Cse291-J: Blockchain Security

Deian Stefan and Stefan Savage, Spring 2024

Blockchain tracing

Up front: today's paper

Several of us were involved and it is the subject of a bestselling book

A Fistful of Bitcoins: Characterizing Payments Among Men with No Names

Sarah Meiklejohn Marjori Pomarole Grant Jordan
Kirill Levchenko Damon McCoy[†] Geoffrey M. Voelker Stefan Savage
University of California, San Diego George Mason University[†]

ABSTRACT

Bitcoin is a purely online virtual currency, unbacked by either physical commodities or sovereign obligation; instead, it relies on a combination of cryptographic protection and a peer-to-peer protocol for witnessing settlements. Consequently, Bitcoin has the unintuitive property that while the ownership of money is implicitly anonymous, its flow is globally visible. In this paper we explore this unique characteristic further, using heuristic clustering to group Bitcoin wallets based on evidence of shared authority, and then using re-identification attacks (i.e., empirical purchasing of goods and services) to classify the operators of those clusters. From this analysis, we characterize longitudinal changes in the Bitcoin market, the stresses these changes are placing on the system, and the challenges for those seeking to use Bitcoin for criminal or fraudulent purposes at scale.

Categories and Subject Descriptors

K.4.4 [**Electronic Commerce**]: Payment schemes

Keywords

Bitcoin; Measurement; Anonymity

1. INTRODUCTION

Demand for low friction e-commerce of various kinds has driven a proliferation in online payment systems over the last decade. Thus, in addition to established payment card networks (e.g., Visa and Mastercard) a broad range of so-called "alternative payments" has emerged including eWallets (e.g., Paypal, Google Checkout, and WebMoney), direct debit systems (typically via ACH, such as eBillMe), money transfer systems (e.g., Moneygram) and so on. However, virtually all of these systems have the property that they are denominated in existing fiat currencies (e.g., dollars), explicitly identify the payer in transactions, and are centrally or quasi-centrally administered.[1]

[1]In particular, there is a central controlling authority who has the technical and legal capacity to tie a transaction back to a pair of individuals.

IMC'13, October 23–25, 2013, Barcelona, Spain.

http://dx.doi.org/10.1145/2504730.2504747.

By far the most intriguing exception to this rule is Bitcoin. First deployed in 2009, Bitcoin is an independent online monetary system that combines some of the features of cash and existing online payment methods. Like cash, Bitcoin transactions do not explicitly identify the payer or the payee: a transaction is a cryptographically-signed transfer of funds from one public key to another. Moreover, like cash, Bitcoin transactions are irreversible (in particular, there is no *chargeback* risk as with credit cards). However, unlike cash, Bitcoin requires third party mediation: a global peer-to-peer network of participants validates and certifies all transactions; such decentralized accounting requires each network participant to maintain the entire transaction history of the system, currently amounting to over 3GB of compressed data. Bitcoin identities are thus *pseudo-anonymous*: while not explicitly tied to real-world individuals or organizations, all transactions are completely transparent.[2]

This unusual combination of features has given rise to considerable confusion about the nature and consequences of the anonymity that Bitcoin provides. In particular, there is concern that the combination of scalable, irrevocable, anonymous payments would prove highly attractive for criminals engaged in fraud or money laundering. In a widely leaked 2012 Intelligence Assessment, FBI analysts make just this case and conclude that a key "advantage" of Bitcoin for criminals is that "law enforcement faces difficulties detecting suspicious activity, identifying users and obtaining transaction records" [7]. Similarly, in a late 2012 report on Virtual Currency Schemes, the European Central Bank opines that the lack of regulation and due diligence might enable "criminals, terrorists, fraudsters and money laundering" and that "the extent to which any money flows can be traced back to a particular user is unknown" [6]. Indeed, there is at least some anecdotal evidence that this statement is true, with the widely publicized "Silk Road" service using Bitcoin to trade in a range of illegal goods (e.g., restricted drugs and firearms). Finally, adding to this urgency is Bitcoin's considerable growth, both quantitatively — a merchant service, Bitpay, announced that it had signed up over 1,000 merchants in 2012 to accept the currency, and in April 2013 the exchange rate soared to 235 USD per bitcoin before settling to a more modest 100 USD per bitcoin — and qualitatively via integration with existing payment mechanisms (e.g., Bitinstant offering to tie users' Bitcoin wallets to Mastercard accounts [5] and Bitcoin Central's recent partnership with the French bank Crédit Mutuel Arkéa to gateway Bitcoin into the banking system [16]) and the increasing attention of world financial institutions (e.g., Canada's recent decision to tax Bitcoin transactions [3] and FinCEN's recent regulations

[2]Note that this statement is not strictly true since private exchanges of Bitcoin between customers of a single third party exchange, such as Mt. Gox, need not (and do not) engage the global Bitcoin protocol and are therefore not transparent.

In the Murky World of Bitcoin, Fraud Is Quicker Than the Law

By NATHANIEL POPPER DECEMBER 5, 2013, 6:58 PM 186 Comments

Sandy Huffaker for The New York Times

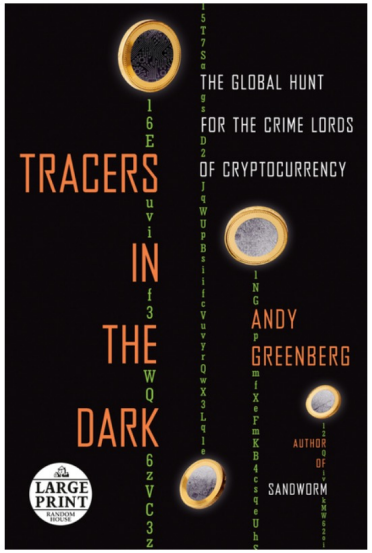
Sarah Meiklejohn, a graduate student at the University of California, San Diego, who researches the Bitcoin economy.

The call went out on Twitter: "For insane profits come and join the pump."

It was an invitation to a penny stock-style pump-and-dump scheme — only this one involved Bitcoin, the soaring, slightly scary virtual currency that has beckoned and bewildered people around the world.

While such bid 'em up, sell 'em off scams are shut down in the financial markets all the time, this one and other frauds involving digital money have gone unchecked. The reason, in no small part: Government authorities do not agree on which laws apply to Bitcoin — or even on what Bitcoin is.

The person behind the recent scheme, a trader known on Twitter as Fontas, said in a secure Internet chat that he operated with little fear of a crackdown.



Historical context: what is the relationship between identity and payments?

Early forms of payment (medium of exchange)

- Barter
- Commodities (e.g., , gold)
- Who are counterparties? Role of identity?

Paper currency

- Starts in 7th century BCE, real economy ~1260CE – both in China; Europe catches up in ~400 years)
- Advantage of paper currency over barter/commodities?
- Any new issues wrt counterparties or identity?

Banking (both holding proceeds and lending)

- Goes back ~2000 CE in Mesopotamia; grew significantly in Greece and then the Roman Empire (developed "bills of exchange")
- Double entry-bookkeeping – popularized by Italian city states in 14th and 15th centuries (Luca Pacioli – father of accounting)
- Any new issues wrt counterparties or identity?

Modern version in US

Cash is effectively anonymous, but hard to move in bulk, typically in-person payment

- Hence the significant problems for criminal economies (e.g., drugs) dealing with cash

Checks, wires and associated payment systems used for larger sums

- Account number to account number; records maintained by financial institution

The growth of anti-money laundering regulation

1970 Bank Secrecy Act

- Not really about secrecy
  - In fact, goal is to make data available to the government
  - Brought to Supreme court on 4th and 5th amendment grounds; CA Bankers Assn v Schultz; upheld
- Creates record keeping/reporting requirements (and penalties) on financial institutions
  - CTR, SAR, FBAR, MIL, CMIR (all go to Financial Crimes Enforcement Network, FinCEN, branch of Treasury)
- Customer Due Diligence (CDD) rule – start of modern "Know You Customer" (KYC) rules
  - Requirement on FS to know and verify identity of **own** customers

Vastly enhanced via amendments and other laws post 9/11

- Cornerstone of modern anti-money laundering (AML) regime
- Used extra-territorially; but also ubiquitous across foreign countries (perhaps except DPRK)

Modern goals

- Identify and target money laundering by criminal enterprises; or around geo-political conflict (e.g., sanctions/OFAC)

Crypto has entered the chat…

Bitcoin (and other kinds of virtual currencies) show up

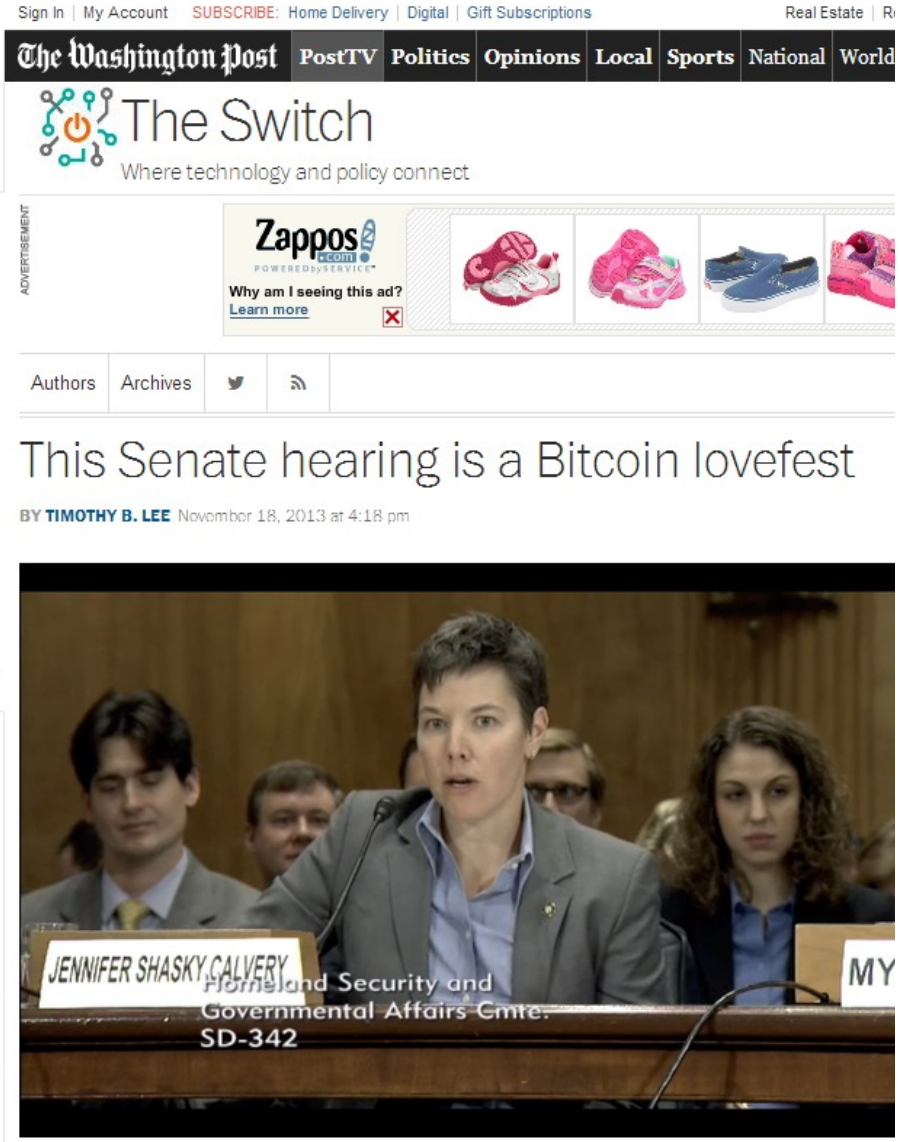
- Its not regulated
- Its not clear what it is even (is it a security? a commodity? a currency?)
- It doesn't matter until people start to use it
  (particularly exchanging it for "fiat" currency)

Issue

- No regulated entity is maintaining records
- There are no "accounts" per se
- No KYC
- What happens if there is drug, terrorism, etc money laundering via Bitcoin?

Much wringing of hands…

- Don't want to overregulate new innovative technology, but don't want rampant fraud



US Govt takes light touch in 2013

- Hands off on pure-crypto transactions

Only requirement is on money service brokers (MSBs) to have KYC

- i.e., if you exchange Bitcoin for $, then you need to know who your counterparties are

The paper you read for today is an integral part of the reason why

Blockchains and identity

We'll use Bitcoin has the example (UTXO style)

- Most other blockchains similar or easier to trace (e.g., ETH)
- A few exceptions (Monera, Zcash)

Bitcoin is not anonymous like cash, but it is pseudonomous

- Transactions are transparent on blockchain (i.e., everyone sees X sent q BTC to Z)

  More transparent than existing banking actually – all transactions public (modulo private, off chain, transactions)

- But counterparties are just public key hashes (i.e., don't known who is X or Z)

Key question: how anonymous is Bitcoin-style pseudonymity?

Transaction inference

What do we know about this transaction?

Not much, address A sent 3 BTC to address B. That's it.

Transaction inference

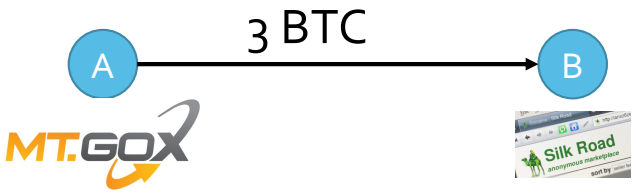
What do we know about this transaction?

What if **I** tell you that A is an address operated by MtGox?

and B is an address operated by Silk Road?

Transaction inference

What do we know about this transaction?

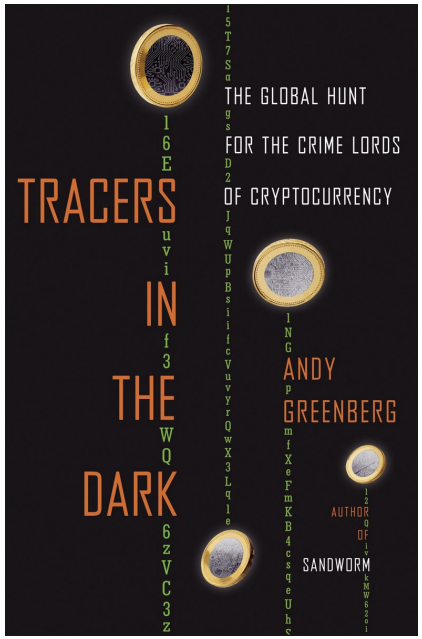


But, how would I know such things?

- **Reidentification via Side effects**

  Crypto is anonymous until you use it, but once used its associated with whatever side effects it had (i.e., drug package is shipped, dollars exchanged for Bitcoin, etc)

- Ok, but that's very abstract – what does one concretely do to figure this out?



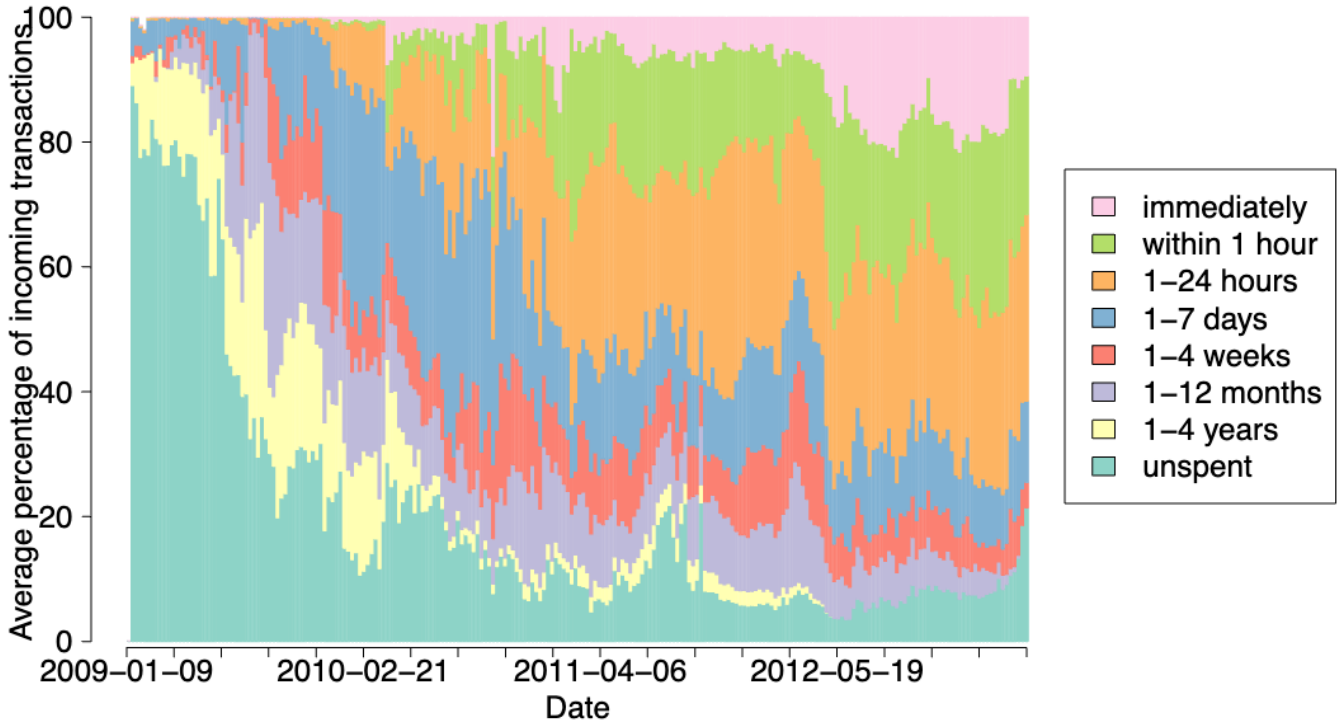
researchers' tracing work. As another UCSD professor working on the project, Geoffrey Voelker, described it at the time, "Our secret weapon is shopping."

So when Meiklejohn began talking over her Bitcoin tracking project with Savage, the two agreed she should take the same approach: She would manually identify Bitcoin addresses one by one by doing transactions with them herself, like a cop on the narcotics beat carrying out buy-and-busts. To

Mining		
50 BTC	BTC Guild	Itzod
ABC Pool	Deepbit	Ozcoin
Bitclockers	EclipseMC	Slush
Bitminter	Eligius	
Wallets		
Bitcoin Faucet	Easywallet	Strongcoin
My Wallet	Flexcoin	WalletBit
Coinbase	Instawallet	
Easycoin	Paytunia	
Exchanges		
Bitcoin 24	BTC-e	Aurum Xchange
Bitcoin Central	CampBX	BitInstant
Bitcoin.de	CA VirtEx	Bitcoin Nordic
Bitcurex	ICBit	BTC Quick
Bitfloor	Mercado Bitcoin	FastCash4Bitcoins
Bitmarket	Mt Gox	Lilion Transfer
Bitme	The Rock	Nanaimo Gold
Bitstamp	Vircurex	OKPay
BTC China	Virwox	
Vendors		
ABU Games	BTC Buy	HealthRX
Bitbrew	BTC Gadgets	JJ Games
Bitdomain	Casascius	NZBs R Us
Bitmit	Coinabul	Silk Road
Bitpay	CoinDL	WalletBit
Bit Usenet	Etsy	Yoku
Gambling		
Bit Elfin	BitZino	Gold Game Land
Bitcoin 24/7	BTC Griffin	Satoshi Dice
Bitcoin Darts	BTC Lucky	Seals with Clubs
Bitcoin Kamikaze	BTC on Tilt	
Bitcoin Minefield	Clone Dice	
Miscellaneous		
Bit Visitor	Bitfog	CoinAd
Bitcoin Advertisers	Bitlaundry	Coinapult
Bitcoin Laundry	BitMix	Wikileaks

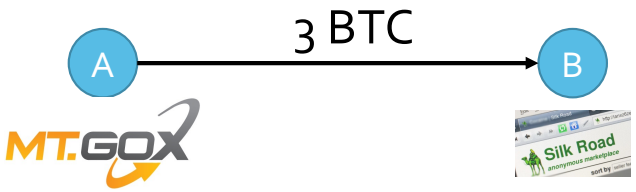
Table 1: The various services we interacted with, grouped by (approximate) type.

Crypto is moving all the time – you tend to hit something



Transaction inference

What do we know about this transaction?



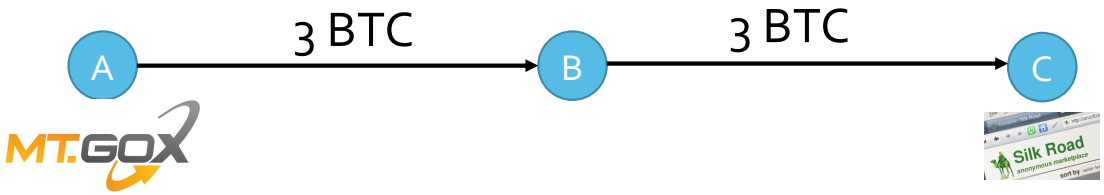
Ok… but so what? MtGox and SilkRoad are not people, they're institutions – those addresses don't correspond to a single persron

- They have records; and legit organizations (e.g., MtGox) are subject to legal process
- Grand Jury subpoenas (and similar) can demand all business records pursuant to a transaction and remember KYC?

  "Please give me all information about the account associated with this transaction, their name/address/real-world bank account, and all other transactions they did"
- Note – identifying recipient or sender may deanonymize **both**

Transaction inference

Ok, but that was a very simple case… what if you don't know both sides?



I don't know who B is

But I can look at what B does in the future…

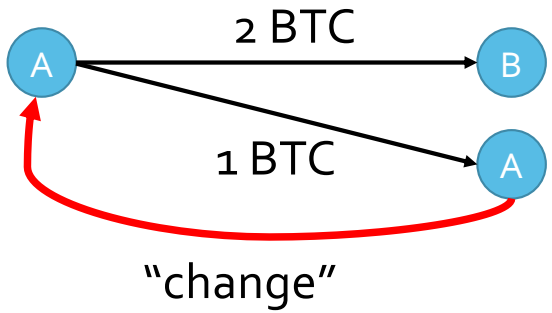
Ok, but then isn't this just trivial?

No… there are several problems.

Remember the UTXO model

- What if A has 3 BTC and only wants to send 2 to B? Still has to **send all three**

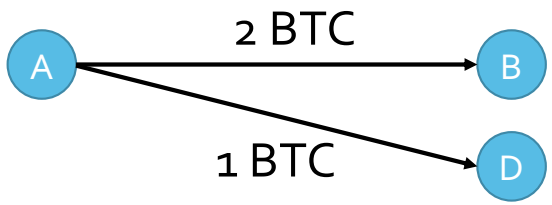
One way to do this makes it clear what happened



But what if A wasn't a dumbass?

Creates a new address D to hold the change

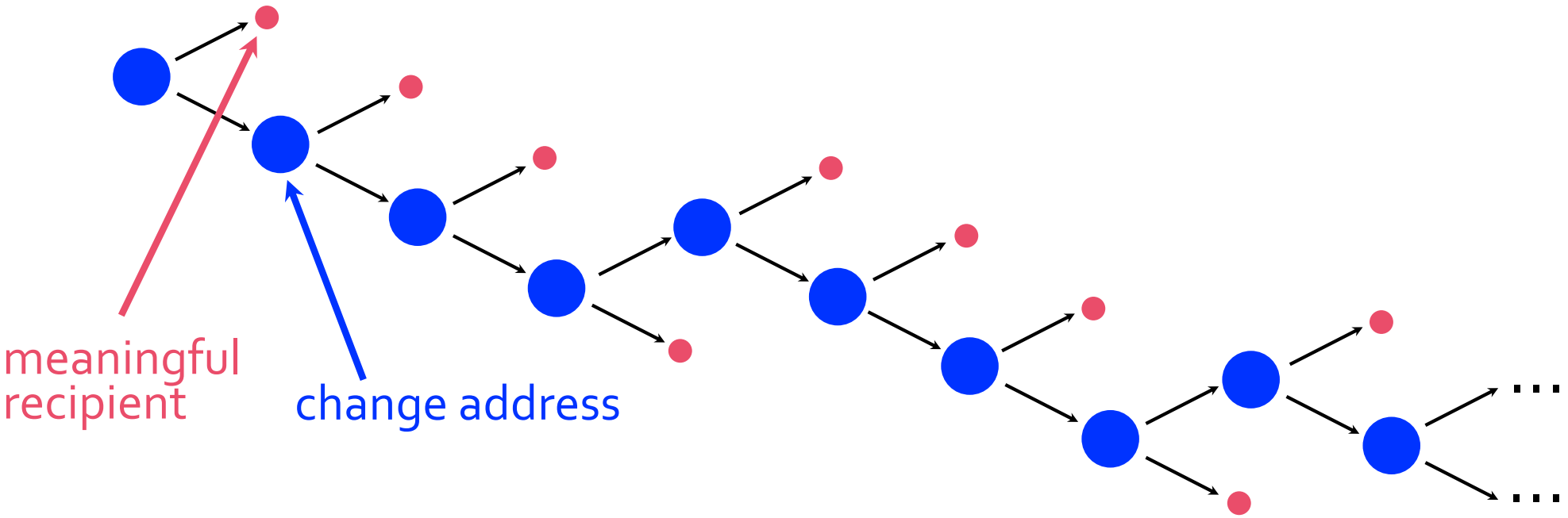
- **We** know that B is the recipient and D is the change, but how does an analyst?



And this could keep going…

Peeling chains – collection of change addresses

- Create new one after each purchase and use until they're empty



Clustering addresses by input user

All these addresses belong to the same user (generally foolproof)

How do we know this?

Clustering addresses by change user

This address **also** belongs to the same user (heuristic)

Change address heuristics

First, if address ever is ever a **co-input** with another address they belong to the same user (again, shared control over private keys)

Fistful of Bitcoins heuristic (2013 Meiklejohn paper)

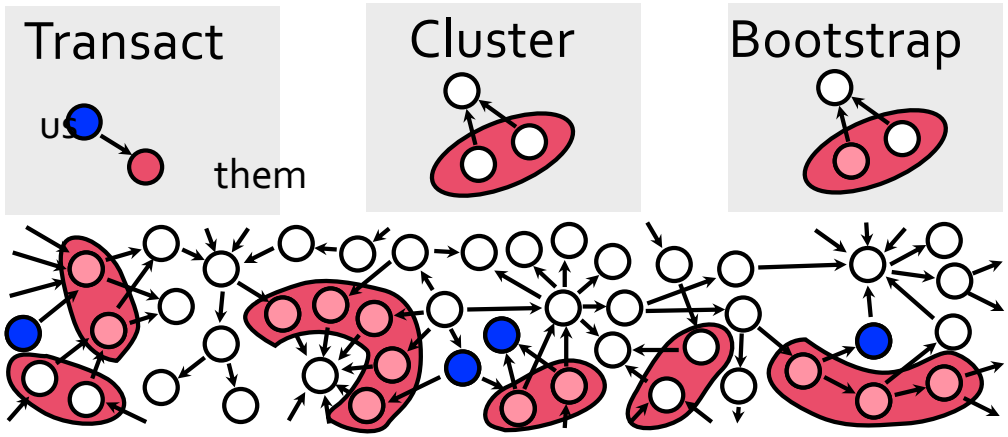
- Its change iff:
  - this is not a block reward transaction (i.e., fresh coin)
  - there is no self-change address (dumbass rule, same address in input and output)
  - it is the only **fresh** address on the output side (never used before)

Heuristic relies on idiom of Bitcoin wallet code/services of the day

- Would generate and manage a single new addresses for change
  - Creates a peel chain of change addresses that get used (input side) as needed in future transactions (as per previous picture)
  - Each change address gets used precisely twice (once in, once out)

But… clustering still requires that you actually name things… how much purchasing is required?

Empirical finding: don't need huge number of "seeds" Huge expansion from graph clustering



Interacted with 31 MtGox addresses, tagged 518,723

Participated in 344 transactions and tagged 1.3M public keys

Can it trace criminal Bitcoin flows?

i.e., Does illicit money tend to hit services with KYC obligations?

Theft	BTC	Date	Movement	Exchanges?
MyBitcoin	4019	Jun 2011	A/P/S	Yes
Linode	46,648	Mar 2012	A/P/F	Yes
Betcoin	3171	Mar 2012	F/A/P	Yes
Bitcoinica	18,547	May 2012	P/A	Yes
Bitcoinica	40,000	Jul 2012	P/A/S	Yes
Bitfloor	24,078	Sep 2012	P/A/P	Yes
Trojan	3257	Oct 2012	F/A	No

2857 BTC (87%) hadn't moved

Other change address heuristics

A bunch of variants

- Not a transaction with a mix (next class) (Goldfeder)
- Value must be significant to the 4^{th} decimal place (filter out noise transactions, Satoshi dice, fees, etc)

How to Peel A Million heuristic (2022 Meiklejohn paper)

- Classifies each transaction by
  - Vector of four binary features (replace-by-fee, locktime, version, segwit)
  - 10 different kinds of address feature combinations (pubkey, hash, witness, compressed vs uncompressed, multisigs of various sizes, segwit, etc)
  - Location of candidate change address in output list (always first, always last, always first or last, other)
- Features used to train ransom forest classifier (manual ground truth for TP/FP lagels)

Basic idea here: there are now a broad range of wallet implementations

- a user will use a particular kind of wallet service/software and all of its transactions will follow the same/similar idiom

Comparing against ground truth today

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

Table 4: The expansion factor and false discovery rate of findNext and findNext2, as evaluated on our 60 TP clusters and as compared with previous change heuristics. Both metrics are averaged across all clusters.

Figure 1: An example of splitting. This address received 22 Bitcoin from the US-based Gemini exchange, and split into 25% and 75%. 1 Bitcoin from this address would eventually be sent to an address in the leak. Other funds were transferred to other illicit entities, such as the sanctioned exchange Garantex.

Exchange	Confirmed Payments	Likely Payments	Total
Unlabeled Cluster	$8.6M	$64.8M	$73.4M
Gemini	$5.9M	$17.4M	$23.1M
Kraken	$1.0M	$0.2M	$1.2M
Coinbase	$0.4M	$0.6M	$1.1M
Binance	$0.6M	$0.0M	$0.6M

Table IV: Top exchanges from which Conti ransom payments originate. Note that "Unlabeled Cluster" represents the unlabaled cluster of bitcoin addresses, discussed in Section IV.

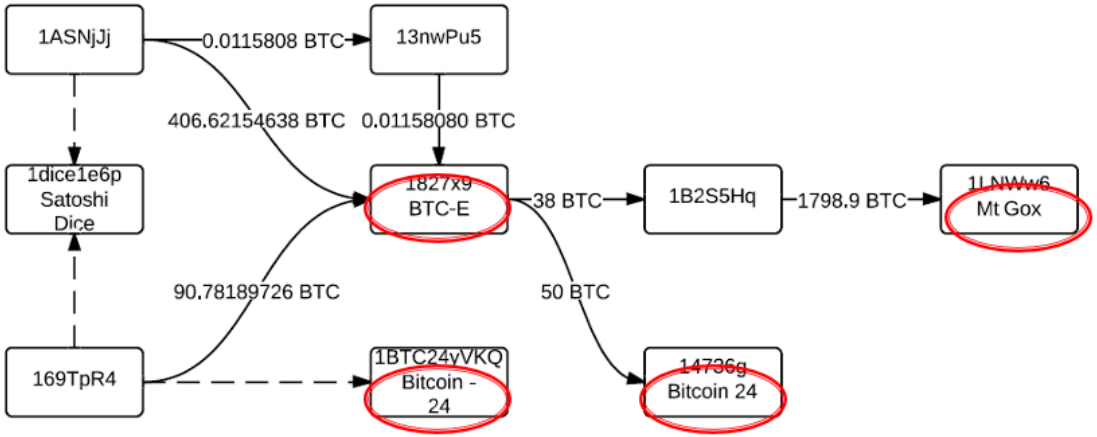


Figure 1: CryptoHitman Outgoing Relationships Graph

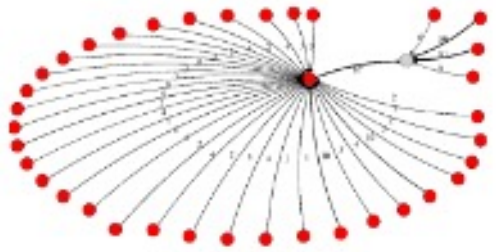


Figure 2: Locky collector address example.

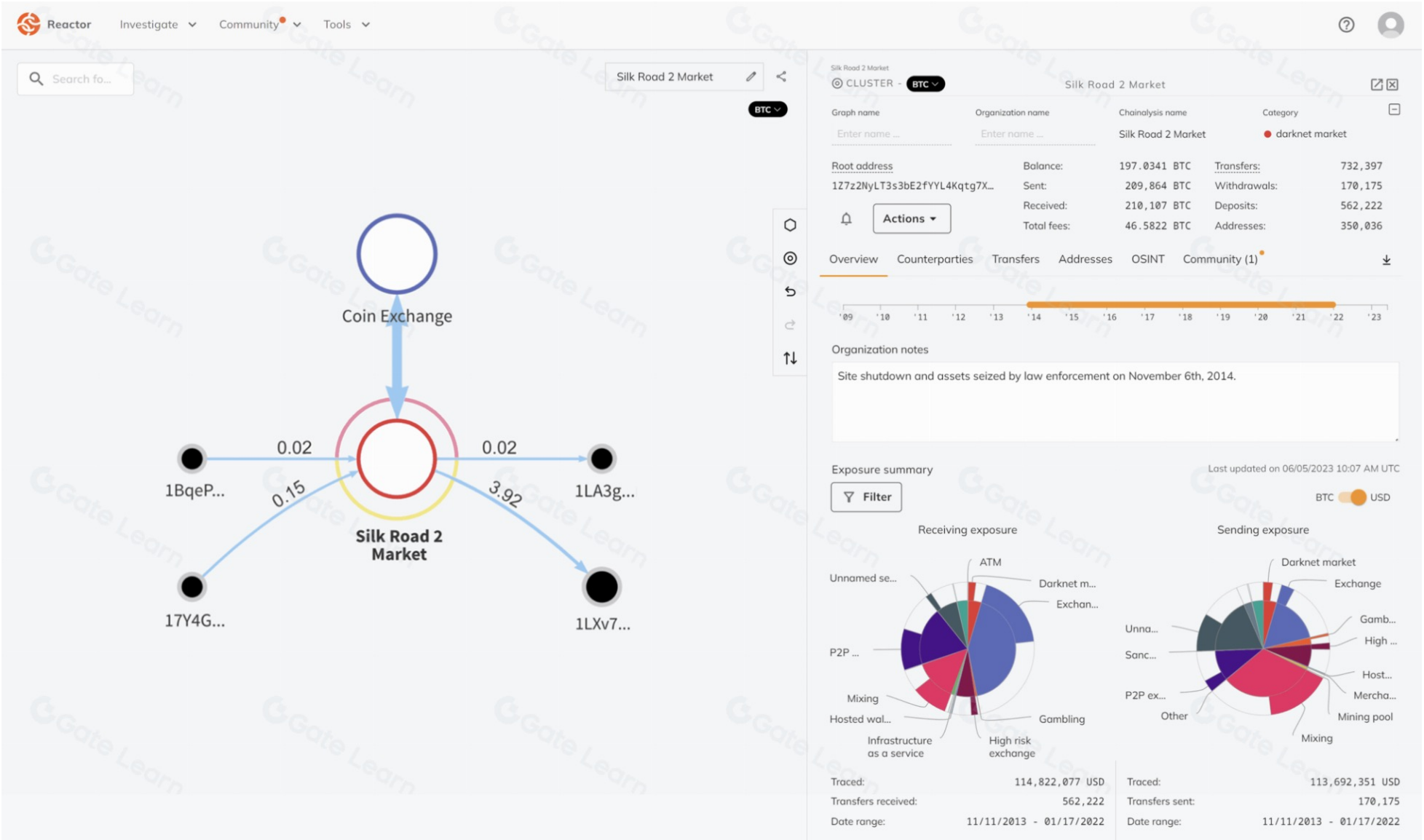
Today a bunch of companies compete in this space

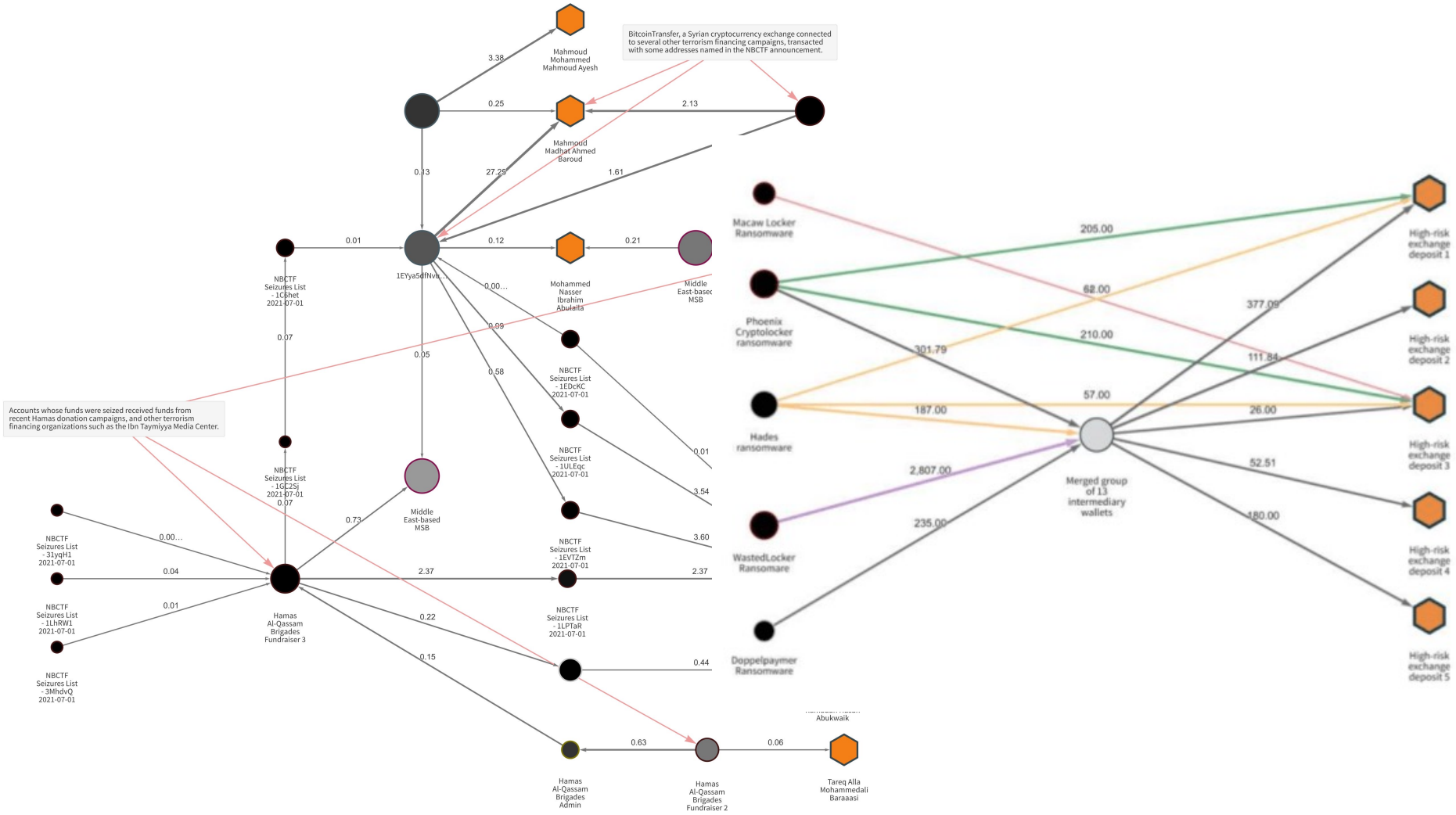
More advanced heuristics

- Also using external data

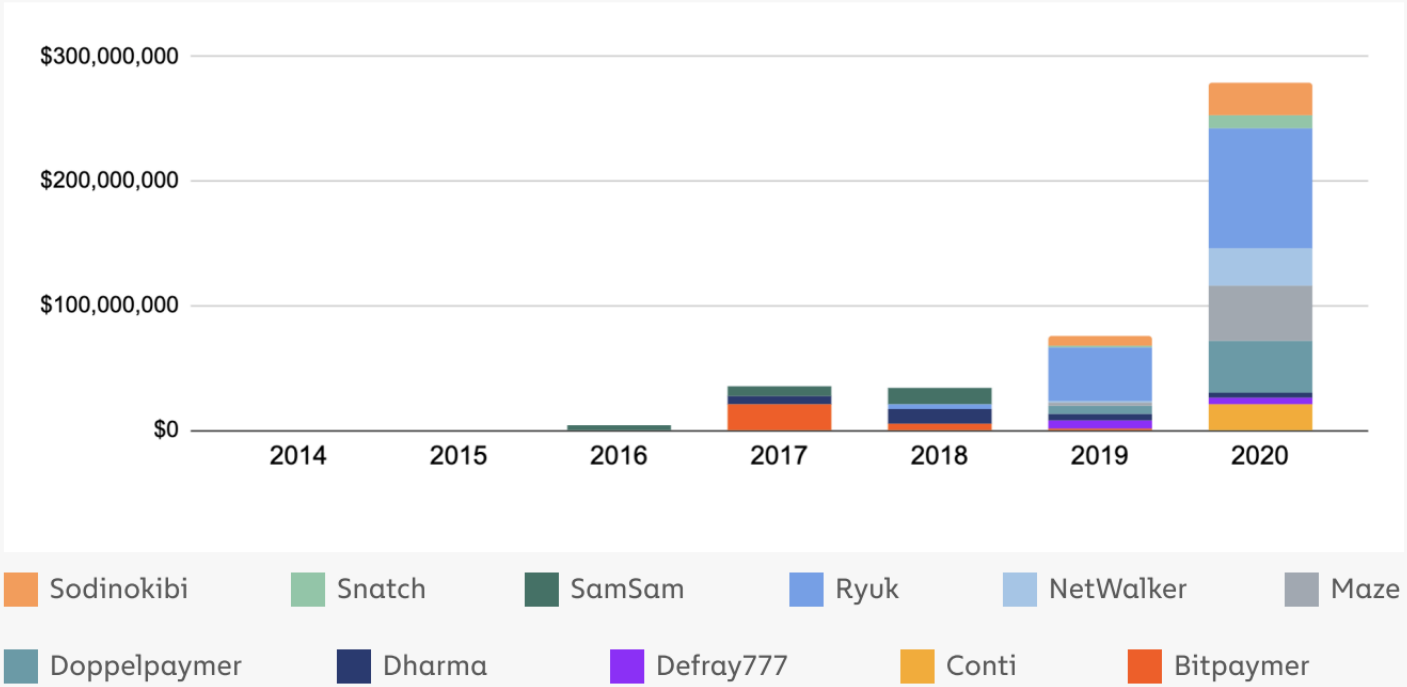
Lots of effort put into labelling (shopping and equivalent)

High-performance and interactive UIs


Reactor
Investigate
Community
Tools
Search fo...
Silk Road 2 Market
BTC
Coin Exchange
0.02
0.02
0.15
3.92
1BqeP...
1LA3g...
17Y4G...
1LXv7...
Silk Road 2 Market
Silk Road 2 Market
CLUSTER - BTC
Silk Road 2 Market
Graph name
Enter name ...
Organization name
Enter name ...
Chainalysis name
Silk Road 2 Market
Category
darknet market
Root address
1Z7z2NyLT3s3bE2fYYL4Kqtg7X...
Balance: 197.0341 BTC
Sent: 209,864 BTC
Received: 210,107 BTC
Total fees: 46.5822 BTC
Transfers: 732,397
Withdrawals: 170,175
Deposits: 562,222
Addresses: 350,036
Actions
Overview
Counterparties
Transfers
Addresses
OSINT
Community (1)
'09 '10 '11 '12 '13 '14 '15 '16 '17 '18 '19 '20 '21 '22 '23
Organization notes
Site shutdown and assets seized by law enforcement on November 6th, 2014.
Exposure summary
Last updated on 06/05/2023 10:07 AM UTC
Filter
BTC USD
Receiving exposure
Sending exposure
ATM
Darknet m...
Exchan...
Unnamed se...
P2P ...
Mixing
Hosted wal...
Infrastructure as a service
High risk exchange
Gambling
Darknet market
Exchange
Gamb...
High ...
Host...
Mercha...
Mining pool
Mixing
P2P ex...
Other
Sanc...
Unna...
Traced: 114,822,077 USD
Transfers received: 562,222
Date range: 11/11/2013 - 01/17/2022
Traced: 113,692,351 USD
Transfers sent: 170,175
Date range: 11/11/2013 - 01/17/2022


BitcoinTransfer, a Syrian cryptocurrency exchange connected to several other terrorism financing campaigns, transacted with some addresses named in the NBCTF announcement.
Accounts whose funds were seized received funds from recent Hamas donation campaigns, and other terrorism financing organizations such as the Ibn Taymiyya Media Center.
Mahmoud Mohammed Mahmoud Ayesh
Mahmoud Madhat Ahmed Baroud
Mohammed Nasser Ibrahim Abulaffa
Middle East-based MSB
NBCTF Seizures List - 1CGhet 2021-07-01
1EYya5dfNvb...
NBCTF Seizures List - 1EDcKC 2021-07-01
NBCTF Seizures List - 1ULEqc 2021-07-01
NBCTF Seizures List - 1GC2Sj 2021-07-01
NBCTF Seizures List - 1EVTZm 2021-07-01
NBCTF Seizures List - 1LPTaR 2021-07-01
NBCTF Seizures List - 31yqH1 2021-07-01
NBCTF Seizures List - 1LhRW1 2021-07-01
NBCTF Seizures List - 3MhdvQ 2021-07-01
Hamas Al-Qassam Brigades Fundraiser 3
Hamas Al-Qassam Brigades Admin
Hamas Al-Qassam Brigades Fundraiser 2
Tareq Alla Mohammedali Baraasi
Abukwaik
Macaw Locker Ransomware
Phoenix Cryptolocker ransomware
Hades ransomware
WastedLocker Ransomare
Doppelpaymer Ransomware
Merged group of 13 intermediary wallets
High-risk exchange deposit 1
High-risk exchange deposit 2
High-risk exchange deposit 3
High-risk exchange deposit 4
High-risk exchange deposit 5

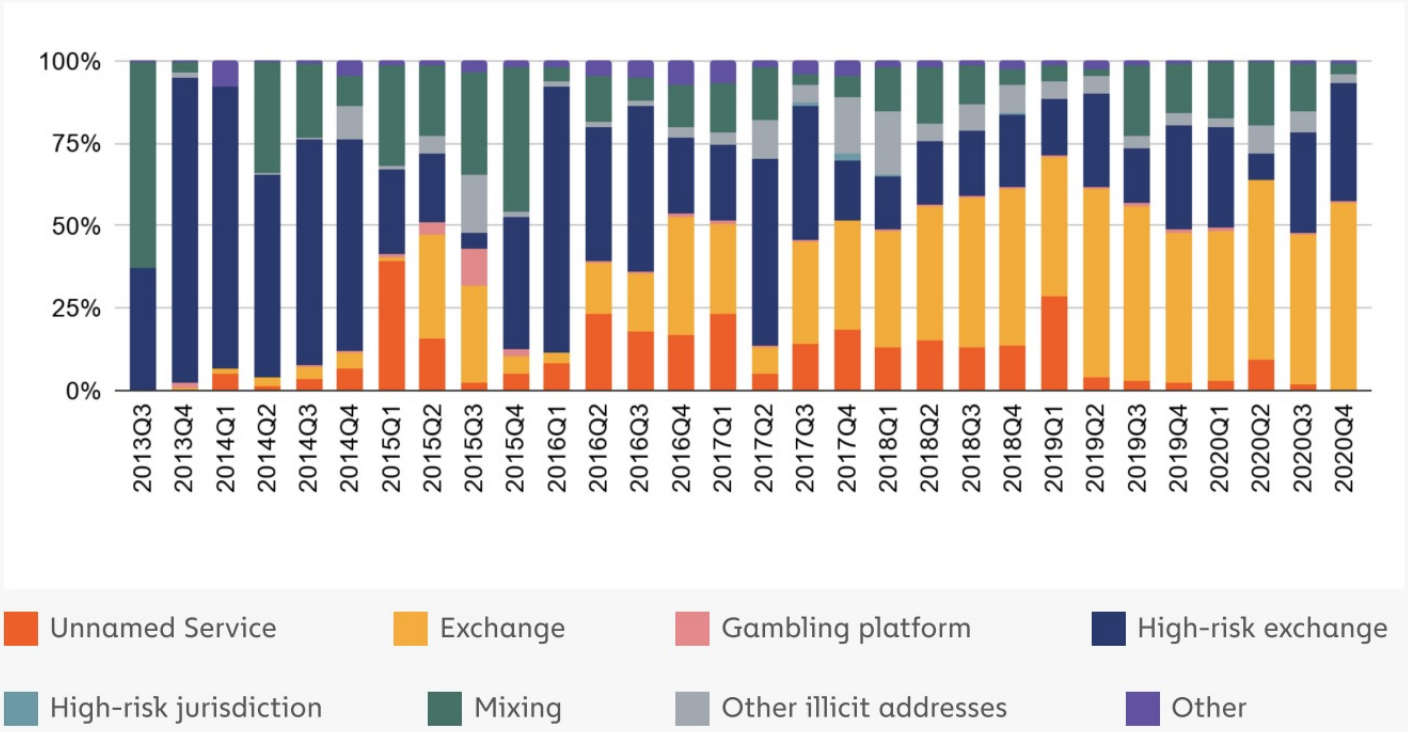
Top 10 ransomware strains by revenue by year | 2014 - 2020



Currencies included: BCH, BTC

Ransomware strains don't operate consistently, even month-to-month. Below, we see that the top-earning strains have ebbed and flowed throughout 2020.

Destination of funds leaving ransomware wallets | Q3 2013 - Q4 2020



Can you use this in court? It's a heuristic after all

Discovered chains are typically very clear
Typically evidence discovered at exchanges are crisp and damning
(i.e., result doesn't rest on clustering itself)

But are warrant searches legit?
(i.e., does clustering meet probable cause standard of 4th amendment)

Ok, but what about 4th amendment?

Can government use blockchain clustering analysis to establish probable cause? (DC Circuit opinion)

UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

IN THE MATTER OF THE SEARCH OF MULTIPLE EMAIL ACCOUNTS PURSUANT TO 18 U.S.C. § 2703 FOR INVESTIGATION OF VIOLATION OF 18 U.S.C. § 1956 et al

Case No. 20-sc-3310 (ZMF)

MEMORANDUM OPINION

In January 2021, the government submitted an Application for a Warrant ("Application") to search certain email accounts (the "Target Accounts"). *See* ECF No. 3 (Application). The Court subsequently posed questions to the government about this request. In June 2021, the government submitted a memorandum of law in support of the Application. *See* ECF No. 8 (Mem. in Supp. Of Appl.) ("Memo"). The Court's concerns included whether: (1) it had venue; (2) the government's previous collection of evidence complied with the Fourth Amendment; and (3) the software the government used to establish probable cause was reliable. For the reasons below, this Court granted the Application.[1]

D. Blockchain Analysis

Cryptocurrency transactions that occur on a blockchain are, by design, publicly available, and thus are pseudoanonymous. *See* Sarah Meiklejohn et al., *A Fistful of Bitcoins: Characterizing Payments Among Men with No Names*, IMC '13: Proceedings of the 2013 Conference on Internet Measurement Conference, Barcelona, Oct. 23-25, 2013, at 1, 1 (Association for Computing Machinery), https://doi.org/10.1145/2504730.2504747 [hereinafter *Fistful*]; *One Address*, *supra*, at *4. "Ironically, the public nature of the blockchain makes it exponentially easier to follow the flow of cryptocurrency over fiat funds." *One Address*, *supra*, at *3. Repeated government seizures and forfeiture actions should disabuse the uninformed of the myth that BTC is untraceable, yet this myth abides. Indeed, the IRS alone seized $1.2 billion worth of cryptocurrency in fiscal year 2021. *See The IRS has seized $1.2 billion worth of cryptocurrency this fiscal year – here's what happens to it*, https://www.cnbc.com/2021/08/04/irs-has-seized-1point2-billion-worth-of-cryptocurrency-this-year-.html.

B. Probable Cause Determination

i. Clustering as the basis for probable cause

The instant affidavit intricately follows the theft of funds using subpoena and prior search warrant returns. *See* Aff. Yet underlying this all is the clustering analysis, which empowered the government to defeat a variety of alleged money laundering techniques. For example, "investigators have been able to trace the stolen funds moving from Victim VCE to a cluster of BTC addresses, where they remained dormant until January 2017. Then, after the stolen funds began to move again, investigators traced them as follows:

ii. Reliability of clustering software

Probable cause "is a fluid concept—turning on the assessment of probabilities in particular factual contexts—not readily, or even usefully, reduced to a neat set of legal rules." *Maryland v. Pringle*, 540 U.S. 366, 370–71 (2003) (citation omitted). The probable-cause standard is "incapable of precise definition or quantification into percentages because it deals with probabilities and depends on the totality of the circumstances." *Id.* at 371. "[W]hen the majority of the information in the affidavit comes from confidential sources, as it does in this case, courts must consider the veracity, reliability, and the basis of knowledge for that information as part of the totality of circumstances." *United States v. Dyer*, 580 F.3d 386, 390 (6th Cir. 2009) (cleaned up). "While independent corroboration of a confidential informant's story is not a *sine qua non* to a finding of probable cause, in the absence of any indicia of the informants' reliability, courts insist that the affidavit contain substantial independent police corroboration." *Id.* at 390–91.

The sources of information here are the blockchain and the clustering software. "[T]here are no published decisions analyzing the weight or reliability of blockchain evidence in a search warrant application." C. Alden Pelker et. al., *Using Blockchain Analysis from Investigation to Trial*, 69 DOJ J. Fed. L. & Prac. 59, 68 (2021). Not until now. "There is no serious question that the blockchain accurately captures the transactional data used in blockchain analysis. In a similar vein, the blockchain is the product of an automated process (for example, the Bitcoin protocol), so it makes little sense for a court to question the veracity of the data the way it might inquire into the

It is human nature to assess technological confidential sources with greater skepticism. Yet humans are "Flawed. Weak. Organic," Star Trek First Contact (Paramount Pictures, 1996) (Borg Queen), whereas clustering software strives for perfection. To address concerns about human confidential sources' reliability, courts look to prior success. Courts want multiple prior tips, for which the source was "truthful and reliable," and that yielded evidence of the crime and/or arrests. *United States v. Brundidge*, 170 F.3d 1350, 1353 (11th Cir. 1999). There is no hard baseline for these categories. In one case, a source who provided information at least eight prior times, which information was truthful and reliable, and such tips led to the arrest of five persons and the recovery of $3,500 in illegal drugs was reliable for future tips. *See id.* For another court, "[f]ive or six tips

supra at 69–70. Yet "[b]lockchain analysis software does not only aggregate blockchain data; it also applies heuristics and other analytical tools to cluster addresses into related groups." *Id.* at 70. Still, there is no evidence in the government's affidavit that such software reports "false or misleading information," or that it was unreliable. *U.S. v. Thomas*, No. 5:12-cr-37, 2013 WL 6000484, at *6 (D. Vt. Nov. 8, 2013); *see* ECF 3 at 21. Far from it, the government's data reveals only overwhelming reliability of this software. *See* ECF 3 at 21. "Because probable cause does not require scientific certainty, no more was [required]." *United States v. Chiaradio*, 684 F.3d 265,

IV. CONCLUSION

Cryptocurrency and related software analytics tools are "[t]he wave of the future, Dude. One hundred percent electronic." The Big Lebowski (Polygram Filmed Entertainment & Working Title Films 1998).

To summarize

Tracing crypto payments are about identifying clear flows to/from entities that can tie you to an identity (e.g., via KYC obligations)

- Blockchain itself identifies all on-chain transactions (A sends to B)
- Leverage shared control to cluster things like deposit/aggregation addresses

UTXO-style chains are a bit trickier to trace

- Key issue is the algorithm for identifying the change address
- If combined with existing input address is easy
- Otherwise, leverage commonality of transaction type and address type

Ok, so are criminals screwed?

- Dirty exchanges (e.g., BTC-e)
- Blockchains designed for anonymity (Zcash, Monero)
- Mixes/Tumblers
- Next time…