On the Dark Side of the Coin: Characterizing Bitcoin use for Illicit Activities

Hampus Rosenquist[1], David Hasselquist[1], Martin Arlitt[2], and Niklas Carlsson[1]

[1] Linköping University, Linköping, Sweden
[2] University of Calgary, Calgary, Canada

Abstract. The decentralized nature of Bitcoin allows for pseudonymous money exchange beyond authorities' control, contributing to its popularity for diverse illegal activities such as scams, ransomware attacks, money laundering, and black markets. In this paper, we characterize this landscape, providing insights into similarities and differences in the use of Bitcoin for such activities. Our analysis and the derived insights contribute to the understanding of Bitcoin transactions associated with illegal activities through three main aspects. First, our study offers a comprehensive characterization of money flows to and from Bitcoin addresses linked to different abuse categories, revealing variations in flow patterns and success rates. Second, our temporal analysis captures long-term trends and weekly patterns across categories. Finally, our analysis of outflow from reported addresses uncovers differences in graph properties and flow patterns among illicit addresses and between abuse categories. These findings provide valuable insights into the distribution, temporal dynamics, and interconnections within various categories of Bitcoin transactions related to illicit activities. The increased understanding of this landscape and the insights gained from this study offer important empirical guidance for informed decision-making and policy development in the ongoing effort to address the challenges presented by illicit activities within the cryptocurrency space.

1 Introduction

Bitcoin, a decentralized digital currency, is attempting to revolutionize finance by facilitating pseudonymous exchanges outside the oversight of traditional authorities. Bitcoin's unique pseudonymous design, coupled with its lack of regulation, has propelled its widespread popularity. However, this has also made it a favored tool for illicit activities such as scams, ransomware attacks, money laundering, and black market transactions. As a result, Bitcoin poses significant challenges to law enforcement agencies globally, straining traditional legal frameworks.

Furthermore, as our paper demonstrates, Bitcoin use for illicit activities is widespread and turns over large sums of money. With many of the abuse types studied here preying on the weak, it is clear that these activities have increasingly negative societal effects. While the public often focuses on Bitcoin's energy consumption, much less attention has been placed on Bitcoin's role in various illicit activities affecting large numbers of victims. With the effect of Bitcoin abuse on humans being both *apparent* and *current*, we argue that it is important to shine a light on the Bitcoin patterns associated with different illicit activities.

Despite prior works having considered a wide range of criminal activities with relations to Bitcoin [22,28,27,41,10], most prior works either try to estimate the global cybercrime Bitcoin revenue [32,14,19] or focus only on a single category of illegal activities, including money laundering (using tumblers) [23], ransomware [35,29,37,17,42,13,9,20], sextortion [26,25], cryptojacking [38], darknet markets [8,7,18], and human trafficking [30]. In contrast to these works, we present a comprehensive characterization of the money-flow to and from a large set of addresses linked to *various categories of illegal activities,* and provide insights into both similarities and differences in the Bitcoin transactions and Bitcoin flows associated with the addresses of the different categories.

This paper uses the Bitcoin Abuse Database and Bitcoin's blockchain as its primary data sources for analysis. The Bitcoin Abuse Database [4] provides information on attacks and related Bitcoin addresses from reports submitted by victims and other individuals or organizations, detailing the attack type and often including additional information such as email examples. By identifying Bitcoin addresses used by attackers and extracting information about these addresses (and addresses they send money to) directly from the Bitcoin blockchain, we provide a comprehensive comparison of the quantity of funds directed to addresses associated with different types of attacks. Importantly, this methodology enables the observation of transactions involving a larger number of victims beyond those who reported an attack, acknowledging that many actual victims may not report their experiences and that some reports may come from individuals who did not transfer funds themselves. We next outline our main contributions.

First, we perform a high-level characterization of the transactions received by the Bitcoin addresses reported to the Bitcoin Abuse Database [4] from May 16, 2017 to April 25, 2022; both as an aggregate across all reported addresses (Section 3) and on a per-category basis (Section 4). Our characterization reveals a high skew in the distribution of funds attracted by different Bitcoin addresses involved in illicit activities. Our observations also highlight significant variations in the success of different abuse categories, with "Blackmail scams" and "Sextortion" receiving numerous reports but attracting smaller funds, while categories like "Ransomware" and "Darknet markets" receive fewer reports but attract substantial funds, indicating differences in effectiveness and financial impact. The category attracting the most transactions and funds, however, is the "Other" category. While only the fourth-most reported category, this category includes many of the top addresses attracting the most and the biggest transactions.

Second, we perform a temporal analysis (Section 5) that captures both long-term trends, differences in the weekly patterns associated with the different categories, and temporal correlations with when reports of an illicit address are first reported. While the number of reports in the Bitcoin Abuse Database remained relatively steady between 2019 and 2022, the daily number of bitcoins received by reported addresses increased by a factor of 100 over the same period, indicating a substantial rise in funds transferred to these addresses. Weekly variations were observed, with higher volumes and more funds transferred during weekdays compared to weekends, with notable patterns in “Ransomware”, “Darknet markets”, and the “Other” category. Although the reports typically were obtained around the time that the addresses saw peak activity and there were significant variations between abuse categories, most transactions occurred before the first report was filed, suggesting that victims may not report abusive addresses. This raises questions about the effectiveness of using reports to ban addresses.

Third, we analyze the outflow of bitcoins from the reported addresses associated with each abuse category (Section 6). This analysis reveals several additional interesting observations. For example, when considering the outgoing money from reported addresses, there is a concentration of funds towards specific addresses, while the majority of receiving addresses have a node degree of one, indicating a dispersion of funds after the first step. In our multi-step tracking of money flows, Bitcoin tumblers stand out with higher node degrees and concentration, suggesting relatively fewer actors are involved in this category or that many Bitcoin-using miscreants use tumblers. There are significant differences in graph structure and transaction patterns between categories, with Bitcoin tumblers having more connecting edges and loops, and the “Other” category receiving significantly more transactions going back to reported addresses. Finally, transactions between categories show increased inflow/outflow to/from Bitcoin tumblers, indicating interest in their money laundering services.

Summary of contributions: We present a comprehensive characterization of money flows to and from a large set of Bitcoin addresses associated with different categories of illegal activities. Our analysis and the derived insights contribute significantly to the understanding of Bitcoin transactions associated with illegal activities through three main aspects. First, our aggregate and category-based characterizations reveal variations in flow patterns and success rates both within and across addresses associated with the different categories. Second, our temporal analysis captures long-term trends, weekly patterns, and the relative timing of when illicit addresses of each category are first reported. Finally, our analysis of the outflow from reported addresses uncovers differences in graph properties and flow patterns among illicit addresses and between abuse categories. Overall, the paper provides comprehensive insights into the money-flow characteristics in Bitcoin transactions linked to various illegal activities, revealing differences and similarities in distributions, temporal patterns, and interconnections among different categories.

Outline: After presenting our data collection methodology and dataset (§2), we first present a brief aggregate characterization (§3), followed by a category-based characterization (§4), a temporal analysis (§5), and an outflow analysis (§6). Finally, we present related work (§7) and conclusions (§8).

Ethics: Our research discusses the significant challenges presented by Bitcoin, including its pseudonymous transactions and lack of regulation, particularly in facilitating illicit activities. It emphasizes that Bitcoin’s use for illegal purposes is widespread and has negative societal impacts, often harming vulnerable individuals. Our study respects privacy and confidentiality by using data from

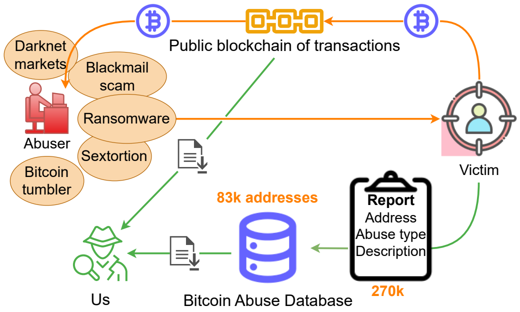


Fig. 1. Our data collection framework in the context of Bitcoin Abuse reports (filled by victims and other entities) and the public blockchains storing information about all transactions to/from the attacker operated addresses.

the Bitcoin Abuse Database while adhering to legal and ethical guidelines, and our insights can be used by law enforcement, regulatory bodies, and cryptocurrency service providers to develop strategies for safer and more secure financial practices. We discuss ethical concerns in more detail in Appendix A.

2 Data Collection Methodology

We rely on two primary data sources in this paper: the Bitcoin Abuse Database and Bitcoin's blockchain. Figure 1 presents an overview of our data collection framework in the context of these two information sources.

High-level overview: First, we use the "abuse reports" collected by the Bitcoin Abuse Database to obtain knowledge about attacks and the Bitcoin addresses that the attackers used in these attacks. These reports are typically submitted by victims and other persons/organizations and contain information about what type of attack was performed (e.g., blackmail scam, ransomware, sextortion, etc.) and typically some additional information about the attack (e.g., an email example) and the Bitcoin address(es) used in the attack.

Second, we use a series of tools to extract various information about the identified Bitcoin addresses that the attackers used directly from the Bitcoin blockchain itself. Using this information, we compare and contrast how successful attackers were in attracting funds from potential victims to the addresses associated with different types of attacks.

The above methodology allows us to observe transactions made by victims beyond those who report an attack and accounts for reports sometimes being filled by people who did not fall victim themselves. This is important since many victims never report that they have been attacked.

Data sharing: To ensure reproducibility and help others continue this line of research, the combined dataset will be shared with other researchers as per steps outlined here: `https://www.ida.liu.se/~nikca89/papers/pam24a.html`.

Table 1. Summary of primary dataset.

Time period of reports	**2017-05-16 to 2022-04-25**
Reports	267,708
Unique addresses	82,527
Transactions	5,092,489
Received bitcoins	31,346,586
Approximate value in USD	815,011,236,000

2.1 Bitcoin Abuse Database

The Bitcoin Abuse Database [4] contained reports dating back all the way to 2017-05-16; with new reports still being added in the summer of 2023. (The website `bitcoinabuse.com` is now merged/integrated with `chainabuse.com`.) For our main dataset, we obtained all records reported to this database between 2017-05-16 and 2022-04-25. The first two rows in Table 1 summarizes the number of reports (267,708) and unique Bitcoin addresses (82,527) in this dataset.

In addition to our main dataset, we also collected data and analyzed the records reported between 2022-12-20 and 2023-05-19. A brief discussion of these results are presented in Appendix B. The reason for the gap in the dataset were an API issue (appearing in 2022) that only allowed access to recent reports combined with a gap in our data collection.

Dataset information: The Bitcoin Abuse dataset includes the following fields: id, address, abuse type id, abuse type other, abuser, description, from country, from country code, and created at. "id" is a unique number assigned to each report by the database. "Address" is the Bitcoin address that is reported for abuse. "Abuse type id" is a number representing the abuse type (category): Ransomware (1), Darknet markets (2), Bitcoin tumbler (3), Blackmail scam (4), Sextortion (5), and Other (99). "Abuse type other" is an optional free text column where the reporter may describe the abuse type whenever choosing "Other". A closer look at the free text classification of the top-1K accounts from the "Other" category (in terms of funds transferred) reveal that many reporters selected to list terms such as "scam", "investment scam", "ponzi", or use words such as terror, fraud/fake/phishing, hacker/attacker, or stolen/theft in their free text answers. "Abuser" is a free text field where the reporter may describe the abuser's identity. "Description" is a free text field where the reporter usually describes the abuse in more detail; many simply paste an email they have received from a perpetrator. Here, we have observed high similarity between some reports (e.g., copied text), suggesting they may be submitted by the same person, but also quite specific descriptions that appears to be provided by first-hand victims. "From country" represents the reporter's (victim's) home country. "From country code" is the reporter's home country code. "Created at" is a date and time field representing the time the *report* was made.

Limitations: First, as noted above, an API issue and a gap in our data collection prevented us from obtaining data for the time period 2022-04-26 to 2022-12-19. Due to this gap, we decided to focus our analysis on our main dataset: 2017-05-16 to 2022-04-25. Second, `bitcoinabuse.com` has merged with chainabuse.com and is no longer making its full database freely available, making it difficult to fully extend the analysis into 2023 and beyond. While the new site provides an API, this API has a restrictive, rate-limited interface and uses a different categorization of the reported addresses. For these reasons, and after some exploration with the new API, we found the analysis presented here (of our complete main dataset) more comprehensive and insightful than we at this time can achieve with alternative or augmented datasets using data from the new API. Third, we only consider reported Bitcoin addresses. While other cryptocurrencies may also be used for their illicit activities, Bitcoin is still the dominating currency for such activities. For example, chainabuse.com has received close to nine times as many reports for Bitcoin as for the second-most reported cryptocurrency (Ethereum). Fourth, Bitcoin Abuse has a limited category selection, with the “Other" category being the fourth reported category. While it would be interesting to see a finer split of the “Other" category, we felt that the quality of the free-text answers were not sufficient to provide an accurate sub-classification here, and note that only two categories “Darknet markets" and “Bitcoin tumbler" saw fewer reports (cf. Figure 4). Finally, we acknowledge that Bitcoin Abuse only collects reports in English, which potentially causes some biases in the reported addresses and note that it is difficult to know to what degree the full set of *reported* addresses accurately represent the addresses *most used* for illicit activities.

2.2 Blockchain information

We tested several APIs to retrieve information about each observed address from the blockchain. For the analysis presented here, we used the Blockchain.com API [5]. The data for each address was saved to two files each. One includes the raw JSON form retrieved from the API and one contains a list of the address' transactions in CSV format with the headers: hash, timestamp, received/sent, address, and value. The list was created to summarize the data points of interest in an easy-to-process format for later analysis.

2.3 Pre-processing and summary files

To simplify our data analysis, we created two summary files: one with summary information about each reported address and one with information about each transaction associated with these addresses. Both these files contained summary information based on (1) all reports from the Bitcoin Abuse Database containing the address, (2) the raw JSON files from the blockchain API that were associated with the address, and (3) the list of transactions (and their properties).

Per-address summary file: For each address, this file contains the following information (additional fields): received BTC, sent BTC, balance in BTC, # of received transactions, # sent transactions, # total transactions, average received BTC/transaction, average sent BTC/transaction, median received BTC/transaction, median sent BTC/transaction, date of last transaction, most common abuse type id, abuse type ids, abuse type freetext, # reports, date of the first report, country that the address was most commonly reported in, abuser identity, abuse description. Here, we note that the field "most common abuse type id" captures only the most common abuse type (i.e., the category selected by the reporter). To capture the full set of abuse types that an address has seen reports associated with, we included the field "abuse type ids", which contains a list of all different abuse types that the address has been labeled with. The same applies to the fields for in which country the abuse was reported in.

Per-transaction summary file: This file contains all transactions that were made, with some additional information about the involved address (from the per-address summary file) about the abuse type (category) that the transaction's address belongs to and the date that the transaction's address was first reported. The complete list of headers for the transaction file were: the hash, timestamp, value, address1 (the reported address), received/sent, address2 (other sender/receiver), most common abuse type id, first reported.

2.4 Dataset summary

Table 1 presents an overview of the dataset for our analysis (based on reports made May 16, 2017, to April 25, 2022). We also use an extra dataset (based on reports made the last five months) to confirm that we observe similar reporting rates (114 vs. 147 reports/day) and transactions per address (62.9 vs. 61.7 transactions/address) as seen in the past few years. The statistics for this dataset are provided in Appendix B.

Focusing on our main dataset, based on 268K reports, we identified 83K unique Bitcoin addresses that together have received 31M bitcoins across 5M transactions. Using the average price of a bitcoin (BTC) from 2023 (estimated at roughly $26K USD/BTC) [16], this amounts to a staggering $815 billion USD in total transaction value, and using the daily closing prices (given by Yahoo finance data) the day of each transaction, this amounts to $687 billion USD.

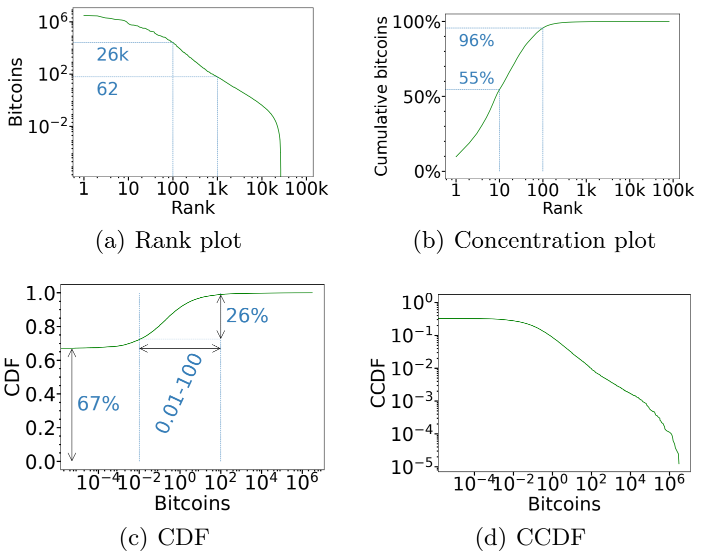
At this time it should also be noted that the value of the cumulative transferred funds should not be seen as a revenue estimate. As we show later in the paper, funds are often moved between several accounts (possibly owned by the same entity), complicating revenue estimates as well as the identification of the initial money transfers made by victims.

3 Aggregate High-level Characterization

This section presents an aggregate analysis of how successful each address is (§3.1) and examines how to best model the amount of BTC obtained by the set of addresses (§3.2).

3.1 How successful is each address?

In our Bitcoin Abuse dataset, 83K unique Bitcoin addresses were reported, and as one may expect, their success in (illicitly) attracting funds varied.



(a) Rank plot

(b) Concentration plot

(c) CDF

(d) CCDF

Fig. 2. Distribution statistics. Rank plots of (a) the received BTC per address and (b) the cumulative fraction of received BTC. The last two sub-plots show (c) the CDF and (d) CCDF of received BTC per address.

High skew: We observed a significant skew in the distribution of funds, with a relatively small subset of addresses attracting the bulk of the attracted funds. This skew is characterized and quantified in Figures 2(a) and 2(b). Figure 2(a) shows the total bitcoins received per address (reported to bitcoinabuse.com) as a function of the rank of each address. Figure 2(b) shows the cumulative fraction of the total observed bitcoins that the top-X addresses have attracted as a function of the cumulative rank X. The top-10 addresses each received more than 700K bitcoins; together these ten addresses were responsible for 55% of the total bitcoins received across all 83K addresses (i.e., 17M out of 31M bitcoins). Similarly, the top-100 all have each received more than 26K bitcoins, combining for more than 29M bitcoins or 96% of the total observed bitcoins in the dataset. The top-1K have each received more than 62 bitcoins (together being responsible for 99.8% of the total bitcoins observed). While 62 bitcoins may seem small relative to the most successful addresses, we note that this still suggests that there are more than 1K abusive addresses that may have attracted over $1.6M USD (based on the $26K USD/BTC estimate of the average Bitcoin price during 2023) and that these 1K addresses together have attracted an estimated $814B USD. For a more granular estimation, using the daily transaction values, we refer to Section 4.1. Sorting the top-1K addresses based on the daily price of Bitcoins for transaction value estimation reveals that all of them have received transactions exceeding $2.8M USD.

Big hitters: Table 2 provides an overview of the number of bitcoins that each of the top-10 accounts have received and the type of reports that have been filed against these accounts. Perhaps most noteworthy, the address that received

Table 2. Overview of the top-10 highest receiving reported addresses.

Received [BTC]	Median [BTC]	Category	Description
3,048,040	40.0	Other	Trading investment scam.
2,845,086	18.0	Other	Forex trading scam, "investment in terror".
2,009,608	25.0	Other	"Investment in terror", begs for treatment money.
1,815,619	800	Other	"Investment in terror".
1,535,341	45.0	Other	"Investment in terror".
1,459,182	160	Ransomware	"Investment in terror".
1,378,975	800	Other	"Investment in terror".
1,259,824	0.50	Other	"Investment in terror".
1,030,376	505	Other	"Inhumane" bank account theft via remote desktop.
724,340	1,150	Other	"Investment in terror", begs for treatment money.

the most bitcoins during our study received more than 3M bitcoins (worth $79B USD in 2023). While it is difficult to convert Bitcoin to usable funds (e.g., without impacting its value) and some of these funds are likely being double-counted (e.g., due to use of Tumblers), this staggering amount is of the same order of magnitude as the Gross Domestic Product (GDP) of entire US states such as Maine and North Dakota [39] or European countries such as Luxembourg [15]. The reports associated with this attack list it as being associated with trading investment scams and link it to services such as CapitalBullTrade [36]. The second-ranked address is reported to be associated with a foreign exchange trading scam (ROFX). The sixth-ranked address has primarily been associated with many ransomware attacks, and the ninth-ranked address is associated with "inhumane" bank account theft through remote desktop software. The remaining addresses on the top-10 list have been reported as organized Bitcoin scam groups that also make worldwide financial "investment in terror", especially in the US, Russia, and Eastern and Central Europe. One reporter of several of these reports claims to have worked for the organized criminals using these addresses for various illicit Bitcoin abuse (e.g., financial scams, begging scams, etc.) and for financial support of "terror".

The most common cases: We next identify the most common cases and how much money these accounts attract. For this, refer to the cumulative distribution function (CDF) of the total received bitcoins per address shown in Figure 2(c). 67% of the reported addresses did not receive any bitcoins at all (i.e., the CDF starts at 0.67), suggesting that many of the reported addresses were not successful in attracting funds. Furthermore, among the addresses that received some funds, most received between 0.01 and 100 bitcoins, with the frequency in this interval being s-shaped on log-scale, suggesting a log-normal-like distribution for this region. In total, these addresses make up 26% of all the reported addresses.

Table 3. Summary of distributions and their model fits with individual x_{min} and their goodness-of-fit (Kolmogorov-Smirnov distance) to the empirical data.

Distribution	$f(x)$	$x_{\min}$	Shape parameter(s)	KS
Power law	$f(x) = Cx^{-\alpha}$	1	$\alpha = 1.35$	0.057
Power + Cutoff	$f(x) = Cx^{-\alpha}e^{-x/\beta}$	1	$\alpha = 1.36$, $\beta = 4.71 \cdot 10^{-7}$	0.099
Lognormal	$\frac{1}{x} \cdot e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$	8,372	$\mu = 9.72$, $\sigma = 2.17$	0.035
Stretched Exponential	$\lambda\beta x^{\beta-1}e^{-\lambda x^{\beta}}$	9,444	$\beta = 0.30$, $\lambda = 1.01 \cdot 10^{-4}$	0.036

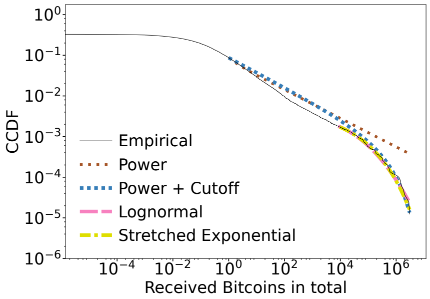


Fig. 3. Curve fitting comparison of the CCDF when computing x_{min} for each class.

Heavy-tailed distribution: As seen in Figures 2(a) and 2(b), a smaller subset of addresses are responsible for the majority of the received bitcoins, suggesting that the distribution may be heavy-tailed. This is confirmed in Figure 2(d), where we plot the CCDF of the amount of bitcoins received per address (with both axes on log-scale). While the distribution clearly is heavy-tailed (i.e., heavier than an exponential), the curvature towards the end suggests that the tail is not power law (as often seen in the wild). We next model this tail behavior and discuss potential implications.

3.2 Model of the tail distribution

To better understand the shape of the tail, we applied model fitting using the following probability distributions: (1) power law, (2) power law with exponential cutoff, (3) lognormal, and (4) stretched exponential. For each class, we determined both the $x_{\min}$ from which the distribution gave the best goodness-of-fit (using the Kolmogorov–Smirnov test [21]) and the best model parameters (using maximum likelihood estimation [24]). Table 3 summarizes the selected parameters and Figure 3 shows the curves fitted from their respective $x_{\min}$ values. While lognormal and stretched exponential provide the best fits for the range they are fitted, we note that they only capture the very end of the tail (as they use $x_{\min}$ values of 8,372 and 9,444, respectively). In contrast, the power-law-based distributions capture a much bigger portion of the tail properly (both models using $x_{\min} = 1$). Of these two distributions, we note that the power-law distribution

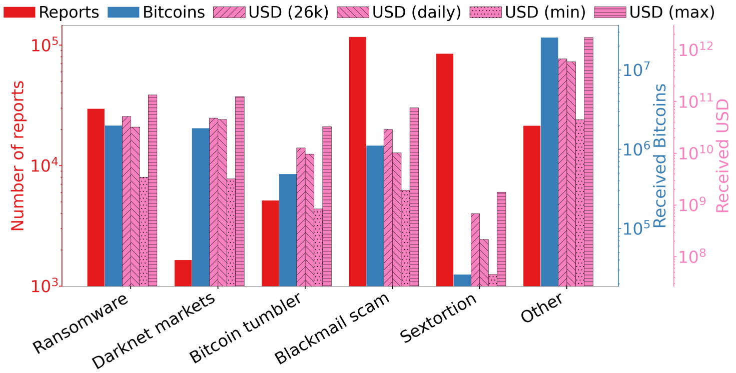


Fig. 4. Categories overview.

with the exponential cutoff better captures the shape of the distribution visually, while the pure power-law function has a smaller Kolmogorov–Smirnov distance (as it better captures the convex-shaped portion of the body of the distribution, which due to the shape of the distribution is given more weight).

The presence of an exponential cutoff (as observed here) typically suggests the presence of a finite resource (e.g., funds from victims) or some form of constraints that limits extreme events (e.g., how much funds can be practically be obtained from victims). We also expect (conjecture) that some of the most successful actors may be responsible for several addresses, over which the funds may be distributed, both as a way to spread risks and as a means to make it more difficult to track funds). This is also suggested by several of the top addresses being reported by a person claiming to have worked for the fake companies responsible for these addresses. One positive aspect here is that the presence of cutoffs can make predictions more reliable compared to systems with a power-law tail (often observed in nature, for example), as there appears to be some rough upper limit on how much funds these accounts have attracted.

4 Category-based High-level Characterization

This section examines the Bitcoin usage as seen across the different address categories reported. After a high-level comparison of the relative Bitcoin usage associated with the different categories (§4.1), we turn our attention to the transactions (§4.2) and reports (§4.3) themselves.

4.1 High-level comparisons

Consider first the number of reports received by Bitcoin Abuse regarding each abuse category and the number of bitcoins that each category of addresses received. Figure 4 summarizes these statistics for each of the abuse categories used by Bitcoin Abuse. To put the number of received bitcoins (shown in blue) in perspective we also show estimates and bounds of the corresponding transfer amounts in USD (shown in pink). Here, we include four estimates: (1) based on our 26k USD/BTC estimate of the average price during 2023, (2) based on daily closing value of BTC price at the day of each individual transaction, as estimated using Yahoo finance source, (3) based on the minimum value of BTC during our abuse report's collection period (2017-05-16 to 2022-04-25), and (4) based on the maximum value of this same time period. While the first two values are used as estimates, the latter two can seen as very rough lower and upper bounds of the price of Bitcoins over the period of interest ($1,734 and $67,567, respectively).

The two most reported abuse categories (i.e., "Blackmail scam" and "Sextortion") are among the three categories that attracted the least funds to the addresses associated with their reported attacks. While this may suggest that these attacks are not very successful, we note that the amount of money they attracted still are non-negligible. For example, while the addresses reported in the 120K reports about Blackmail scams "only" received a modest 1.1M bitcoins, this still corresponds to $29B USD (based on 26k/BTC) or $10B (based on daily closing price). Similarly, the modest 26k bitcoins obtained via addresses associated with Sextortion campaigns are still worth roughly $0.69B USD (based on 26k/BTC) or $0.22B (based on daily closing price).

Having said that, these amounts are very small compared to the amounts paid to the addresses of the reported Ransomware attacks (2.0M bitcoins worth $52B USD or $32B USD based on the average 2023 price and daily closing estimates, respectively) or Darknet markets (1.9M bitcoins worth $49B USD or $45B USD, respectively), not to mention the "Other" category (26M bitcoins worth $670B USD or $590 USD, respectively). As noted, this category includes the top addresses observed in our dataset (e.g., Table 2), including trading/investment scams, remote bank account theft, and "investment[s] in terror".

Also, the reported Bitcoin tumbler sees significant funds passing through them. Here we note that Bitcoin tumbler, also known as a mixing service, combines and shuffles bitcoins from different sources to obscure their original origin, making them an attractive service to be used by the organizations behind many of the illicit addresses. The propensity to employ Bitcoin tumblers for such purposes is evident, as detailed and explored further in Section 6, where we trace bitcoin flows within and across addresses linked to various illicit activities.

Fraction of addresses not attracting any funds: The relative success of the addresses associated with each abuse category becomes even clearer when looking at the distribution statistics. Figure 5 shows a CDF for each category. The leftmost point on each CDF indicates the fraction of accounts in that category that did not receive any bitcoins. Sextortion (93% of addresses), Blackmail scams (79%) and Ransomware (72%) had significantly more accounts that did not receive any funds (i.e., zero bitcoins) compared to the addresses associated with Bitcoin tumbler (12%), "Other" category (10%), and Darknet markets (6%).

Distribution comparisons: While most addresses that received funds ranged between 0.01 and 100 bitcoins (seen by the s-shaped step in the CDFs for this region), we observe a noticeable shift in the distributions. For example, referring to the CDFs in Figure 5, we observe a clear separation between where the dif-

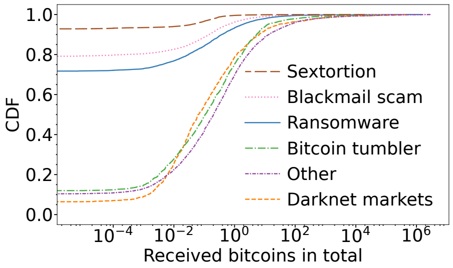


Fig. 5. CDF of received bitcoins.

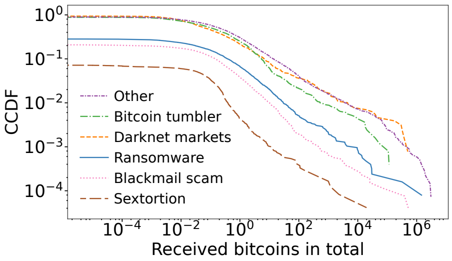


Fig. 6. CCDF of received bitcoins.

Table 4. Power-law fitting of per-category CCDFs.

	Slope estimate		Confidence interval
Category	$x_{\min}$	α (σ)	95%
Sextortion	1	1.423 (0.041)	$\alpha \pm 0.000518$
Blackmail	1	1.419 (0.013)	$\alpha \pm 0.000161$
Ransomw.	1	1.388 (0.013)	$\alpha \pm 0.000234$
Darknet	1	1.309 (0.019)	$\alpha \pm 0.00101$
Tumbler	1	1.391 (0.016)	$\alpha \pm 0.000612$
Other	1	1.329 (0.005)	$\alpha \pm 0.0000851$

ferent distributions approach one. This separation is more visible in the CCDFs shown in Figure 6. Here, the labels of each class are ordered based on the number of accounts that received at least one bitcoin. We observe three distinct groups: (1) Sextortion addresses obtained the least funds, (2) Blackmail scams and Ransomware addresses in general received distinctly more but typically not as much as (3) the addresses associated with Darknet markets, Bitcoin tumbler, and the addresses those in the “Other” category.

Furthermore, when broken down on a per-category basis, the CCDFs become significantly more power-law-like (compared with the aggregate curve in Figure 3), with clear straight-line behavior when plotted on log scale. This is further confirmed by the power-law fitting of each curve. Table 4 summarizes these fittings, with corresponding confidence intervals on the slope parameter. The slopes are relatively clustered around the range $1.31 \leq \alpha \leq 1.42$, each with a relatively tight confidence interval, and together encompassing the slope of the aggregate curve (α=1.35). Rather than the small slope variations, the most visible difference between the CCDFs is instead their relative shift to each other.

4.2 Transactions-based analysis

There are two primary contributors to the differences seen in the distributions of the number of bitcoins received per address when comparing the different abuse categories: the transaction sizes and number of transactions. First, as shown in Figure 7, the size distributions of individual transactions differ substantially between the categories. Here, we note a clear shift in the size distributions; e.g., captured by noticeable differences when looking at the upper percentiles.

Second, in addition to bigger transactions, the most successful addresses in these categories also received more transactions. To capture the strong correla-

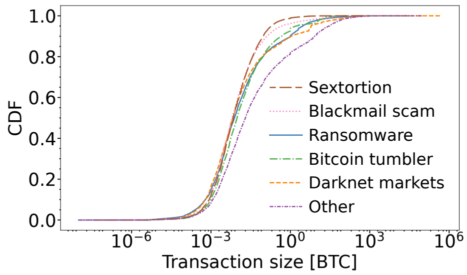


Fig. 7. CDF of received transaction sizes per category.

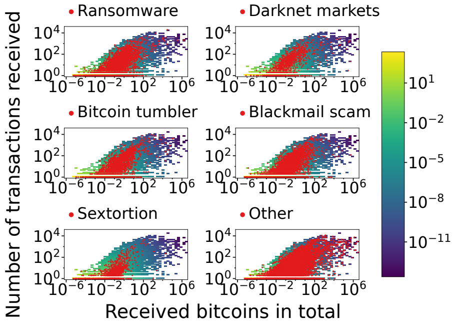


Fig. 8. Received bitcoins vs. the number of received transactions. Per-category scatterplots (red) overlayed on overall heatmaps (gradient color).

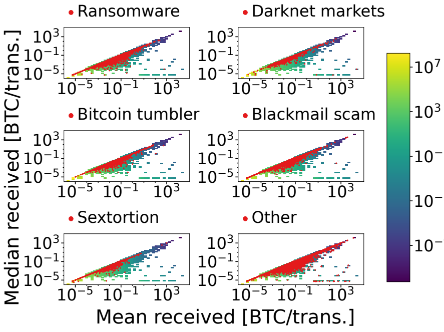
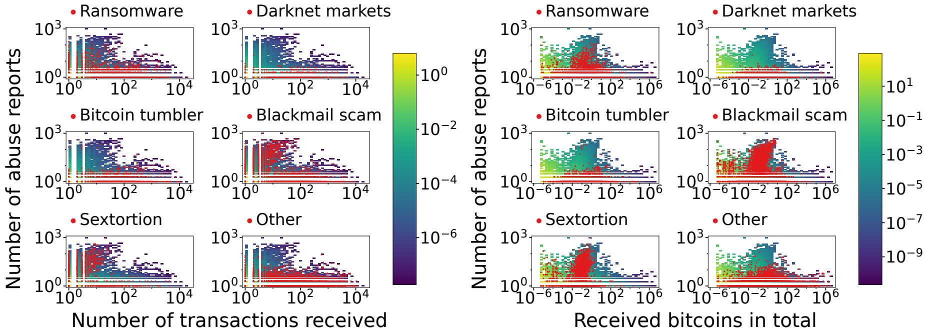


Fig. 9. Mean vs. median number of received bitcoins per transaction. Per-category scatterplots (red) overlayed on overall heatmaps (gradient color).

tion between how successful individual addresses were at attracting funds and the number of victims, Figure 8 shows per-category scatterplots of the received bitcoins (per address) and the number of received transactions (per address) for each category. To simplify comparisons between categories, the scatterplots (shown using red points) are overlayed on a heatmap of the overall per-address distribution (across all categories). Here, the color in the heatmap shows the probability density function (PDF) of addresses observed with that combination. While the distributions for the first four categories (i.e., Ransomware, Darknet markets, Bitcoin tumbler, and Blackmail scams) look relatively similar, with a clear cluster receiving up-to 100 bitcoins spread over up-to 1K incoming transactions, Sextortion and the “Other” category stand out. Again, Sextortion addresses receive fewer bitcoins and fewer transactions compared to other categories. Notably, the “Other" category includes many of the the highest receiving addresses (e.g., with 1K—2M received bitcoins) and some of the addresses with the highest transaction counts.

High skew in transaction sizes within each category: While we observe a high correlation between the number of transactions and the number of received bitcoins, especially when looking at individual categories, there are several noticeable exceptions. One reason for this is the high skew in the size distribution of transactions (CDFs in Figure 7).



(a) Received transactions vs. # reports (b) Received bitcoins vs. # reports

Fig. 10. Successfulness of addresses with different numbers of reports. Per-category scatterplots (red) overlayed on overall heatmaps (gradient color).

High skew in transaction sizes for individual addresses: We have also seen that the sizes can differ substantially for the transactions of individual addresses, best visualized in Figure 9, where we plot the median vs. mean number of received bitcoins per transaction and address. Here, the addresses with relatively symmetric size distributions fall close to the diagonal and those with high skew fall below the diagonal. Perhaps most noticeable is a "line" of addresses at the bottom right of "Other". Those addresses have a very high mean but low median, due to a few very large incoming transactions driving up the mean significantly together with many smaller transactions dragging down the median. We expect that addresses with higher skew may be used for a more diverse set of abuses, targeting both "big" and "small" actors.

4.3 Report frequencies

It is expected that (low-effort) attacks targeting many users will see many reports. It is therefore not surprising to see the much higher report frequencies of Blackmail scams and Sextortion abuse in Figure 4. What is perhaps more interesting is that all categories, including these two categories, include a noticeable mix of low-effort attacks (e.g., spam campaigns) and high-effort, directed attacks. In addition to explaining the high skews observed in how successful different addresses within a category are at attracting funds (both in terms of bitcoins and incoming transactions), we note that these differences also can be observed in the relatively lower correlation between the number of transactions and reports (Figure 10(a)), as well as between the number of received bitcoins and the number of observed reports (Figure 10(b)).

Addresses best at attracting funds are not highly reported: Referring to Figure 10(b), we note that the most successful addresses at attracting funds are only reported a few times (perhaps representing targeted efforts; e.g., of big companies that do not want the attack to be known) and that the number of reports per address are relatively independent of the quantity of received

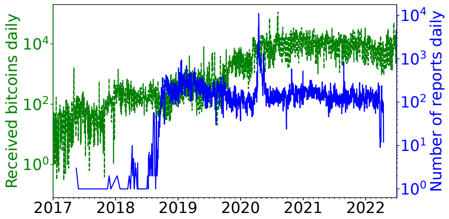


Fig. 11. Timeline of received bitcoins and number of reports between Jan 2017 and July 2022.

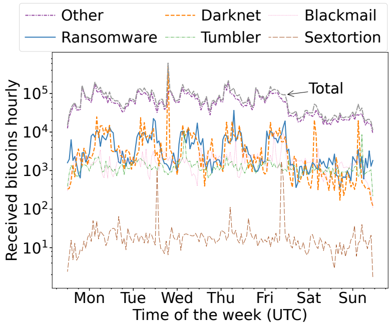


Fig. 12. Received bitcoins by time of the week. Weekday ticks show 12 pm UTC.

funds when considering the three most successful categories: Darknet markets, Bitcoin tumbler, and "Other". In contrast, the most reported addresses (most belonging to the other three categories: Ransomware, Blackmail, Sextortion) typically received less than 100 bitcoins.

5 Temporal Analysis

5.1 Longitudinal timeline

High-level timeline: Figure 11 shows the daily number of reports between January 2017 and July 2022 (blue) together with the total daily number of bitcoins received by the reported set of addresses (green). We note that the Bitcoin Abuse Database was created in 2017 and gained popularity in late 2018 when it saw a steep rise in the number of reports (blue curve). The daily report count has remained relatively steady (at an order of 100's per day) since the beginning of 2019, with exceptions for some temporal peaks and dips. The timeline of the number of received bitcoins per day is more concerning, as there has been a substantial (rough $100x$) increase from O(100) to O(10,000) of bitcoins transferred to these addresses per day over the three-year period that reporting has been relatively stable (i.e., 2019–2022).

Noteworthy spikes: There are several noteworthy spikes in the reporting. The biggest spike by far was observed on April 16th, 2020. On this day, 11K reports were filed, which is roughly 100 times more than the daily average (of 100) for the surrounding days. Our investigation revealed that many news articles around that time warned about a particular style of scam email reported by both the US [34] and Australian [2] governments. In these emails, the attacker (falsely) claims that they have recorded the victim visiting an adult website, while also showing the victim one of their passwords (likely from a leak) in the email.

Looking at the reports for this day, it is clear that a lot of the reports are talking about the same type of attack, matching the descriptions in the articles mentioned. We observed a mix of descriptions written by the victims themselves as well as copies of the emails they received. In many cases, the scammers asked for $1,000 or $2,000 to be paid in bitcoins and displayed one valid password belonging to the targeted victim as "proof" that they know the password.

5.2 Time of the week

Figure 12 shows the time of the week that bitcoins were received for each category. This figure reveals both daily diurnal patterns (with much bigger volumes during daytime/evenings (UTC)) and more funds being transferred during weekdays than weekends. These observations are clearly seen by looking at total volume transferred per hour (grey line at the top marked with "Total") in Figure 12 as well as the Ransomware, Darknet markets, and "Other" categories. With the victims of these categories more often paying during daytime and during the weekdays, suggests that these attacks may be more likely to hit victims on their work computers or that the scammers work during business hours. In contrast, the other categories (e.g., sextortion, tumbler) do not have a strong time of day or day of week pattern, with Bitcoin tumbler seeing the least pronounced patterns, possibly suggesting some level of automation. Here it should be noted that Bitcoin tumbler typically aim at pooling and redistributing the funds at random intervals, with the aim to enhance the anonymity of Bitcoin and achieve effective money laundering. In the case of sextortion, we also expect these scams to reach people on personal devices (or any device) and not just during work hours.

Finally, the biggest relative spikes can be seen for Sextortion, Blackmail scams, and Darknet markets. These spikes are due to large individual transactions affecting these typically smaller volume categories more. For example, the two by far largest transactions (1,000 BTC and 650 BTC, respectively) in the Sextortion category directly line up with the two biggest Sextortion spikes.

5.3 Initial report date analysis

We next consider the timing of the payments to an address relative to the first time that the address was reported to Bitcoin Abuse. This is illustrated on a per-category basis in Figure 13. This shows the CDFs of the relative time each transaction was made in relation to the *first* time the address was reported.

This figure provides several interesting insights. First, addresses are reported around the time that their incoming transaction count is high, indicating that the first report often is made around the time of the abuse's highest activity. This may be a reflection of a significant portion of the addresses only being used for specific attacks. However, significant category differences are observed, with Ransomware displaying the highest concentration around the time of the address first being reported, while Darknet markets exhibit the least concentration.

Second, for all categories except Darknet markets, most transactions take place before the first report is even filled. This may in part be a reflection of most victims not reporting addresses engaging in illicit behaviors.

While some might report the addresses somewhere other than Bitcoin Abuse, the large share of transactions before the first time an address is reported (to Bitcoin Abuse) also suggests that unless reporting behaviors change, there may

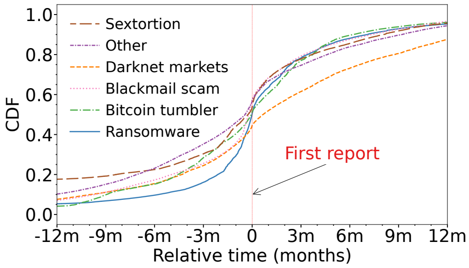


Fig. 13. CDFs of the relative timing of the transactions compared to the first reporting date of an address.

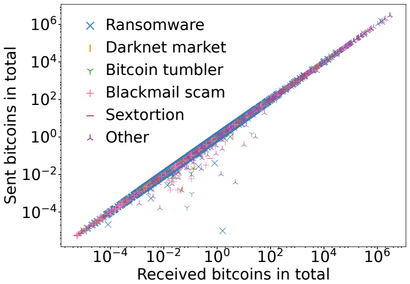


Fig. 14. Scatterplot of received and sent bitcoins.

be limited effectiveness to using such reports to “ban” addresses. As seen here, there is a need for a recognized central anti-fraud mechanism for cryptocurrencies. Without a centralized mechanism, malicious activity will exploit the gaps.

6 Following the money

In this section, we analyze and share insights learned from following the outflow of bitcoins from the reported addresses.

Bitcoins temporarily passing through reported addresses: This analysis is of particular interest since nearly all reported addresses have sent as many bitcoins as they received, leaving a balance of zero. This is illustrated in Figure 14, where we show the total number of bitcoins sent vs. received (per address). This shows that the bitcoins only temporarily pass through the reported addresses, suggesting that these addresses typically are not the wallets that the perpetrators use to store their ill-obtained monetary gains. The following sections compare and share insights for different categories of reported addresses.

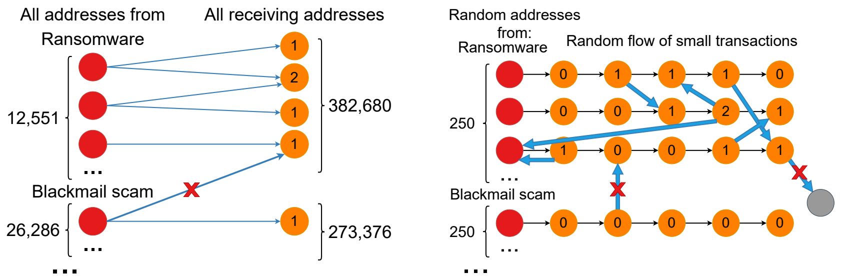
Scope of analysis: The main goal of this analysis is to study and compare the potential concentration or dispersion of money for different abuse address categories. We defer to government agencies and law enforcement to identify individuals or organizations that extract or use the money.

6.1 Following the money methodology

The analysis thus far has not required us to keep track of who has transferred funds to whom. We simply counted the funds transferred. However, when following the money paid from one address to another (as done next), greater attention to detail is needed. We next describe the main challenge with this type of analysis and the decisions (and their limitations) we made to address this.

Basics: All Bitcoin transactions consists of one or more inputs (previous transactions) that are transferred to one or more outputs.

Trivial cases: Since each input and output is labeled with how many bitcoins an address is contributing/receiving as the result of a transaction, for cases where there is a single input the transactions can easily be determined regardless if


All addresses from Ransomware
All receiving addresses
12,551
382,680
Blackmail scam
26,286
273,376
Random addresses from: Ransomware
Random flow of small transactions
250
Blackmail scam
250

(a) One-step analysis (b) Multi-step analysis

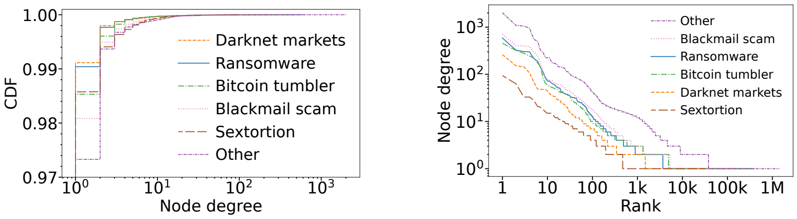
Fig. 15. Visualization of the follow the money analysis. Here, the node degrees are in the orange circles. For the multi-step analysis, we show both the basic chain of transactions (black) and "concentration" edges (blue arrows).

there are one or more outputs, since these can be split into several separate (virtual) transactions. By symmetry, the same approach works when there is a single output but one or more inputs that are part of the transaction.

Challenging case and our solution: However, when there are multiple inputs *and* multiple outputs, the situation is like a melting pot and it is not obvious which bitcoins ended up where. In some cases, it is still possible to determine who transferred the funds to whom (with some accuracy) using heuristics based on the input/output sizes. However, to avoid introducing potential inaccuracies, for the analysis presented here we opted to not use any transactions matching for the last case of our "follow the money" analysis (coming next) and instead only consider the transactions for which we are sure exactly who sent what bitcoins to whom. Fortunately, there are very many transactions that have limited inputs or outputs, allowing us to identify how funds flow between a series of Bitcoin addresses. (We again note that this limitation only is applied from here on and that it does not impact any of the analysis presented earlier in the paper.)

6.2 One-step concentration or dispersion

Let's first follow the outgoing money from the reported addresses only *one* step. In particular, we consider the concentration of addresses that the reported addresses transfer funds directly to. Figure 15(a) illustrates how we did this analysis. First, for each category, we added a link from the reported addressees (red circles on the left) belonging to that category to any address that it directly transferred funds to (conclusively). Second, for each receiving first-hop address (orange circles) we count and report how many reporting addresses each such address has received at least some funds from. This corresponds to the in-degree of each (orange) address to the right in the graph. For example, if two reported addresses both send money to address x, then address x has a node degree of two. Please note that this metric does not consider how many transactions an address x receives, only how many unique senders (in this case from the set of reported addresses) that it receives some funds from.



(a) CDF

(b) Rank plot

Fig. 16. Node degrees distributions of receiving addresses being sent money from addresses in each category.

Table 5. Address expansion ratio comparison of categories when going one step deep.

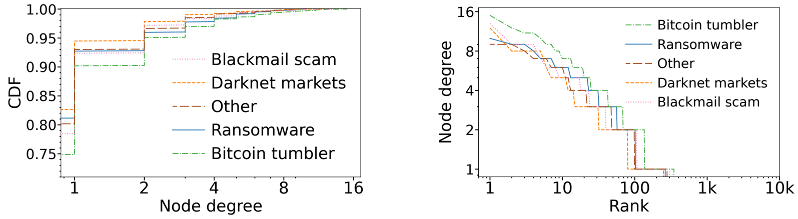
Category	Abuse addr. (#)	Addr. one level deep (#)	Expansion ratio
Sextortion	24,218	32,982	1.36
Blackmail scam	26,286	273,376	10.40
Ransomware	12,551	382,680	30.49
Darknet markets	1,289	168,542	130.75
Bitcoin tumbler	2,507	334,160	133.29
Other	13,736	1,419,902	103.37

Figures 16(a) and 16(b) show the CDFs and rank plots, respectively, of the per-category node degrees. First, we note a Zipf-like distribution (i.e., relatively straight-line behavior in the rank plots shown on log-log scale), capturing a high skew among the nodes that do receive money flows from multiple abuse addresses. For example, each category includes an address that received funds from 100 or more abuse addresses. Notably, an address in the “Other" category received funds from 2,000 abuse addresses, aligning with reports of organized crime using such addresses to pool funds from various attack vectors for global financial “investment in terror".

However, perhaps the main observation is the very long tail of addresses with node degree one. For example, looking at the CDF, 97-99% of the addresses associated with each category have a node degree of one (the minimum), meaning that nearly all receiving addresses are not visibly related when only tracing the money one step. This suggests that the money mostly is spread out across even more addresses after the first step, when going only one step deep from the reported addresses. This is confirmed when looking at the total addresses seen one level deep and the relative expansion ratio of each category, shown in Table 5. Notably, Bitcoin tumblers, Darknet markets, and “Other" categories exhibit significantly larger expansion ratios compared to the rest.

6.3 Multi-step analysis

Having seen little concentration when following the transactions only one step, we next performed a multi-step analysis to track the money flow several steps deep. For this analysis, we again wanted to compare the categories fairly head-



(a) CDF (b) Rank plot

Fig. 17. CDF and rank plot of the node degrees of addresses in the chain, for respective category, when counting only blue "concentration" edges.

to-head and be able to answer questions such as whether some categories are more likely to shuffle around the money a couple of steps only to collect the money a few steps later.

For fair head-to-head comparison, we developed a "tracking the pennies" approach in which we tracked an equal amount of "penny flows" (flow of small transactions) as bitcoins were moved five steps deep. Figure 15(b) provides a visual overview of our sampling and tracking methodology. More specifically, for each category, we used a random (depth first) search to find 250 randomly selected reported addresses from which we were able to trace back at least one random chain of money five steps deep from that address. For each step in this search, a new random transaction was chosen from the last address in the chain. We typically gave preference to a small transaction (under 0.1 bitcoins), and if no such transaction existed, we used any random transaction as a fallback. Finally, if a chain of transactions reached a dead end (i.e., where there are no more outgoing transactions), the latest address was removed from the chain and a new random transaction (new path) was chosen.

The choice of picking 250 was made to ensure that we have a substantial number of random paths for each category (so that numerical properties can be compared with some statistical confidence). However, we note that this choice forces us to drop Sextortion from the analysis since we did not find 250 full five-step paths for this category. Therefore, the following analysis focuses only on the other five categories.

Furthermore, for the node degree analysis we did not count all types of address relationships that we identify (but report on these separately). For example, as illustrated with an × in Figure 15(b), we did not count the basic chain (black arrows), cross-category links (e.g., vertical arrow in the figure) or links to addresses further away than five steps deep.

Fewer addresses with in-degree one: While all categories still had nodes with an in-degree of one, compared to the one-step analysis, this fraction reduced noticeably (from 97-99% to 83-89%). Here, Bitcoin tumbler saw the biggest reduction (from 98.5% down to 83%).

Bitcoin tumblers: In general, Bitcoin tumbler stands out in our multi-step analysis. For example, in addition to the above finding, it stands out with higher node degrees (e.g., see CDFs of "concentration" edges in Figure 17(a)) and sub-

Table 6. Comparison of graph and transaction metrics calculated on each category's flow graph.

	Graph metrics		Transaction metrics		
Category	Connecting edges	Loops	To reported	To reported category	To reported 250
Ransomware	254	4	2,526	399	169
Darknet	209	17	2,240	269	263
Tumbler	354	25	2,524	585	97
Blackmail	267	16	2,239	307	154
Other	247	10	5,164	3,155	1,299

stantially higher concentration (e.g., see the rank plot in Figure 17(b)) than the other categories. These findings suggest that fewer actors are involved with this category, typically used for anonymity/money laundering. These differences can perhaps be explained by the increased effort associated with running such addresses. In particular, we note that tumbling bitcoins typically requires more effort and expertise than simply sending blackmail scam emails (which has the lowest node degree and concentration of the categories). Another potential explanation is that most Bitcoin-using miscreants use tumblers, and there are more users of tumblers than tumblers.

Money-flow comparisons: When looking closer at the structure of the graphs formed by the 1,250 edges (250 chains × 5 steps) we observed significant differences between the categories. The first two columns of Table 6 summarize some of these properties. Here, the "Connecting edges" represent the blue arrows in Figure 15(b), which are transactions among the set of addresses in the graph, excluding the "penny flow" itself (the white arrows in Figure 15(b)) and "Loops" measures the number of distinct cycles that exist in the graph structure for that category. Looking at these two metrics, Bitcoin tumbler again stands out with significantly higher "connecting edge" (354) and "loop" (25) counts than the other categories. This again matches the intuition that the addresses in this category are more likely to send money among a relatively smaller set of addresses. In contrast, Darknet markets have the fewest "connecting edges" (209) and Ransomware has by far fewest "loops" (4).

Transaction-based analysis on the graph: Finally, we have found that some of the edges go back to the original reported addresses, and that these in some cases carry a non-negligible number of transactions. The remaining columns of Table 6 summarizes the metrics we used here, where "to reported" counts the number of transactions going back to any of the 267K reported addresses, "to reported category" counts transactions to any reported address in the category, and "to reported 250" only counts transactions back to the 250 randomly picked reported addresses of that category.

Here, we again observe some major differences between the categories. First, the "Other" category has much more transactions going back to reported addresses, especially to its own category (3,155 transactions compared to 585 for the second-ranked category) as well as back to the 250 random (reported) addresses of its own category (1,299 compared to 263 for the second-ranked category).

Table 7. Transactions across categories.

From/to	Ransomware	Darknet	Tumbler	Blackmail	Other
Ransomware	-	982	2,658	1,566	1,228
Darknet	767	-	2,126	664	915
Tumbler	2,173	1,638	-	2,282	2,588
Blackmail	1,512	914	4,087	-	2,230
Other	1,086	1,825	3,015	1,903	-

These findings suggest that a significant number of transactions are directed towards some of the reported addresses in the “Other” category.

Transactions across categories: To better understand how money flowed between addresses associated with the different categories, we next counted the transactions made between the subgraphs of each category (note that these were not included above, since we did not include that type of cross-category “connecting edges” in the original graph analysis (marked with × in Figure 15(b)). Table 7 summarizes the total number of transactions over such cross-category edges. Here, Bitcoin tumbler again stands out with both a higher inflow and outflow of transactions to/from the category compared to to/from the other categories. Given the nature of Bitcoin tumbler (money laundering), it also makes sense that other categories are interested in their service, which may explain why all categories have more outgoing transactions to Bitcoin tumbler to any of the other categories; e.g., Ransomware (2,658 vs. 1,566 for 2nd ranked), Darknet markets (2,126 vs. 915 for 2nd ranked), Blackmail scams (4,087 vs. 2,230 for 2nd ranked), and “Other” (3,015 vs. 1,903 for 2nd ranked).

Finally, we note that the “Other" category receives the most transactions (2,588) from Tumbler addresses (of all categories) and send the most transactions (1,825) of any category outside Tumblers to Darknet addresses and the second-most transactions (3,015) after Blackmail (4,087) to Tumbler addresses. While it is difficult to label these accounts or pinpoint where their funds are directed based only on this analysis, it is clear that several of these accounts play a central role in the transfer of illicit funds (matching some reporters’ claims that some of these addresses are associated with organized crime). Here, it should also be noted that our methodology did not track the amount of funds transferred between the address types, so the skew in transferred funds between the account types may be substantially bigger (and different) when accounting for the transactions of the “Other" category generally being bigger than for the rest (e.g., Figure 7).

7 Related Work

Anonymity: Many works study anonymity aspects of Bitcoin and identify ways to deanonymize users. Reid and Harrigan [31] present an early study of the anonymity aspects of Bitcoin. Herrera-Joancomart [12] provides an exhaustive review of Bitcoin anonymity. Biryukov et al. [3] show that combining Bitcoin with the anonymizing service Tor creates a new attack vector, jeopardizing their privacy. Androulaki et al. [1] investigate user privacy in Bitcoin by simulating usage of Bitcoin in accordance with Bitcoin's recommended privacy measures, finding that almost 40% of the simulated participants could be accurately profiled using behavior-based clustering techniques. Meiklejohn et al. [22] discuss the challenges Bitcoin's public flow of transactions causes for larger-scale criminal and fraudulent activity. Harrigan et al. [11] explain how "unreasonably" effective address clustering is — i.e., heuristics that group addresses together.

Criminal activities: As we are not the first to characterize the landscape of Bitcoin abuse, many previous works have studied criminal activities related to Bitcoin [28,27,41,10,32,14,19]. For example, Pastrana et al. [28] perform a large measurement of 4.5M crypto-mining malware samples, revealing campaigns with multi-million dollar earnings. Pastrana et al. [27] measure the practice of "eWhoring" (selling photos and videos with sexual content of another person). While most transactions involved PayPal and Amazon giftcards, Bitcoin was found to be a popular tool for offloading eWhoring profits.

Money laundering and tumbling: Möser et al. [23] present the first study on Bitcoin money laundering (tumblers) and conclude that applying a Know-Your-Customer principle to Bitcoin is likely not possible. Others have created protocols that facilitate the service of mixing (tumbling) transactions. Bonneau et al. [6] propose a protocol called MixCoin, later improved to BlindCoin by Valenta and Rowan [40]. Concurrently, Ruffing et al. [33] proposes a decentralized mixing system called CoinShuffle.

Ransomware: Bitcoin use with ransomware [35,29,37,17,42,13,9,20] is well studied. For example, Kharraz et al. [17] present a long-term study of observed ransomware attacks between 2006 and 2014. More recently, Wang et al. [42] present a large-scale empirical analysis based on data from 2012–2021. Huang et al. [13] study the landscape of ransomware and trace the money-flow from when the victim acquires Bitcoins to when the perpetrator converts it back to fiat. For a two-year period, they trace 19,750 victims' (likely) ransom payments of more than $16M. Conti et al. [9] conduct a large study of the economic impact of many different ransomwares from a perspective of Bitcoin transactions, including ransomwares like WannaCry, Jigsaw and many more. Liao et al. [20] focus on one particular family of ransomware called CryptoLocker, i.e., ransomware that simply encrypts files until the ransom is paid.

Sextortion: Paquet-Clouston et al. [26] study sextortion spam that requires a payment in Bitcoin using a dataset of 4M entries, concluding that one entity is likely behind the majority of them and has gained around $1.3M over an 11-month period. Oggier et al. [25] also analyze sextortions, but focus on those where the victim is blackmailed (scammed) *with* compromising sexual information, rather than being blackmailed *into* committing sexual actions.

Darknet markets: Christin [8] perform a measurement analysis of the darknet market Silk Road over an 8-month period in 2012. Broséus et al. [7] study the structure and organization of darknet markets from a Canadian perspective. Lee et al. [18] study how criminals abuse cryptocurrencies on the dark web using over 27M dark webpages and 10M Bitcoin addresses, learning that more than 80% of the addresses on the dark web were used for malicious activity.

8 Conclusion

This paper presented a comprehensive analysis of money flows to and from Bitcoin addresses linked to different abuse categories. Our analysis revealed valuable insights for understanding Bitcoin transactions linked to illicit activities, guiding future efforts to combat cryptocurrency-related illicit activities, with significant implications for legitimate users and stakeholders.

First, our high-level characterization of money flows revealed substantial variations in flow patterns, report rates, and success rates within and across addresses associated with different abuse categories (e.g., high skew, heavy tails, and big differences between categories). This understanding aids law enforcement and regulators in identifying patterns and trends in illegal Bitcoin activities, and improve their strategies to detect, prevent, and mitigate illicit activities.

Second, our temporal analysis captured long-term trends, weekly patterns, and the relative timing of when illicit addresses of each category are reported or receive funds. The observed increase in the daily number of Bitcoins received by reported addresses over time (e.g., $\approx \times 100$ over three years) indicates a significant rise in funds transferred to these addresses. This calls for continuous vigilance and adaptive approaches to keep up with the evolving landscape of illegal transactions. Moreover, the weekly and daily variations in activity levels and transaction volumes highlight the importance of targeted enforcement efforts during periods of heightened activity. To counter dynamic illicit transactions and improve reporting, stakeholders should adopt agile, real-time monitoring, proactive intervention, and international cooperation for early warnings.

Third, our analysis of the outflow of bitcoins from reported addresses sheds light on significant differences in graph properties and flow patterns among illicit addresses and between abuse categories. For example, the concentration of funds toward specific addresses, the dispersion of funds after the initial step, and the presence of loops highlight the complexity of money laundering schemes and the need for enhanced measures to track and disrupt these networks. The significant differences in graph structure and transaction patterns between categories, particularly the prominence of Bitcoin tumblers, underscore the importance of addressing the role of specific services in facilitating illicit financial flows.

Finally, our results also highlight that authorities need a coordinated effort to monitor all Bitcoin activities, and cryptocurrency activities in general. While there is a lot of illicit activity on Bitcoin, only a fraction of it gets reported, and at least to Bitcoin Abuse many of the reports come in relatively late. Furthermore, while databases like Bitcoin Abuse are great data sources, we note that researchers (as we have found here) are limited to the availability and APIs provided by those data sources. When they are not accessible, transparency is hurt, emphasizing the need for transparent methods to monitor fraud on these networks. We advocate for global government collaboration on an official centralized abuse monitoring effort for all cryptocurrencies to capture money flows across diverse illicit activities and national borders. Our follow-the-money analysis avoids pointing to specific actors, yet highlights the potential use of sophisticated techniques to identify threat actors involved in multiple illicit activities.

Acknowledgments. This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. We would also like to thank our shepherd Lucianna Kiffer and the anonymous reviewers, who provided valuable feedback as well as several students that have contributed to early results in this project, including Axel Flodmark, Markus Jakum, Vera Antonov, and Karl Söderbäck.

References

1. Androulaki, E., Karame, G.O., Roeschlin, M., Scherer, T., Capkun, S.: Evaluating user privacy in bitcoin. In: Proc. Financial Cryptography and Data Security (FC) (2013)
2. Australian Cyber Security Centre: Sextortion email campaign impacting Australians. `https://www.cyber.gov.au/about-us/alerts/sextortion-email-campaign-impacting-australians` (2020), [Accessed 17-May-2023]
3. Biryukov, A., Pustogarov, I.: Bitcoin over tor isn't a good idea. In: Proc. IEEE Symposium on Security and Privacy (S&P) (2015)
4. BitcoinAbuse: Bitcoin Abuse Database. `https://www.bitcoinabuse.com` (2023), [Accessed 08-May-2023]
5. Blockchain: Blockchain Developer APIs. `https://www.blockchain.com/explorer/api` (2023), [Accessed 25-May-2023]
6. Bonneau, J., Narayanan, A., Miller, A., Clark, J., Kroll, J.A., Felten, E.W.: Mixcoin: Anonymity for bitcoin with accountable mixes. In: Proc. Financial Cryptography and Data Security (FC) (2014)
7. Broséus, J., Rhumorbarbe, D., Mireault, C., Ouellette, V., Crispino, F., Décary-Hétu, D.: Studying illicit drug trafficking on darknet markets: Structure and organisation from a canadian perspective. Forensic Science International (2016)
8. Christin, N.: Traveling the silk road: A measurement analysis of a large anonymous online marketplace. In: Proc. World Wide Web (WWW) (2013)
9. Conti, M., Gangwal, A., Ruj, S.: On the economic significance of ransomware campaigns: A bitcoin transactions perspective. Computers & Security (2018)
10. Gomez, G., Moreno-Sanchez, P., Caballero, J.: Watch your back: Identifying cybercrime financial relationships in bitcoin through back-and-forth exploration. In: Proc. ACM Computer and Communications Security (CCS) (2022)
11. Harrigan, M., Fretter, C.: The unreasonable effectiveness of address clustering. In: Proc. UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld (2016)
12. Herrera-Joancomartí, J.: Research and challenges on bitcoin anonymity. In: Proc. Workshop on Data Privacy Management (DPM) (2015)
13. Huang, D.Y., Aliapoulios, M.M., Li, V.G., Invernizzi, L., Bursztein, E., McRoberts, K., Levin, J., Levchenko, K., Snoeren, A.C., McCoy, D.: Tracking ransomware end-to-end. In: Proc. IEEE Symposium on Security and Privacy (S&P) (2018)
14. Huang, D.Y., Dharmdasani, H., Meiklejohn, S., Dave, V., Grier, C., McCoy, D., Savage, S., Weaver, N., Snoeren, A.C., Levchenko, K.: Botcoin: Monetizing stolen cycles. In: Proc. Network and Distributed System Security Symposium (NDSS) (2014)
15. International Monetary Fund: IMF.org (2023), [Accessed 15-Oct-2023]
16. Investing.com: Bitcoin Historical Data. `https://www.investing.com/crypto/bitcoin/historical-data` (2023), [Accessed 17-May-2023]

17. Kharraz, A., Robertson, W., Balzarotti, D., Bilge, L., Kirda, E.: Cutting the gordian knot: A look under the hood of ransomware attacks. In: Proc. Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA) (2015)
18. Lee, S., Yoon, C., Kang, H., Kim, Y., Kim, Y., Han, D., Son, S., Shin, S.: Cybercriminal minds: an investigative study of cryptocurrency abuses in the dark web. In: Proc. Network and Distributed System Security Symposium (NDSS) (2019)
19. Li, X., Yepuri, A., Nikiforakis, N.: Double and nothing: Understanding and detecting cryptocurrency giveaway scams. In: Proc. Network and Distributed System Security Symposium (NDSS) (2023)
20. Liao, K., Zhao, Z., Doupé, A., Ahn, G.J.: Behind closed doors: measurement and analysis of cryptolocker ransoms in bitcoin. In: Proc. Electronic Crime Research (eCrime) (2016)
21. Massey Jr, F.J.: The kolmogorov-smirnov test for goodness of fit. Journal of the American statistical Association (1951)
22. Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G.M., Savage, S.: A fistful of bitcoins: Characterizing payments among men with no names. In: Proc. ACM Internet Measurement Conference (IMC) (2016)
23. Möser, M., Böhme, R., Breuker, D.: An inquiry into money laundering tools in the bitcoin ecosystem. In: Proc. APWG eCrime Researchers Summit (2013)
24. Myung, I.J.: Tutorial on maximum likelihood estimation. Journal of Mathematical Psychology (2003)
25. Oggier, F., Datta, A., Phetsouvanh, S.: An ego network analysis of sextortionists. Social Network Analysis and Mining (2020)
26. Paquet-Clouston, M., Romiti, M., Haslhofer, B., Charvat, T.: Spams meet cryptocurrencies: Sextortion in the bitcoin ecosystem. In: Proc. ACM Advances in Financial Technologies (AFT) (2019)
27. Pastrana, S., Hutchings, A., Thomas, D., Tapiador, J.: Measuring ewhoring. In: Proc. ACM Internet Measurement Conference (IMC) (2019)
28. Pastrana, S., Suarez-Tangil, G.: A first look at the crypto-mining malware ecosystem: A decade of unrestricted wealth. In: Proc. ACM Internet Measurement Conference (IMC) (2019)
29. Pletinckx, S., Trap, C., Doerr, C.: Malware coordination using the blockchain: An analysis of the cerber ransomware. In: Proc. IEEE Communications and Network Security (CNS) (2018)
30. Portnoff, R.S., Huang, D.Y., Doerfler, P., Afroz, S., McCoy, D.: Backpage and bitcoin: Uncovering human traffickers. In: Proc. Knowledge Discovery and Data Mining (KDD) (2017)
31. Reid, F., Harrigan, M.: An analysis of anonymity in the bitcoin system. Springer (2013)
32. Ron, D., Shamir, A.: How did dread pirate roberts acquire and protect his bitcoin wealth? In: Proc. Financial Cryptography and Data Security (FC) (2014)
33. Ruffing, T., Moreno-Sanchez, P., Kate, A.: Coinshuffle: Practical decentralized coin mixing for bitcoin. In: Proc. ESORICS (2014)
34. Small, B.: Scam emails demand Bitcoin, threaten blackmail — consumer.ftc.gov. `https://consumer.ftc.gov/consumer-alerts/2020/04/scam-emails-demand-bitcoin-threaten-blackmail` (2020), [Accessed 17-May-2023]
35. Spagnuolo, M., Maggi, F., Zanero, S.: Bitiodine: Extracting intelligence from the bitcoin network. In: Proc. Financial Cryptography and Data Security (FC) (2014)

36. Staugh, G.: Bull capital trading review 2022: 5 disturbing facts about bullcapitaltrading.com. https://www.forexbrokerz.com/brokers/bull-capital-trading-review (2023), [Accessed 2-Nov-2023]
37. Taniguchi, T., Griffioen, H., Doerr, C.: Analysis and takeover of the bitcoin-coordinated pony malware. In: Proc. ACM Asia Computer and Communications Security (ASIACCS) (2021)
38. Tekiner, E., Acar, A., Uluagac, A.S., Kirda, E., Selcuk, A.A.: Sok: Cryptojacking malware. In: Proc. IEEE European Symposium on Security and Privacy (EuroS&P) (2021)
39. U.S. Bureau of Economic Analysis (BEA): GDP by State . https://www.bea.gov/data/gdp/gdp-state (2023), [Accessed 15-Oct-2023]
40. Valenta, L., Rowan, B.: Blindcoin: Blinded, accountable mixes for bitcoin. In: Proc. Financial Cryptography and Data Security (FC) (2015)
41. Vu, A.V., Hughes, J., Pete, I., Collier, B., Chua, Y.T., Shumailov, I., Hutchings, A.: Turning up the dial: the evolution of a cybercrime market through set-up, stable, and covid-19 eras. In: Proc. ACM Internet Measurement Conference (IMC) (2020)
42. Wang, K., Pang, J., Chen, D., Zhao, Y., Huang, D., Chen, C., Han, W.: A large-scale empirical analysis of ransomware activities in bitcoin. ACM Trans. Web (2021)

A Ethics

Our research highlights the significant challenges posed by Bitcoin's pseudonymous transactions and lack of regulation, particularly in relation to its use in illicit activities. In the paper, we provide a comprehensive analysis of Bitcoin transactions associated with different categories of illegal activities, conducted so as to ensure that ethical considerations were upheld.

Addresses an important societal problem: As the paper demonstrates, Bitcoin use for illicit activities is clearly widespread and turns over very large sums of money. Many of the abuse types prey on the weak, and it is clear that these activities have increasingly negative societal effects. While the general public often focuses on Bitcoin's energy consumption, much less attention has been put on the numerous victims that fall prey to Bitcoin's role in various illicit activities. We note that while Bitcoin's energy consumption *might* have long-term effects on *humans*, the effect of Bitcoin abuse on humans is both *apparent* and *current*.

Respect privacy and confidentiality of individuals: The data used for analysis is sourced from the Bitcoin Abuse Database, which collects information from reports submitted by victims and other individuals or organizations, and the public Bitcoin chain itself. While the database provides insights into attacks and associated Bitcoin addresses, care is taken to ensure the anonymity of victims and attackers. We recognize that not all victims report their experiences and that reports may come from individuals who did not fall victim themselves.

Adhere to legal and ethical guidelines: The analysis focuses on publicly available blockchain data and information obtained from the Bitcoin Abuse Database. No attempts are made to compromise the security or integrity of the Bitcoin network or any other systems.

Promote responsible and ethical use of the findings: Law enforcement agencies, regulatory bodies, and cryptocurrency service providers can leverage our insights to enhance their strategies, policies, and compliance measures. Overall, the study aims

Table 8. Comparison of recent reporting rates and the volume of new addresses being reported in the past five months, and the transactions they receive.

	Primary dataset	Latest reports
Reports time frame	2017-05-16 – 2022-04-25	2022-12-20 – 2023-05-19
Reports	267,708	17,116
Unique addresses	82,527	2,249
Transactions	5,092,489	141,485
Received bitcoins	31,346,586	269,070
Received in USD	815,011,236,000	6,995,820,000

to contribute to the development of effective strategies to mitigate the risks and vulnerabilities associated with Bitcoin's potential for misuse, promoting a safer and more secure financial landscape.

B Additional statistics

Table 8 provides a high-level overview of the recent reporting rates, the volume of new addresses having been reported in the past five months, and the transactions they receive. We note that reporting rates (114 reports/day on average) are similar to those observed in Figure 11 and that the average transactions/address is almost the same (61.7 for primary vs. 62.9 for latest dataset). The main difference is a reduction in the average number of bitcoins received/address (380 for the primary dataset vs. 120 for the latest dataset) and the rate new unique addresses are observed. These later differences are easily explained by (1) these addresses being earlier in their lifecycle (e.g., Figure 13) and/or (2) a bias towards many of the most successful addresses (in terms of attracting funds; e.g., the top-hitters in the "Other" category) already having been reported. Yet, the large amount of funds that these *newly reported* addresses obtain shows that there continually are many more (new) illicit addresses being reported that are attracting significant funds, including at the present moment.