Bitcoin is Traceable Headline

## Traceability is mainstream.

The reality of bitcoin privacy, and what users are capable of when using bitcoin to make payments, lies somewhere between perfectly anonymous and perfectly traceable.

Most tracking techniques are reliant on heuristics and evaluating the flow of bitcoin. Without fundamental privacy enhancements at the protocol level, these techniques must be undermined at the application layer.

Early in bitcoin's history, randomized wallet fingerprinting was used to defeat analytic heuristics. Tools such as custodial tumblers were used to defeat financial flow analyses at the risk of loss of funds. Today coinjoins and collaborative transactions allow for reasonable spending privacy with enhanced security over previous methods. However, much work remains to be done in automating privacy enhancements at the application layer.

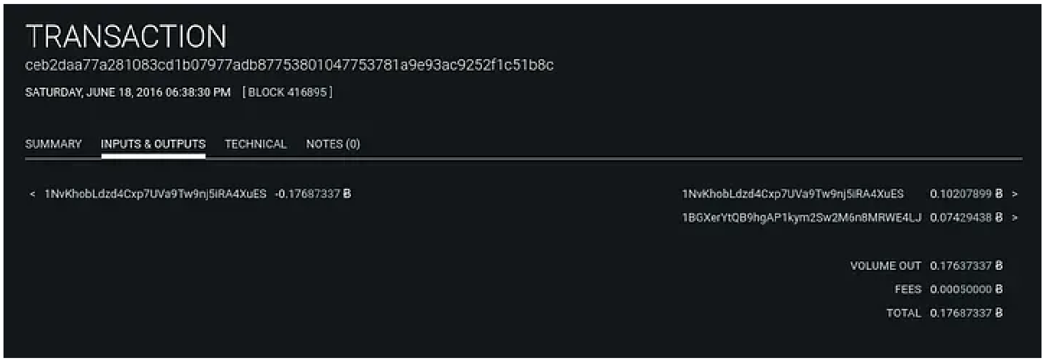
We consider education among the remaining work to aid users in obtaining a basic level of financial privacy they are typically afforded by the legacy financial network. This is our main motivation for creating this documentation. In this guide we will cover the following:

1. What "traceable" bitcoin really means including examples.
2. Defensive measures users can take to defeat tracking and why they work.
3. Practical considerations for users sending and receiving payments.

## PROBLEM STATEMENT — PSEUDONYMOUS BITCOIN

Bitcoin transaction activity is pseudonymous, not anonymous. A user's true name and personal identification information (PII) is not included in the Bitcoin protocol. This provides a very basic level of privacy, as transaction activity must be attributed to an individual.

However, bitcoin transactions are made for transparent amounts to transparent addresses. Addresses are pseudonyms that represent an actual user's activity, or more specifically, a private key's activity.



TRANSACTION

ceb2daa77a281083cd1b07977adb87753801047753781a9e93ac9252f1c51b8c

SATURDAY, JUNE 18, 2016 06:38:30 PM [ BLOCK 416895 ]

< 1NvKhobLdzd4Cxp7UVa9Tw9nj5iRA4XuES -0.17687337 Ƀ

1NvKhobLdzd4Cxp7UVa9Tw9nj5iRA4XuES 0.10207899 Ƀ >

1BGXerYtQB9hgAP1kym2Sw2M6n8MRWE4LJ 0.07429438 Ƀ >

VOLUME OUT 0.17637337 Ƀ

FEES 0.00050000 Ƀ

TOTAL 0.17687337 Ƀ

Fig. 1.3 — Example Bitcoin Transaction

Bitcoin transactions are broadcast to the bitcoin network, and archived on the bitcoin ledger. The ledger, and any transactions included in the ledger, are viewable to anyone running a bitcoin node or with access to a web based blockchain explorer.

The transparency of this public information allows for an unprecedented ease of information access that is not possible with the traditional financial system.

Users can query a blockchain explorer through a web browser and observe the entire bitcoin network's past transaction history. They can easily track the flows of bitcoin across multiple transactions.

Broadly, this practice has become synonymous with the terms traceability and chain analysis.

The most common form of chain analysis is focused on identification of transaction "change" outputs. The process is based on a series of heuristics that can be used to follow a user's activity over multiple transactions.

If this on-chain activity leads to an identified economic entity's wallet cluster, investigators may be able to obtain user PII associated with the observed transaction activity.

**From here we can glean two critical issues that can be used to attack a user privacy:**

1. "Change" outputs can be used to track a users activity on the blockchain.

2. The intersection of this activity with entities that obtain PII links observed blockchain activity with a possible real identity.

While item 2 is a critical part of what turns chain analysis into real world enforcement, this documentation will focus specifically on mitigation of the on-chain tracking.

## Heuristics

Heuristics are rules of thumb used to make decisions under uncertain conditions. They are often based on practical shortcuts. Being shortcuts heuristics are not 100% accurate.

Much of traditional chain analysis is based on heuristics. The primary heuristics have to do with "change detection" in simple spends and the clustering of separate addresses by the common input ownership heuristic.

In isolation, these heuristics can be misleading and inaccurate. But when combined with additional transaction patterns or external data such as wallet cluster labeling, the short comings of these heuristics become less damaging to attempted tracking.

## Unspent Transaction Outputs (UTXOs)

Bitcoin transactions are records of the flows of bitcoin between addresses.

Unspent transaction outputs (UTXOs) are "pieces of bitcoin" that are used to construct transactions. Transaction structure can be divided into inputs and outputs.

A user constructs a transaction by designating a payment address(es) and payment amount(s). These will be the *outputs* of the newly constructed transaction.

Typically wallet software will complete the transaction by algorithmically including UTXOs from previous transactions as *inputs* to the new transaction.

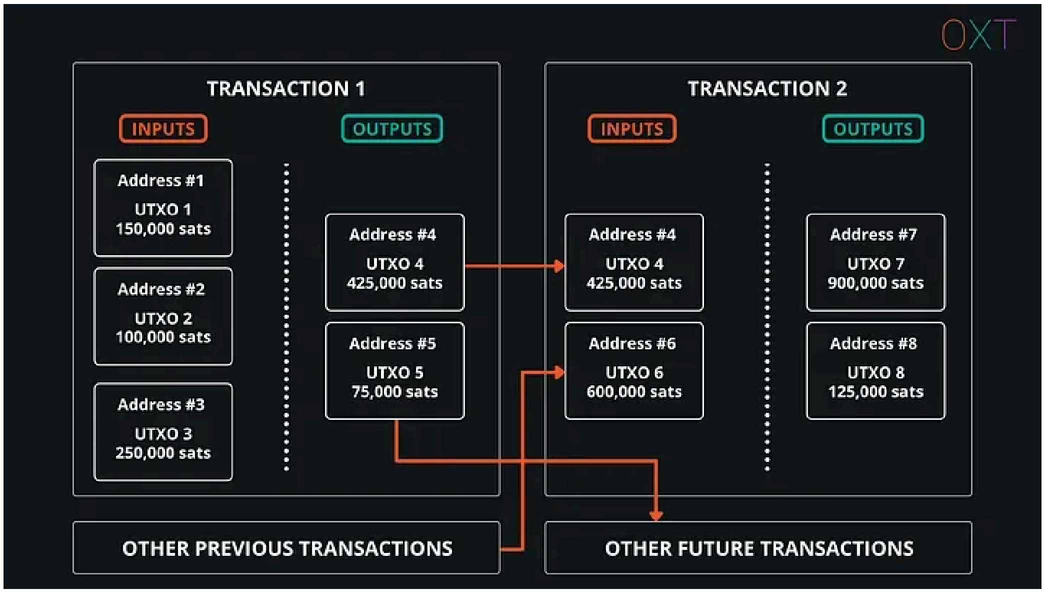


Fig. 1.4 — Unspent Transaction Outputs (UTXOs)

There is a difference between an address and a UTXO. UTXOs are "pieces of bitcoin" paid to an address. The simplest way to visualize the difference is to understand that an address can receive multiple UTXOs, a process detrimental to privacy referred to as "address reuse".

## Example Transactions

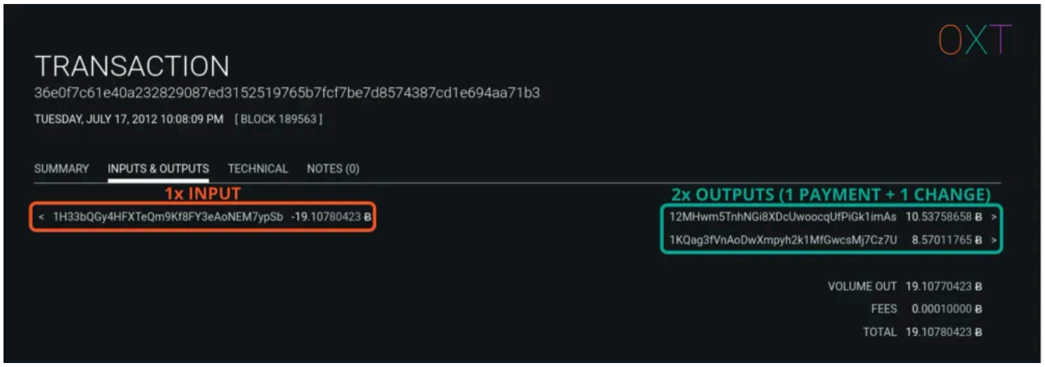
There are several common categories of bitcoin transactions. These categories are inferred based on a transactions input and output profile along with our experience and observed cluster labeling. Examples of the major transaction types are provided below.

## Simple Send

Simple spends are some of the most common transaction types and make up roughly 50% of bitcoin transactions in recent blocks (Source: transactionfee.info).

These transaction types are evidence of typical user behavior, where a user makes a payment and receives a change output.

***Traits:***
- Number of Inputs: 1 (or more)
- Number of Outputs: 2
- Common Interpretation: 1 payment output, 1 change output



TRANSACTION
36e0f7c61e40a232829087ed3152519765b7fcf7be7d8574387cd1e694aa71b3
TUESDAY, JULY 17, 2012 10:08:09 PM [ BLOCK 189563 ]

< 1H33bQGy4HFXTeQm9Kf8FY3eAoNEM7ypSb -19.10780423 B

12MHwm5TnhNGi8XDcUwoocqUfPiGk1imAs 10.53758658 B >
1KQag3fVnAoDwXmpyh2k1MfGwcsMj7Cz7U 8.57011765 B >

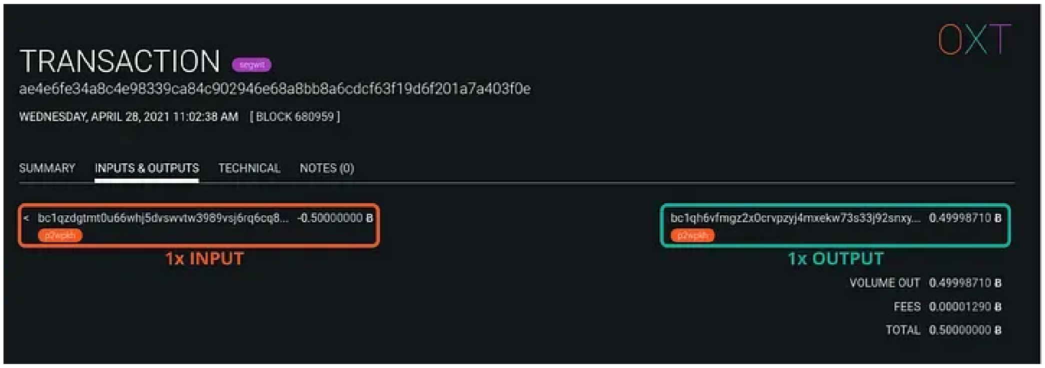
VOLUME OUT 19.10770423 B
FEES 0.00010000 B
TOTAL 19.10780423 B

Fig. 1.5 — Simple Spend Transaction Example

## Sweep

“Sweeps” spend the entirety of a single UTXO to a new address.

***Traits:***
- Number of Inputs: 1
- Number of Outputs: 1
- Common Interpretation: Possible self-transfer



TRANSACTION
ae4e6fe34a8c4e98339ca84c902946e68a8bb8a6cdcf63f19d6f201a7a403f0e
WEDNESDAY, APRIL 28, 2021 11:02:38 AM [ BLOCK 680959 ]

< bc1qzdgtmt0u66whj5dvswvtw3989vsj6rq6cq8... -0.50000000 B

bc1qh6vfmgz2x0crvpzyj4mxekw73s33j92snxy... 0.49998710 B

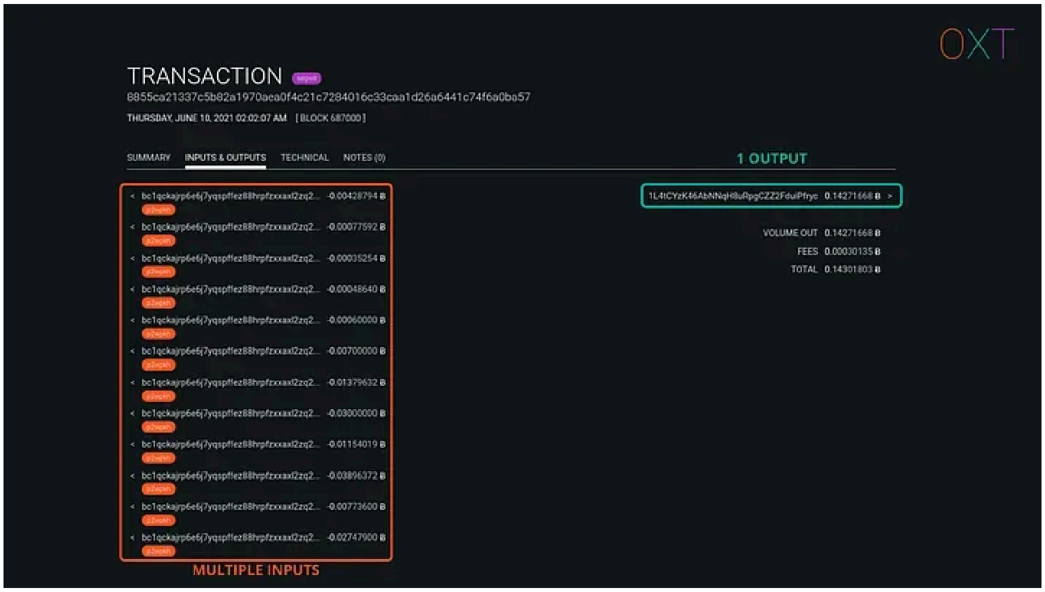
VOLUME OUT 0.49998710 B
FEES 0.00001290 B
TOTAL 0.50000000 B

Fig. 1.6 — “Sweep” Example Transaction

## Consolidation Spend

Consolidation transactions combine multiple UTXOs into a single UTXO. These are rarely “true payments” because a normal payment has a change output.

***Traits:***
- Number of Inputs: >1
- Number of Outputs: 1
- Common Interpretation: Possible self-transfer



TRANSACTION
8855ca21337c5b82a1970aea0f4c21c7284016c33caa1d26a6441c74f6a0ba57
THURSDAY, JUNE 10, 2021 02:02:07 AM [BLOCK 687000]

SUMMARY | INPUTS & OUTPUTS | TECHNICAL | NOTES (0)

| | |
|---|---|
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.00428794 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.00077592 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.00035254 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.00048640 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.00060000 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.00700000 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.01379632 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.03000000 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.01154019 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.03896372 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.00773600 ฿ |
| < bc1qckajrp6e6j7yqspflez88hrpfzxxaxl2zq2... | -0.02747900 ฿ |

MULTIPLE INPUTS

1 OUTPUT

1L4tCYzK46AbNNqH8uRpgCZZ2FduiPfryc 0.14271668 ฿ >

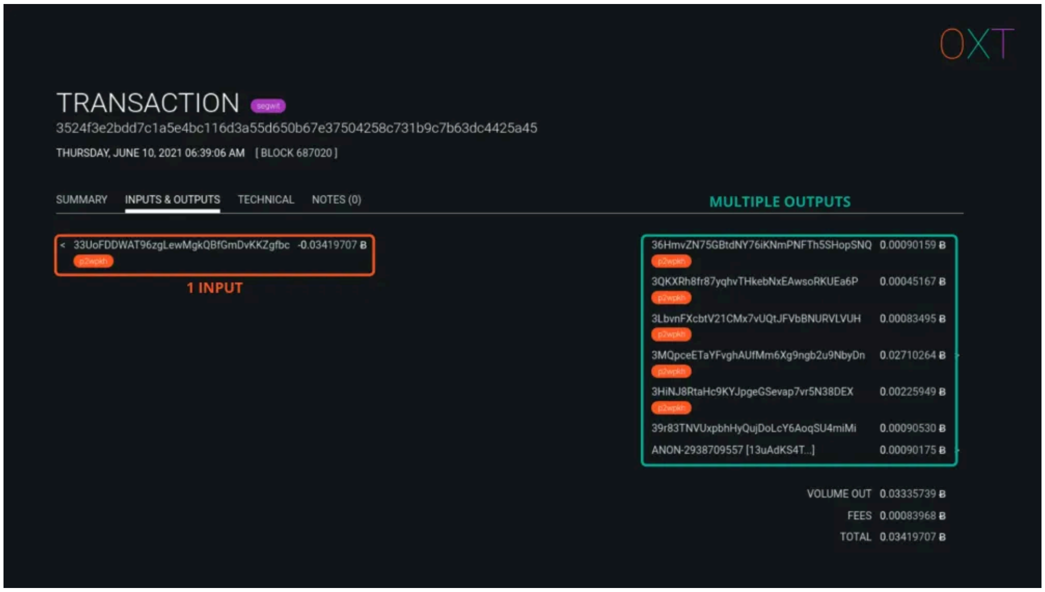
| | |
|---|---|
| VOLUME OUT | 0.14271668 ฿ |
| FEES | 0.00030135 ฿ |
| TOTAL | 0.14301803 ฿ |

Fig. 1.7 — Consolidation Spend Example Transaction

## Batch Spend

Batch spends are most likely performed by exchanges and include 1 or more inputs and many outputs. These transactions aim to save on miner fees by making as many payments as possible in a single transaction.

***Traits:***
- Number of Inputs: ≥1
- Number of Outputs: MANY
- Common Interpretation: Large economic activity, likely exchange



TRANSACTION
3524f3e2bdd7c1a5e4bc116d3a55d650b67e37504258c731b9c7b63dc4425a45
THURSDAY, JUNE 10, 2021 06:39:06 AM [ BLOCK 687020 ]

```
< 33UoFDDWAT96zgLewMgkQBfGmDvKKZgfbc  -0.03419707 Ƀ
```

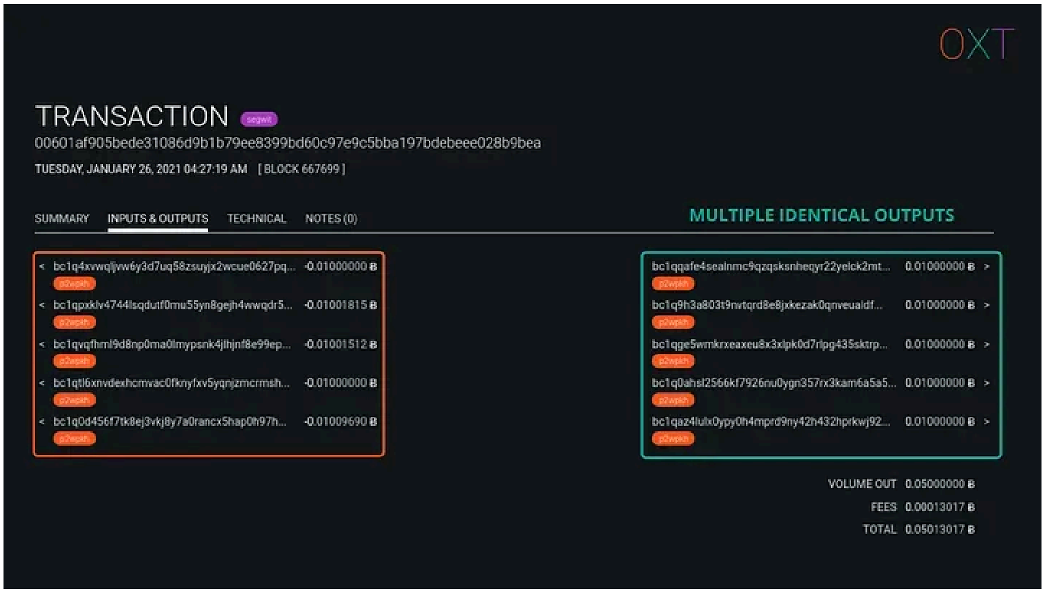
```
36HmvZN75GBtdNY76iKNmPNFTh5SHopSNQ  0.00090159 Ƀ
3QKXRh8fr87yqhvTHkebNxEAwsoRKUEa6P  0.00045167 Ƀ
3LbvnFXcbtV21CMx7vUQtJFVbBNURVLVUH  0.00083495 Ƀ
3MQpceETaYFvghAUfMm6Xg9ngb2u9NbyDn  0.02710264 Ƀ
3HiNJ8RtaHc9KYJpgeGSevap7vr5N38DEX  0.00225949 Ƀ
39r83TNVUxpbhHyQujDoLcY6AoqSU4miMi  0.00090530 Ƀ
ANON-2938709557 [13uAdKS4T...]      0.00090175 Ƀ

                        VOLUME OUT  0.03335739 Ƀ
                              FEES  0.00083968 Ƀ
                             TOTAL  0.03419707 Ƀ
```

Fig. 1.8 — Batch Spend Example Transaction

## Multi-party Transactions (Coinjoin)

Multi-party transactions involve collaboration between many users to perform a single transaction that improves participant privacy. These transactions are easily identified by their equal output amounts.

***Traits:***
- Number of Inputs: MANY
- Number of Outputs: MANY
- Output Profile: Number of identical outputs is a proxy for number of participants.



TRANSACTION
00601af905bede31086d9b1b79ee8399bd60c97e9c5bba197bdebeee028b9bea
TUESDAY, JANUARY 26, 2021 04:27:19 AM [ BLOCK 667699 ]

```
< bc1q4xvwqljvw6y3d7uq58zsuyjx2wcue0627pq...  -0.01000000 Ƀ
< bc1qpxklv4744lsqdutf0mu55yn8gejh4wwqdr5...  -0.01001815 Ƀ
< bc1qvqfhml9d8np0ma0lmypsnk4jlhjnf8e99ep...  -0.01001512 Ƀ
< bc1qtl6xnvdexhcmvac0fknyfxv5yqnjzmcrmsh...  -0.01000000 Ƀ
< bc1q0d456f7tk8ej3vkj8y7a0rancx5hap0h97h...  -0.01009690 Ƀ
```

```
bc1qqafe4sealnmc9qzqsksnheqyr22yelck2mt...  0.01000000 Ƀ  >
bc1q9h3a803t9nvtqrd8e8jxkezak0qnveualdf...  0.01000000 Ƀ  >
bc1qge5wmkrxeaxeu8x3xlpk0d7rlpg435sktrp...  0.01000000 Ƀ  >
bc1q0ahsl2566kf7926nu0ygn357rx3kam6a5a5...  0.01000000 Ƀ  >
bc1qaz4lulx0ypy0h4mprd9ny42h432hprkwj92...  0.01000000 Ƀ  >

                              VOLUME OUT  0.05000000 Ƀ
                                    FEES  0.00013017 Ƀ
                                   TOTAL  0.05013017 Ƀ
```

Fig. 1.9 — Coinjoin Example Transaction

## Change Detection — Simple Spend Interpretations

In this section we are going to introduce transaction interpretations for simple spends (1 input, 2 outputs). This section will focus on the most common interpretation of these transactions, where one output is interpreted as a payment and the remaining output is interpreted as change back to the original wallet. A complete list of interpretations under this model are presented in Part III. The goal of this section is to familiarize the reader with the heuristics used to evaluate simple spends and detect a change output.

When observing a single transaction in isolation, we are presented with a limited amount of data that is included within the transaction. We refer to this information as internal transaction data.

**Internal transaction data is limited to the following**:

- amounts (input, output, miner fees)
- input address script type
- output address script type
- transaction version number
- transaction locktime
- replace by fee signalling

## Change Detection

Due to miner fees, a normal bitcoin payment always requires spending a larger output amount than the intended payment amount.

When a single UTXO for a larger amount than the intended payment amount is consumed, the user will receive a change UTXO for the remainder back to an address generated by their wallet's private key.

Much of traditional chain analysis is based on detecting this change output. If a change output can be successfully detected, a single user's activity can be tracked across a series of transactions.

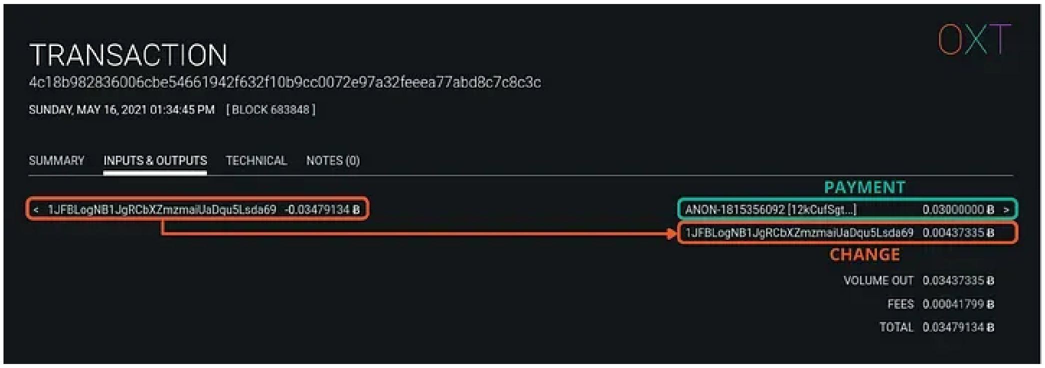
In this section we will detect change outputs based on internal transaction data. The heuristics are presented roughly in the order of decreasing effectiveness accuracy. Examples will be presented and annotated screenshots shown for aiding users in visualizing these concepts.

## Address Reuse

Addresses are created by a single private key. Multiple uses of the same address are signs of activity of the same private key.

In a simple spend, if one output is to a new address, and the remaining output is to the same address as the input address, we know that the reused address is the change output.

Today, most bitcoin wallets will automatically generate a new address for receiving change outputs. However, a wallet can be configured to receive change to the same address as the transaction input. This behavior is typically indicative of centralized service wallet activities or old versions of the bitcoin core wallet.



TRANSACTION
4c18b982836006cbe54661942f632f10b9cc0072e97a32feeea77abd8c7c8c3c
SUNDAY, MAY 16, 2021 01:34:45 PM [ BLOCK 683848 ]

< 1JFBLogNB1JgRCbXZmzmaiUaDqu5Lsda69 -0.03479134 Ƀ

ANON-1815356092 [12kCufSgt...] 0.03000000 Ƀ >
1JFBLogNB1JgRCbXZmzmaiUaDqu5Lsda69 0.00437335 Ƀ

VOLUME OUT 0.03437335 Ƀ
FEES 0.00041799 Ƀ
TOTAL 0.03479134 Ƀ

Fig. 1.10 — Address Reuse Example Transaction

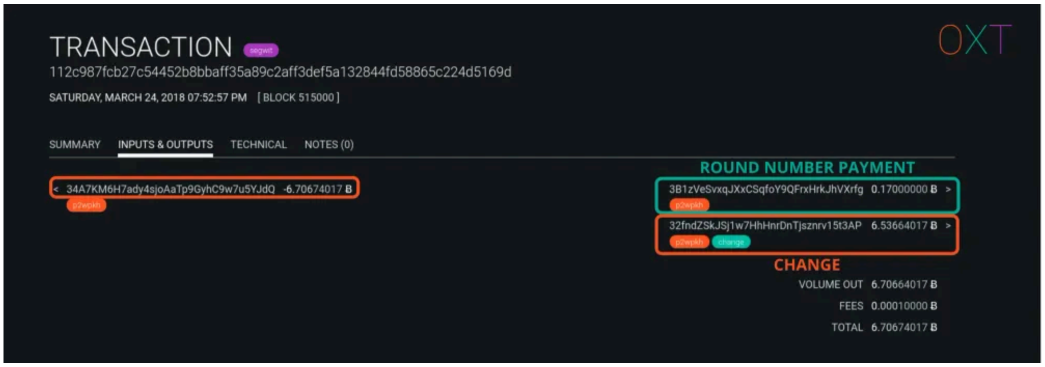
## Round Number Payment Heuristic

When a user initiates a payment, they input the payment destination (address), payment amount, and a network transaction fee into their wallet software. The wallet software will select input UTXOs and generate the change output (if there is one).

The change amount in a simple spend is calculated as follows:

> input amount - payment amount - tx size (in vbytes) * network fee rate (in sat/vbyte)

It's difficult for a user to generate a change output for a "round number amount" on purpose. In a simple spend, the round number output is the likely payment which makes the remaining output the change output.



TRANSACTION
112c987fcb27c54452b8bbaff35a89c2aff3def5a132844fd58865c224d5169d
SATURDAY, MARCH 24, 2018 07:52:57 PM [ BLOCK 515000 ]

< 34A7KM6H7ady4sjoAaTp9GyhC9w7u5YJdQ -6.70674017 Ƀ

3B1zVeSvxqJXxCSqfoY9QFrxHrkJhVXrfg 0.17000000 Ƀ >
32fndZSkJSj1w7HhHnrDnTjsznrv15t3AP 6.53664017 Ƀ >

VOLUME OUT 6.70664017 Ƀ
FEES 0.00010000 Ƀ
TOTAL 6.70674017 Ƀ

Fig. 1.11 — Round Number Payment Heuristic Example Transaction

## Different Script Type Heuristic

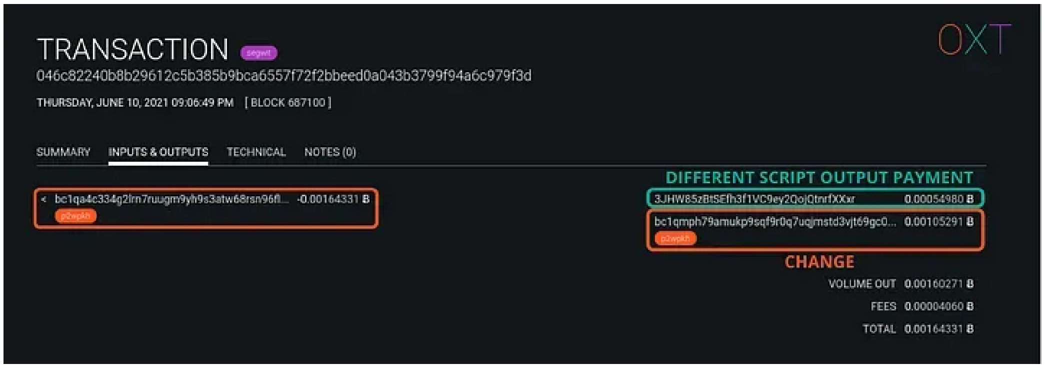
There are several bitcoin address script types. Readers will be most familiar with the following script types:

- Pay to pubkey hash (P2PKH): Addresses starting with a 1
- Pay to script script hash (P2SH): Addresses starting with a 3

- Native Segwit (version 0)(bech32): Addresses starting with bc1q

For a given input script type, if one output is to the same type as the input and the remaining output is to a new address script type, the output to the new address script type is the likely payment.Which makes the output to the same address address script type the likely change output.

In other words, an output to a different script type is the likely payment output. This heuristic can also be combined with the round number payment heuristic.



TRANSACTION
046c82240b8b29612c5b385b9bca6557f72f2bbeed0a043b3799f94a6c979f3d
THURSDAY, JUNE 10, 2021 09:06:49 PM [ BLOCK 687100 ]

< bc1qa4c334g2lrn7ruugm9yh9s3atw68rsn96fl... -0.00164331 B

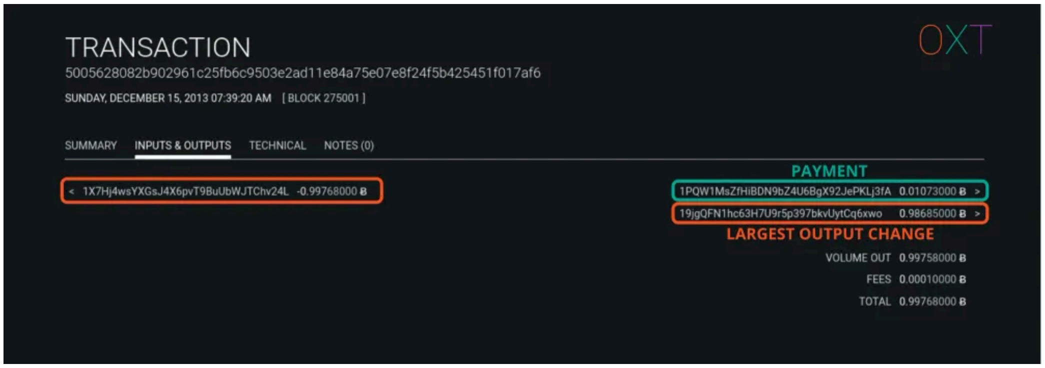
3JHW85zBtSEfh3f1VC9ey2QojQtnrfXXxr 0.00054980 B
bc1qmph79amukp9sqf9r0q7uqjmstd3vjt69gc0... 0.00105291 B

VOLUME OUT 0.00160271 B
FEES 0.00004060 B
TOTAL 0.00164331 B

Fig. 1.12 — Different Script Output Type Heuristic Example Transaction

## Largest Output Amount Heuristic

Another simple heuristic, assumes that the largest output amount is the likely change output. This is one of the weakest heuristics, particularly when taken in isolation, but as we will discuss in Part 2, this heuristic can be helpful in a transaction graph analysis.



TRANSACTION
5005628082b902961c25fb6c9503e2ad11e84a75e07e8f24f5b425451f017af6
SUNDAY, DECEMBER 15, 2013 07:39:20 AM [ BLOCK 275001 ]

< 1X7Hj4wsYXGsJ4X6pvT9BuUbWJTChv24L -0.99768000 B

1PQW1MsZfHiBDN9bZ4U6BgX92JePKLj3fA 0.01073000 B >
19jgQFN1hc63H7U9r5p397bkvUytCq6xwo 0.98685000 B >

VOLUME OUT 0.99758000 B
FEES 0.00010000 B
TOTAL 0.99768000 B

Fig. 1.13 — Largest Output Heuristic Example Transaction

## Review And Preview

In this section we introduced many of the basic concepts around bitcoin transaction privacy. Pseudonymous bitcoin provides a basic level of privacy that is not directly associated with personal identifiable information.

However, the openness of the bitcoin protocol allows for the tracking of bitcoin flows. This concept is broadly referred to as “chain analysis”. A core focus of chain analysis focuses on detection of change outputs by making use of several heuristics. If successful, these heuristics allow for tracking a single users activity over multiple transactions.

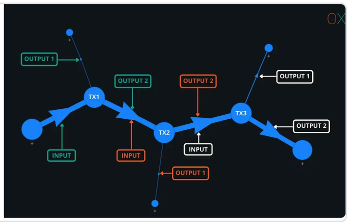
**Part 2 will introduce core concepts of chain analysis such as:**

1. External transaction data, which can be used to weaken transaction privacy.
2. Transaction graph analysis, a main tool for tracking the flow of bitcoin over multiple transactions.
3. The common input ownership heuristic, also known as "wallet clustering" and its implication on the bitcoin network.
4. **This section also includes a walk-through and examples of OXTs transaction graph tool and wallet clustering scheme.**

**Understanding Bitcoin Privacy with OXT — Part 2**

Introduction Now that we have presented the basic concepts in Part 1, we can move on to the core concepts that underpin chain analysis by…

OXT ErgoBTC • OXT Research



**Part 3 cover core concepts of improving bitcoin's privacy including:**

1. Randomized wallet fingerprinting for defeating change detection.
2. UTXO flows and the fundamental link between inputs and outputs.
3. How equal output coinjoins address the issue of deterministic flows.
4. Transaction entropy
5. How payjoin defeats the common input ownership heuristic

**Understanding Bitcoin Privacy with OXT — Part 3**

Introduction With the foundational concepts of chain analysis introduced in Part 1 and Part 2, Part 3 will discuss the methods for…

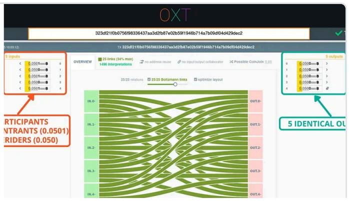
OXT ErgoBTC • OXT Research

**Part 4 discusses:**

1. Analyses needing a "starting point"
2. The privacy implications of sending and receiving payments
3. How existing privacy techniques can mitigate many of the issues discussed throughout the guide.

**Understanding Bitcoin Privacy with OXT — Part 4**

Introduction So far this guide has introduced the basics concepts used by chain analysis and the concepts used to undermine chain analysis…



OXT ErgoBTC • OXT Research

Published by:

## You might also like...

| | | |
|---|---|---|
| AUG 11 | Understanding Bitcoin Privacy with OXT — Part 4 | 10 min read |
| AUG 09 | Understanding Bitcoin Privacy with OXT — Part 3 | 10 min read |
| AUG 06 | Understanding Bitcoin Privacy with OXT — Part 2 | 11 min read |

Sign up

Powered by Ghost

Subscribe