Cybercriminal Minds: An investigative study of cryptocurrency abuses in the Dark Web

Seunghyeon Lee[†‡] Changhoon Yoon[‡] Heedo Kang[†] Yeonkeun Kim[†]
Yongdae Kim[†] Dongsu Han[†] Sooel Son[†] Seungwon Shin[†‡]
[†]KAIST [‡]S2W LAB Inc.
{seunghyeon, kangheedo, yeonk, yongdaek, dhan.ee, sl.son, claude}@kaist.ac.kr {cy}@s2wlab.com

***Abstract*—The Dark Web is notorious for being a major distribution channel of harmful content as well as unlawful goods. Perpetrators have also used cryptocurrencies to conduct illicit financial transactions while hiding their identities. The limited coverage and outdated data of the Dark Web in previous studies motivated us to conduct an in-depth investigative study to understand how perpetrators abuse cryptocurrencies in the Dark Web. We designed and implemented MFScope, a new framework which collects Dark Web data, extracts cryptocurrency information, and analyzes their usage characteristics on the Dark Web. Specifically, MFScope collected more than 27 million dark webpages and extracted around 10 million unique cryptocurrency addresses for Bitcoin, Ethereum, and Monero. It then classified their usages to identify trades of illicit goods and traced cryptocurrency money flows, to reveal black money operations on the Dark Web. In total, using MFScope we discovered that more than 80% of Bitcoin addresses on the Dark Web were used with malicious intent; their monetary volume was around 180 million USD, and they sent a large sum of their money to several popular cryptocurrency services (e.g., exchange services). Furthermore, we present two real-world unlawful services and demonstrate their Bitcoin transaction traces, which helps in understanding their marketing strategy as well as black money operations.**

I. Introduction

Anonymity is a double-edged sword. On the one hand, it protects the privacy of people, fostering freedom of speech and democracy in oppressive regimes [26]. On the other hand, it is misused to conduct illegal behaviors and even acts of (cyber-) terrorism, while the perpetrators often go unaccounted for their acts. The problem becomes increasingly sophisticated as technology advances and multiple technologies are used in combination.

Today, we are facing two up-to-date techniques for hiding identity: (i) Dark Web and (ii) Cryptocurrency. The Dark Web leverages anonymous routing techniques (e.g., Tor [38]) to conceal the user's identity. While the Dark Web was first proposed to support the freedom of the press and guarantee open discussions without political pressure [49], it is also misused for malicious purposes, such as advertising harmful content [34], [30] and command-and-control servers (C&C). For example, an e-commerce market in the Dark Web is known as one of the major drug trading sites [13], [22], and WannaCry malware, one of the most notorious ransomware, has actively used the Dark Web to operate C&C servers [50]. Cryptocurrency also presents a similar situation. Apart from a centralized server, cryptocurrencies (e.g., Bitcoin [58] and Ethereum [72]) enable people to conduct peer-to-peer trades without central authorities, and thus it is hard to identify trading peers.

Similar to the case of the Dark Web, cryptocurrencies also provide benefits to our society in that they can redesign financial trading mechanisms and thus motivate new business models, but are also adopted in financial crimes (e.g., money laundering) [37], [57]. In fact, several recent studies have pointed out that Bitcoin is used for Ponzi fraud [69], [29] and payment for ransomware [59].

While the abuses of either the Dark Web [34], [30] or cryptocurrency [69], [29] have already been investigated by some researchers, we note that most of them mainly examine either the Dark Web or cryptocurrency separately, and only a few recent studies consider them together [34], [43]. Christin *et al.* measured how Bitcoin has been used in a well-known black market [34], while Foley *et al.* measured the Bitcoin volume used in trading illicit goods and their characteristics [43]. Indeed, they conducted pioneering research work, but it still has critical limitations. They both mainly focus on well-known Dark Web markets (e.g., Silkroad [22]) and thus their analysis results are quite limited to surveying specific market-related operations. Importantly, no previous research has addressed the question: *Where does the money go from online merchants trading illicit services and goods?* The answer to this question advances our understanding of how perpetrators capitalize their money while minimizing the risk of being tracked. Moreover, their data for analysis are quite old (e.g., data collected in 2012 [34]) or only based on known blacklist information (e.g., FBI Bitcoin blacklist [43]). Thus, we believe that their analysis results cannot present recent trends or diverse characteristics of usages of cryptocurrencies in the Dark Web.

As noted, the main goal of this paper is to provide in-depth analysis on the usages of cryptocurrencies, focusing on misuse cases for illicit intent. However, conducting this research is not an easy process, because it presents three key research challenges. First, collecting large-scale data of cryptocurrency on the Dark Web is difficult due to the nature of the Dark Web. Moreover, before collecting cryptocurrency data on the Dark Web, one must also first find a way of collecting Dark Web information. Second, because cryptocurrency is designed for people seeking pseudonymity (i.e., hiding who

Network and Distributed Systems Security (NDSS) Symposium 2019
24-27 February 2019, San Diego, CA, USA
ISBN 1-891562-55-X
https://dx.doi.org/10.14722/ndss.2019.23055
www.ndss-symposium.org

is sending/receiving the money), it is not easy to identify the user/owner of cryptocurrency accounts. Such pseudonymity also exacerbates the manifesting of an entire money flow chain among its transaction participants. Third, even after collecting data related to cryptocurrency in the Dark Web, we still need to gather more information that can be used to reveal its identity for further analysis.

To address the challenges, we design a Dark Web data collection and analysis platform, MFScope. Our platform first extracts seed dark website addresses[1] by leveraging Dark Web indexing services (e.g., Ahmia [1]) and crawls those extracted sites. It also extracts links to other dark websites from the crawled data to increase our data corpus. With this platform, we collect a large number of dark websites (around 27 millions of pages) and cryptocurrency addresses (around 10 millions of unique cryptocurrency addresses). We believe that this large-scale data collection makes our analysis much more solid. *Note that we do NOT claim our data covers most of the Dark Web and cryptocurrency usage within it*, because it is hard to estimate the exact size of the Dark Web. However, we argue that the analysis of a large amount of data will provide a better understanding of the Dark Web and its usage of cryptocurrency than that of other works relying on a small set of data.

On the basis of our analysis, we find that 99.8% of collected cryptocurrency addresses was Bitcoin, which indicates that Bitcoin is the most popular cryptocurrency on the Dark Web. This motivates us to rigorously conduct in-depth analysis of the illicit uses of Bitcoin on the Dark Web. From Bitcoin addresses collected from the Dark Web, we identify more than 80% of Bitcoin addresses used for illicit intent by classifying their usages (e.g., drug dealing and financial fraud). We also estimate how much value (in USD) has been traded through those Bitcoin addresses by considering their market value. We demonstrate that the market size of trading in illicit goods and services via the Dark Web is approximately 180 million USD.

We also compute the money flows from such illicit Bitcoin addresses. We propose a novel algorithm, *Taint-based Bitcoin flow analysis*, which models the volume of illicit Bitcoins transferred from an illicit Bitcoin address to their destinations. This helps us to understand illicit financial transactions on the Dark Web. Based on our taint analysis, we find that the perpetrators tend to send a large sum of their money to several popular Bitcoin exchange services.

We conduct a correlational analysis with the Surface Web to obtain missing information in the Dark Web. By using a combination of data obtained from the different domains, the cross-referencing clues provide crucial information that contributes to demystifying the entire cybercrime scheme. Based on our analysis, we reveal two real-world Dark Web value chains involving Bitcoin, *Bitcoin investment scam and trafficking*.

Our contributions are summarized as follows.

• We design and implement a platform that collects a large number of dark websites and extracts useful cryptocurrency information in them automatically. Here, we also present several techniques to identify valid cryptocurrency addresses and introduce efficient analysis methods of cryptocurrency transactions (See Section IV).

• We provide diverse case studies of how cryptocurrency has been used in the Dark Web. The results include correlation analysis with Surface Web data to disclose other crime activities in which cryptocurrency has been involved (or their related information) and financial flow analysis to track how much money has been transferred to whom/where, which provides an in-depth understanding on the usage of cryptocurrency in the Dark Web (See Section V and VI).

• We identify a real Bitcoin scam value chain and a weapon trading value chain in the Dark Web and provide in-depth analysis information on those chains, and demonstrate the importance of cross-referencing clues from the Dark Web to the Surface Web and vice versa. We believe that our work is the first exemplary investigation revealing the real value chains using cryptocurrency in the Dark Web (See Section VII).

We lastly discuss the ethical considerations in conducting our research, and a possible solution to mitigate illicit trades behind the state-of-the-art anonymity techniques (i.e., anonymous network and cryptocurrency) in Section VIII.

II. Background

A. Bitcoin

Bitcoin [58] is a decentralized digital cryptocurrency that relies on cryptography algorithms and a peer-to-peer network to manage a fully distributed ledger without a central authority.

Unlike the traditional banking system, the absence of a central authority means that financial activities have remained under a pseudonym. Bitcoin users can generate multiple accounts (i.e., public addresses) with corresponding verifiers of the ownership (i.e., private keys) to send/receive bitcoins (BTCs)[2] through a wallet software, which makes a payment as well as manages key pairs. Thus, payments in Bitcoin can be transferred over the Bitcoin network without revealing the real identities of the participants involved in each transaction.

Payment in Bitcoin starts by broadcasting a transaction over the Bitcoin network by Bitcoin users. Suppose that Alice sends BTCs to Bob. Alice's wallet software first searches *unspent transaction outputs* (UTXOs) that contain amounts of BTCs and conditions to spend corresponding BTCs. Each UTXO can be spent on other Bitcoin addresses as an input in a new transaction. If Alice has authentication information (i.e., private keys) to ensure ownership of Bitcoin addresses having valid UTXOs, Alice's wallet software creates a transaction signed by her private keys and broadcasts it over the Bitcoin network. Bitcoin users can transfer arbitrary valid public addresses to receive/send BTCs with other users, but the address reuse is not recommended for privacy and security reasons [32].

After receiving a transaction request, Bitcoin nodes first check whether the requested transaction is cryptographically acceptable (valid) and register the transaction into the Bitcoin Mempool if it is verified. For creating a new Bitcoin block, Bitcoin nodes collate a set of transactions from the Mempool, form them into a block, then perform PoW to solve a

[1]There has been no official definition for "Dark Web", but often referred by the security community and the popular press [65]. In the paper, we use the Dark Web to refer to the Tor anonymous network [70] (i.e., .onion addresses).

[2]The unit of account of the Bitcoin system. We use BTC as a ticker symbol.

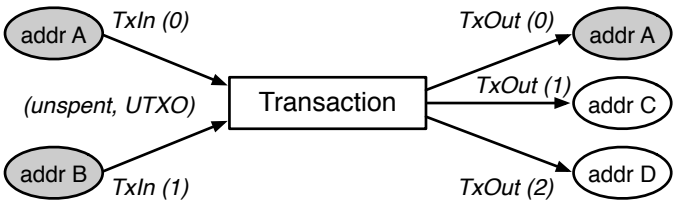

Figure 1: An example of a Bitcoin transaction between Alice and Bob. The gray and white ovals indicate the public Bitcoin addresses owned by Alice and Bob respectively.

mathematical equation, called a *mining process*. If a Bitcoin node solves the math problem and it is verified by other Bitcoin nodes, the new block is finally linked to the Bitcoin Blockchain.

Figure 1 illustrates an example transaction in which Alice sends BTCs to Bob and sends back the remainder of the BTCs to Alice. This Bitcoin transaction consists of a list of inputs ($TxIn$), which are referenced to Alice's public addresses (the gray oval) connected to unspent transaction outputs ($UTXO$), and a list of outputs ($TxOut$) - the destination public addresses belonging to Alice and Bob. In this example, Alice transfers certain BTCs to Bob's public addresses ($addr\ C$ and D). Since the total input value should equal the total output value according to the Bitcoin protocol, Alice sends the rest of the BTCs back to the same address, used in $TxIn(0)$.

B. Dark Web criminal ecosystem

We detail the procedures for how an illegal underground transaction involving the Dark Web and cryptocurrency operates, which consists of five steps: (i) advertisement, (ii) discovery, (iii) negotiation, (iv) payment, and (v) fulfillment.

Advertisement. Advertising illegal products or services on the Dark Web requires different approaches from promoting legal products or services through the Surface Web since traditional search engines do not index content on the Dark Web. If a dark website is created to promote sales, then this information must be registered with a directory service provided on the Dark Web (e.g., a hidden service directory through Tor). This registration is then advertised to potential visitors by posting access information (e.g., onion domains) on the Surface Web (e.g., SNS and forums). An alternative approach is to advertise dark websites on general purpose Dark Web search engines (e.g., Ahmia [1] and Haystak [14]) or market platforms (e.g., Silkroad [22] and Dream Market [11]).

Discovery. Buyers follow similar approaches from the leads of a seller's advertisement strategies, such as discovering entry points to suppliers selling illegal offerings through communities or Dark Web search engines. Also, buyers may share access information with other buyers directly.

Negotiation. To proceed with a transaction, a buyer must confer with a seller about the deal regarding shipping method, price, customizing services, and payment methods. These details vary according to the type of product or service. For example, porn dealers receive money from a buyer and send a passcode for accessing a porn archive. In contrast, hacking service providers might require additional details, such as the type of hacking services requested and general information about targets. Typically, guidelines for information needed are included with the seller's sales information.

Payment. Payment through the Dark Web commonly has the two following options depending on the existence of a third party who mediates transactions between the buyers and sellers. Transacting parties without a third-party mediator make agreements to receive and send fees directly where sellers provide a cryptocurrency address to the buyers for collecting fees. Escrow is available to overcome uncertainty in the credentials of transacting parties since established service providers tend to have a higher reputation. Escrow service providers support an automated payment system to buyers and charge service fees to the sellers.

Fulfillment. As the final step, sellers fulfill orders similar to e-commerce services of the Surface Web by sending physical products via an agreed delivery method (e.g., drugs and weapons), providing online services (e.g., hacking and illegal content) or performing criminal activities in real-world environments (e.g., targeted assassinations).

III. Motivation and Challenges

Anonymity network and cryptocurrency have contributed to protecting the privacy of people seeking anonymity. At the same time, perpetrators have abused these to hide their identity. The Dark Web, a Web environment based on anonymity networks, has been infamous for hosting unlawful content and black markets trading illegal goods [30], [34]. Cryptocurrency also plays a role in concealing the identities of people involved in illegal monetary transactions. The pseudonymous Bitcoin address and the decentralized nature of Bitcoin make it difficult to dissect illicit financial activities [42].

Despite the wide attention on the cryptocurrency and Dark Web from law enforcement and the research community [31], [30], [28], [64], [34], [39], [43], [47], no previous research has conducted a large-scale and in-depth measurement study estimating the quantity or popularity of illicit cryptocurrency transactions on the Dark Web. Hence, it cannot clearly answer the following research questions: *What are the popular illicit goods and wide-spread services on the Dark Web? How much money has been spent on trading illicit goods or services on the Dark Web? Is it feasible to track illicit cryptocurrency money flows? How do perpetrators capitalize their goods and services while minimizing the risk of revealing themselves?*

Answering these questions will advance our understanding of (illicit) financial activities in the underground economy and contribute to finding a new way of preventing illicit cryptocurrency activities. However, finding answers to these questions is not a trivial exercise, due to the following technical challenges.

Limited Dark Web Data Accessibility. Unlike the Surface Web, the content of which can be easily searched and accessed via modern search engines (e.g., Google) with extensive coverage, the Dark Web can only be accessed by using a special software/browser and there is no major search engine substantially covering the Dark Web. Moreover, dark websites are quite volatile since site owners can easily create and change domains without restrictions. Therefore, it is challenging to collect

Category	Count
# .onion domains	36,864
# dark webpages	27,665,572
Period	Jan 2017 ~ Mar 2018 (15 months)

Table I: The statistics of crawled dark webpages and .onion domains.

	BTC	ETH	XMR	Total
# domains	2,886	180	121	3,187
# webpages	1,579,047	4,743	4,410	1,588,200
# extracted addresses	34,265,032	12,138	49,852	34,327,022
# distinct addresses	9,906,129	649	38,440	9,945,218
# preprocessed addresses	5,440	50	61	5,551

Table II: The statistics of cryptocurrency addresses (Bitcoin, Ethereum and Monero) extracted from the Dark Web.

extensive coverage of dark websites and to track changed content over time. Section IV describes how we obtain hidden service addresses for our crawlers and what we implement to improve our coverage on the Dark Web.

Pseudonymity of cryptocurrency. Considering that Bitcoin, the most popular cryptocurrency, is often used for purchasing illegal goods [34], [43], tracking transaction parties of Bitcoin on such unlawful activities may provide clues for examining the underground ecosystem. However, understanding transaction participants in illicit financial activities is still challenging, because Bitcoin transactions are commonly operated under a pseudonym, not revealing the identities of participants. Furthermore, the disposable nature of Bitcoin addresses makes it even harder to ensure that consistent entities are involved in financial activities. Section V explains how we identify pseudonymous financial transaction entities and correlate the external information from the Surface Web to gain more insights to discover their activities.

Obscure cryptocurrency money flows. Tracing the flow of funds in Bitcoin remains a challenge due to the design. Specifically, if a transaction has more than two inputs and outputs, it is hard to determine how much BTCs in each input are sent to which outputs due to the lack of explicit links between inputs and outputs in a transaction. Moreover, anonymization methods for Bitcoin transactions make it more difficult to analyze fund flows of Bitcoin. For example, CoinJoin [53] combines multiple Bitcoin payments from multiple senders into a single transaction as if one user owns all input addresses in the transaction. CoinShuffling [61] actively mixes funds through random-like transactions to prevent fund tracing. We tackle such challenges and track illegal fund operations to understand the characteristics of *black money* flows by employing the concept of *taint analysis*. Section VI describes our methodologies in detail.

IV. Collecting cryptocurrency addresses on the Dark Web

MFScope. To facilitate our large-scale and in-depth study of cryptocurrency usages in the Dark Web, we design and implement MFScope, and its overall architecture and workflow is presented in Figure 2. MFScope consists of two main components; *data collection*, collecting illicit cryptocurrency addresses from the Dark Web and *analysis*, analyzing the cryptocurrency addresses and tracking their illicit money flows. This section mainly focuses on describing the data collection part, and the analysis part will be presented in Section V and VI.

Data Collection Overview. MFScope starts by collecting seed onion addresses from Tor hidden service search engines such as Ahmia [1] and FreshOnions [12]. From the collected seed addresses, MFScope crawls text contents and traverses onion links on visited dark websites until there are no more links to traverse (Section IV-A). From the crawled websites, MFScope extracts cryptocurrency addresses and performs preprocessing to filter out invalid or unnecessary addresses (Section IV-B). We then label whether such collected cryptocurrency addresses are indeed used for selling illegal goods and services (Section IV-C).

A. Crawling the Dark Web

MFScope starts by crawling the Dark Web with 10K onion addresses that we have obtained from two popular Tor hidden service indexing services: Ahmia [1] and FreshOnions [12]. Since they provide the list of indexed onion addresses on the Tor anonymity network, those 10K onion addresses are not biased by our selections of search keywords. For each onion address, crawlers visit its webpages and traverse all of the onion links that appear on the webpages simultaneously. They then extract text information from visited pages and store the information to a distributed database. As shown in Table I, we have collected over 27M distinct webpages from 36,864 distinct onion domains.

B. Extracting cryptocurrency addresses

In our analysis, we choose Bitcoin [58] and Ethereum [72] for their exceptional popularity [17]. Monero [18] is also selected because of its intrinsic support for privacy-conscious users, which may attract perpetrators who seek anonymity for illegal activities [54]. We have extracted Bitcoin, Ethereum, and Monero cryptocurrency addresses from the collected 27M dark webpages via the *Address Extraction* module in MFScope.

This module extracts Bitcoin, Ethereum and Monero addresses with regular expressions. As shown in Table II, the module discovers over 34M cryptocurrency addresses and finally captures about 10M distinct Bitcoin addresses from 2,886 onion domains, 649 distinct Ethereum addresses from 180 domains, and about 38K distinct Monero addresses from 121 domains, respectively.

To accurately analyze cryptocurrency usages in the Dark Web, the Address Extraction module filters out invalid and unnecessary cryptocurrency addresses. It excludes cryptocurrency addresses that appear on dark websites publishing blockchain data such as a mirror site of *Blockchain.com* [7] on the Dark Web, because Blockchain information provides no clue for inferring the illicit intent of their usages. It also filters out invalid addresses that match the regular expressions of each cryptocurrency, but fail to pass a validation check of

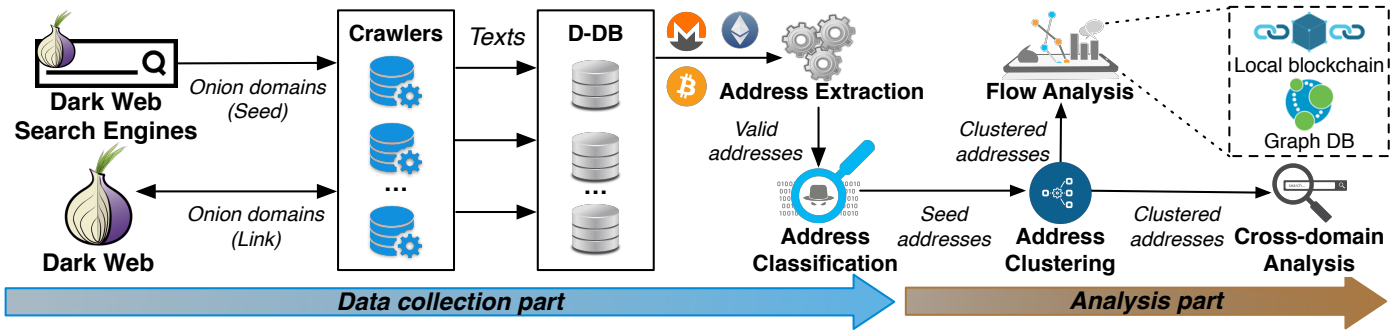

Figure 2: MFScope's workflow; An overview of analysis platform including data collection and analysis modules.

each cryptocurrency client[3]. Lastly, cryptocurrency addresses with no transactions are excluded because these have not yet involved in any financial transactions. However, we are unable to filter Monero addresses with no transactions since Monero transactions are private by design.

After filtering out invalid and unnecessary cryptocurrency addresses, we finally obtain 5,440 Bitcoin, 50 Ethereum, 61 Monero addresses. In particular, over 99.8% of Monero addresses are excluded as garbage strings (e.g., RSS feeds and random strings) and even then it is unknown whether the remaining 61 Monero addresses have been involved in illicit businesses due to its privacy design. In summary, few instances of Ethereum and Monero are discovered compared with Bitcoin. Such trends confirm that Bitcoin is the most popular cryptocurrency on the Dark Web, which leads us to focus on analyzing it. In the rest of the paper, we investigate 5,440 Bitcoin addresses for characterizing cryptocurrency usages in the Dark Web.

C. Classifying illicit Bitcoin addresses

In the *Address Classification* process[4], we manually check whether those 5,440 Bitcoin addresses have indeed been used for trading illicit goods or services by reviewing dark webpages containing each of 5,440 Bitcoin addresses. For instance, if a Bitcoin address is extracted from a drug trafficking site, we are able to consider that the address is used for an illicit purpose, illegal drug trade. In this context, we ask 10 security researchers[5] to classify whether a given Bitcoin address is used for illicit purposes. We ask them to review all webpages where any of the 5,440 Bitcoin addresses appear. Because one Bitcoin address may appear at numerous webpages, we review up to 20 pages for each address. We specifically ask them the following question: *"Do any of the dark webpages containing a Bitcoin address promote trading illicit goods or services?"* For cross-checking, we make each participant label every 5,440 Bitcoin address.

We initially classify 5,440 Bitcoin addresses into the two categories: benign and potentially illicit addresses. We consider a given Bitcoin address as *potentially illicit* if at least two out of ten participants label it as *illicit*. We find that 4,556 (83.75%) addresses are used for trading potentially illicit goods or services and the remaining 884 (16.25%) addresses are benign, as described in Table III.

Category	Count	Ratio (%)
Potential illicit addresses	4,556	83.75%
Legitimate addresses	884	16.25%
Total	5,440	100.00%

Table III: Cryptocurrency distribution over the Dark Web.

We further classify the 884 legitimate Bitcoin addresses in Table IV (a) by asking the following question: *"What is the usage for a legitimate address based on the content where the address appears?"* We classify addresses largely into seven categories. Most of the addresses are discovered in the contents of requesting donations, sharing knowledge, verifying escrow, identifying users, advertising products, and providing legal services.

Among 4,556 potentially illicit addresses, we conservatively pick illicit Bitcoin addresses that more than seven out of ten researchers (70%) mark as *illicit*. The remaining addresses with less than seven votes are referred to as *possible illicit addresses*. We categorize the 4,471 possible illicit Bitcoin addresses into two cases: *Proof* and *Unidentified*, in Table IV (b). The addresses belonging to *Proof* are found on the Ponzi scam sites, which ask victims to invest in cryptocurrencies. To misplace victims' trust that these websites actually return the invested cryptocurrencies, they show valid but unrelated Bitcoin transactions. Illegal keywords include the webpages where the addresses in the *Unidentified* category appear. These websites lack the conclusive proof to decide whether their services are illegal because of their complicated business models and no direct sales of illicit goods.

We classify 85 Bitcoin addresses (Table IV (c)) as *illicit*. We are aware that the law or moral standards for deciding unlawful goods depends on the national and cultural backgrounds of participants. To help participants objectively reason their choices, we ask the participants the second question: *"What goods do they commerce via the illicit Bitcoin addresses?"* We create nine good/service categories and use these categories for the survey. Besides the nine categories, we also create the *Other* category to include content that is hard to be classified into nine categories, such as selling chips for online

[3]We used `bcoin` [2], `geth` [41], `XMR Tools` [52] for Bitcoin, Ethereum, and Monero respectively.

[4]In MFScope, the address classification is a manual task.

[5]They are quite familiar with the Dark Web and cryptocurrency and have at least more than 2 years' experience in conducting security research.

(a) Legitimate addresses

Category	Description	Count
Donation	Donation requests	277
Escrow	Escrow services	11
Informative	Information delivery content (e.g., cryptocurrency wallet guide)	343
Identification	Contents of own information	60
Product	General goods (e.g., CD, sportswear, etc.)	14
Service	Legitimate services (e.g., hosting, website selling, etc.)	21
Unidentified	Non-existence of illegal keywords and goods name, and not understandable context	158
Total		884

(b) Possible illicit addresses

Category	Description	Count
Proof	Contents for proof of transaction (e.g., Ponzi scheme)	4,171
Unidentified	Existence of illegal keywords, but non-existence of goods name or not understandable context	300
Total		4,471

(c) Illicit addresses

Category	Products	Count
Abuse	Illegal sexual and violent content	15
Account selling	Hacked social accounts	6
Counterfeit	Counterfeit money	6
Card dumps	Dumped credit cards	2
Drug	Illegal drugs	4
Investment	Financial investment options	29
Membership	Membership to join private illegal forums	8
Service	Illegal errand services (e.g., hacking and contract killings)	8
Weapon	Unauthorized weapons	1
Others	Etc.	6
Total		85

Table IV: Cryptocurrency usage over the Dark Web

gambling and offering technical tutorials for hacking. One Bitcoin address in *Others* is discovered in 774 dark websites hosted on Freedom Hosting II, which have been compromised by a hacker group [33], and this address is left on the sites as a deposit account. Such exceptional cases are categorized as *Others*.

We call these 85 addresses *seed Bitcoin addresses*. They serve as the ground-truth–denoting that these addresses have been used for commerce in illicit goods or services. In Section V, we leverage these seed addresses to disclose more Bitcoin addresses directly involved in circulating illicitly earned cryptocurrencies.

V. Demystifying Bitcoin Ownerships

The seed Bitcoin addresses have been exposed on the Dark Web with explicit evidence of their involvement in illicit activities. Perpetrators on the Dark Web are the owners of these seed addresses, and it is highly likely that each perpetrator has other Bitcoin addresses that have yet to be exposed. The *Address Clustering* module in MFScope discovers more illicit Bitcoin addresses that the perpetrators have owned. The module clusters Bitcoin addresses by leveraging their ownership.

Information from the Dark Web exposes a tip of the iceberg of perpetrators' illegal activities since they may conceal evidence that possibly reveals their entire illegal businesses or themselves. Based on information from the Dark Web sorely, it is difficult to grasp the entire schemes of their activities. To gain more information, our *Cross-domain Analysis* module performs correlational analysis with the Surface Web to obtain additional information associated with the illicit Bitcoin addresses that perpetrators have owned.

A. Clustering illicit Bitcoin addresses

The pseudonymous nature of Bitcoin addresses hinders the inference of the explicit ownership of different Bitcoin addresses. Despite this, there exist several heuristics that can determine ownership by analyzing how Bitcoin addresses have been controlled [55], [40], [27]. The heuristics from the previous studies infer ownership based on (i) multi-input transactions and (ii) change addresses. Our *Address Clustering* module leverages BlockSci [48], which is a Bitcoin analysis platform that implements both of the heuristics for tracing address ownership.

A multi-input (MI) transaction is a Bitcoin transaction in which multiple input addresses are involved. It is possible to infer that the input addresses in a multi-input transaction are owned by a single entity because one must present all the private keys associated with the input addresses to make such a transaction. However, in the case of CoinJoin transactions [53], although they involve MI transactions, the input addresses are not necessarily owned by a single entity, and therefore, such transactions must be excluded at the time of clustering.

Our *Address Clustering* module groups Bitcoin addresses based on *MI* heuristic. Using BlockSci, which also implements an algorithm [44] that can detect CoinJoin transactions, we discover one CoinJoin transaction out of 3,726 transactions that have at least one of the 85 illicit seed addresses as inputs. Excluding the CoinJoin transaction in the process of clustering, we discover 3,029 additional Bitcoin addresses, which the perpetrators have owned along with the 85 seed addresses.

When classifying input addresses of MI transactions into the same owner cluster, we also take *change addresses (CA)* into account (MI+CA). Since the standard Bitcoin mechanism requires all inputs in a transaction spend all BTCs, wallet software generates a new Bitcoin address to receive the remainder of BTCs after sending the specified amount to the intended address. This newly generated Bitcoin address is referred to as a *change address*, and since it represents the sender's new Bitcoin address, the change address belongs to the owner of the addresses used as the transaction inputs.

The *Address Clustering* module clusters Bitcoin addresses on the basis of ownership by tracing both MI transactions and CAs (MI+CA) and assigns a unique identifier to each cluster (i.e., cluster ID). As Meiklejohn *et al.* stated in [55], falsely identified change addresses may produce large clusters with too many addresses, which results in many false positives. Most of the clusters including each seed Bitcoin address have

Category	# domains (# pages)	Heuristics	# addrs	# transactions In	Out	Total	Market volume BTC (USD) received	BTC (USD) sent	Lifetime TX_{first}-TX_{last}
	33 (76)	Seed	15	673	277	950	26.66 ($41,396)	25.88 ($39,625)	19/03/2015-30/04/2018
Abuse		MI-only	486	9,797	8,580	18,377	3,416.43 ($3,862,983)	3,416.42 ($3,863,185)	19/03/2015-30/04/2018
		MI+CA	539	2,900	1,154	4,054	106.92 ($92,747)	106.13 ($99,853)	17/10/2013-30/04/2018
	11 (56)	Seed	6	28	24	52	1.14 ($741)	1.13 ($1,050)	30/03/2016-24/12/2017
Account selling		MI-only	60	91	83	174	2.01 ($1,811)	2.01 ($2,298)	30/03/2016-24/12/2017
		MI+CA	201	326	294	620	10.60 ($8,949)	10.57 ($16,318)	17/10/2013-30/04/2018
	6 (11)	Seed	6	19	14	33	0.92 ($1,174)	0.92 ($1,195)	26/09/2016-14/02/2018
Card dumps		MI-only	205	4,658	4,458	9,116	2,323.40 ($9,935,313)	2,323.37 ($9,938,336)	17/11/2014-30/04/2018
		MI+CA	833	1,916	1,651	3,567	279.13 ($179,444)	279.11 ($181,709)	17/10/2013-30/04/2018
	3 (3)	Seed	2	8	7	15	0.47 ($511)	0.47 ($534)	18/03/2017-05/07/2017
Counterfeit		MI-only	23	24	24	48	0.49 ($1,129)	0.49 ($1,142)	25/02/2017-05/07/2017
		MI+CA	27	35	33	68	1.01 ($1,701)	1.01 ($1,736)	09/07/2014-30/04/2018
	5 (283)	Seed	4	46	25	71	3.95 ($1,902)	3.95 ($1,923)	06/11/2015-01/03/2017
Drug		MI-only	18	2,509	1,289	3,798	5,245.93 ($14,124,499)	5,245.92 ($14,373,916)	19/07/2014-13/12/2017
		MI+CA	26	1,875	673	2,548	119.92 ($57,867)	119.92 ($58,086)	17/10/2013-30/04/2018
	475 (1,726)	Seed	29	2,258	396	2,654	75.25 ($117,995)	74.79 ($123,486)	21/04/2015-30/04/2018
Investment		MI-only	2,025	93,479	80,026	173,505	32,428.20 ($151,438,331)	32,421.22 ($151,816,053)	04/09/2013-30/04/2018
		MI+CA	204	4,733	918	5,651	188.19 ($211,574)	184.87 ($203,175)	17/10/2013-30/04/2018
	14 (835)	Seed	8	50	38	88	4.43 ($11,441)	4.43 ($13,573)	07/01/2017-20/01/2018
Membership		MI-only	95	504	265	769	29.20 ($85,481)	29.20 ($92,185)	14/11/2016-23/01/2018
		MI+CA	247	769	473	1,242	41.64 ($127,788)	41.64 ($137,228)	17/10/2013-30/04/2018
	9 (74)	Seed	8	30	24	54	6.14 ($5,065)	6.14 ($4,898)	12/01/2015-02/04/2018
Service		MI-only	113	547	432	979	59.39 ($60,141)	59.38 ($59,206)	18/07/2014-29/04/2018
		MI+CA	861	2,083	1,774	3,857	308.77 ($208,761)	308.75 ($211,130)	17/10/2013-30/04/2018
	1 (119)	Seed	1	5	5	10	1.42 ($3,995)	1.42 ($3,820)	20/01/2017-26/04/2018
Weapon		MI-only	42	362	264	626	46.37 ($32,964)	46.35 ($33,028)	18/07/2014-29/04/2018
		MI+CA	754	1,828	1,568	3,396	277.47 ($173,385)	277.46 ($173,782)	09/07/2014-30/04/2018
	786 (1,330)	Seed	6	609	177	786	40.66 ($20,924)	40.66 ($23,211)	14/07/2015-03/01/2018
Others		MI-only	9	1,187	409	1,596	65.91 ($32,043)	65.91 ($32,434)	14/07/2015-03/01/2018
		MI+CA	22	1,968	679	2,647	119.50 ($59,732)	119.50 ($62,463)	17/10/2013-30/04/2018
	1,343 (4,513)	Seed	85	3,726	987	4,713	161.05 ($205,144)	159.80 ($213,314)	12/01/2015-30/04/2018
Total		MI-only	3,029	110,664	94,105	204,769	43,422.64 ($179,317,131)	43,415.62 ($179,954,158)	04/09/2013-30/04/2018
		MI+CA	2,044	12,676	5,208	17,884	776.58 ($712,862)	772.44 ($728,380)	17/10/2013-30/04/2018

Table V: Bitcoin usages and their volumes of perpetrators who trade illicit goods or services on the Dark Web.

at most 1k Bitcoin addresses, but several clusters possess over 350k Bitcoin addresses each. Thus, we exclude such large clusters to avoid false positives. After excluding the large clusters, we discover 2,044 Bitcoin addresses that belong to the owners of the illicit seed addresses, as shown in Table V (Total, MI+CA).

Table V shows the number of illicit seed addresses observed on the Dark Web for each category along with the newly discovered addresses by the ownership tracing heuristic used of multi-input transactions (MI-only) and multi-input transactions and change addresses (MI+CA). Table V also shows how many in and out Bitcoin transactions were made and how much money was sent and received for each category of illicit businesses up until April 2018. The money transferred in USD is calculated based on the market price at the time of each transaction. The total Bitcoin dealt by perpetrators on the Dark Web is approximately 43K BTCs or around 180M in USD.

Of the different categories of illicit businesses operated by the perpetrators on the Dark Web, we observe that *investment* (e.g., Ponzi fraud) is the largest business category in terms of market volume (around 150 million USD). Previously, Massimo *et al.* [29] showed that about 10M dollars in USD have been deposited to 1,211 Bitcoin addresses for Ponzi schemes by analyzing the Bitcoin addresses posted on Bitcoin forums on the Surface Web. Compared to the results shown in the previous study, our collection of 2,258 Bitcoin addresses in the *investment* category and the market volume of 150 million USD are quite extensive. *Drug* and *card dumps* categories have also been relatively active compared to the rest of the illicit business categories on the Dark Web. The market volumes of these categories are 14.4M and 10M USD, respectively.

Category	Seed	MI	MI+CA	Total
Tor proxy	25 (38)	28 (38)	38 (45)	91 (121)
Community	31 (35)	38 (59)	16 (20)	85 (114)
Sales	11 (17)	20 (27)	6 (9)	37 (53)
Media	10 (10)	10 (17)	5 (5)	25 (32)
Archive	3 (4)	7 (12)	2 (6)	12 (22)
Miscellaneous	1 (1)	3 (3)	3 (4)	7 (8)
Unavailable	7 (8)	12 (17)	5 (6)	24 (31)
Total	88 (113)	118 (173)	75 (95)	281 (381)

Table VI: An overview of the number of domains per each category. The values in parentheses indicate the number of webpages.

By analyzing the up-to-date dataset of the Dark Web, we demonstrate recent trends in the illicit businesses on the Dark Web. Considering that its market volume is approximately 180M USD and they are ongoing businesses, this indeed is a concern for the world.

B. Cross-referencing illicit Bitcoin addresses to the Surface Web

By employing two different Bitcoin ownership heuristics (MI-only and MI+CA), we obtain the additional thousands of illicit Bitcoin addresses that the perpetrators on the Dark Web have owned; however, unlike the seed addresses, we do not have any contextual information about these new addresses. In addition, for the seed addresses, we do not know how the perpetrators have used them on the Surface Web.

Accordingly, to gain more insights about the usage of the discovered illicit Bitcoin addresses on the Surface Web, we perform a cross-domain analysis on each one of the illicit Bitcoin addresses.

The *Cross-domain Analysis* module in MFScope conducts a Google search by using the illicit addresses from the Address Clustering module as keywords and publishes search results to a database. While searching, it excludes Blockchain information sites publishing Blockchain data that is mostly out of our concern. Table VI shows the number of surface websites where any of seed, MI-only, MI+CA addresses appear. We collect 381 webpages that include the illicit Bitcoin addresses from 281 websites.

To understand the usage of the illicit Bitcoin addresses on the Surface Web, we manually investigate these webpages and classify them into seven categories, as shown in Table VI. The *Tor proxy* category includes a set of dark webpages exposed to the Surface Web search engines via Tor proxy services such as Tor2Web [66]. Such proxy services allow users to access .onion domains without a Tor browser. The search results in this category complement missing webpages of our Dark Web dataset: the webpages often contain sensitive information about the perpetrators, such as usernames, personal interests, etc. The *Community* category represents Internet forums, where anyone can access and share diverse information. The addresses appearing in community posts are mostly mentioned by third parties who do not own the addresses. Most posts are complaints (fraud reports), feedback or user reviews on the illicit businesses associated with the illicit Bitcoin addresses. Such information reassures that the illicit Bitcoin addresses have been actively used for illicit businesses and many victims have not satisfied goods and services from the perpetrators. The pages in the *Sales* and *Miscellaneous* categories contain a variety of content such as hyperlinks to perpetrator's other illicit businesses (e.g., investment sites), religious propensities, and real-world identities. Finally, the webpages categorized as *unavailable* are no longer accessible due to the dead links.

Performing the cross-domain analysis on the illicit Bitcoin addresses gives us crucial pieces of information including (but not limited to) 1) even more relevant Bitcoin addresses, 2) owner profiles including a forum account ID leading to a personal hacking blog, 3) threads in forums, which help to guess the user's physical location, 4) mail- and web-servers' information including their location and user information, and 5) other fraud campaigns such as hacking services and investment scam sites. Such information is useful to understand the perpetrators or their activities and helps us to reveal illegal value chains in Section VII.

VI. Tracing black money

This section introduces the *Flow Analysis* module that performs *taint-based Bitcoin flow analysis*, which is designed to 1) trace the money that flows from illicit Bitcoin addresses to their destination Bitcoin addresses and 2) quantify how much money flows to the destinations. The characteristics of the illicit money flow revealed in this study are then analyzed.

A. Building transaction graphs

In order to trace the money flows from the illicit Bitcoin addresses to their destinations, the Flow Analysis module constructs a Bitcoin transaction graph for each illicit address based on the Bitcoin transaction information retrieved from the blockchain.

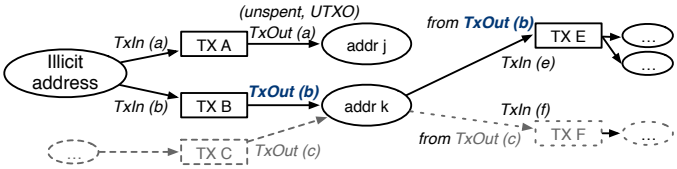

Figure 3: An example transaction graph starting from an illicit Bitcoin address. Solid lines indicate *inputs* and *outputs* involved in transferring coins from the illicit Bitcoin address. Dot lines are *inputs* and *outputs*, which are not involved in illicit money flows and will not be traced.

Our Bitcoin transaction graph is a rooted directed graph with an illicit Bitcoin address as a root node. As shown in Figure 3, for a given illicit Bitcoin address, the Flow Analysis module creates a root $addr$ node and finds the transactions having the illicit address as an input address. It then creates the transaction (TX) nodes for each of the transactions and adds directed $TxIn$ edges from the root node to the TX nodes. For each TX node, it also creates $addr$ nodes for its output addresses and connects them to the TX node with $TxOut$ edges pointing towards $addr$ nodes. For all the edges, it labels the amount of Bitcoins transferred, and particularly for $TxOut$ edges, it additionally labels the edge with a $UTXO$[6] tag only if the transaction output has not been spent.

Again, the Flow Analysis module starts following subsequent transactions, *next transactions*. *Next transaction* is a subsequent transaction (t) in which an output ($TxOut$) of the current transaction (t') is spent as an input ($TxIn$) of the next transaction (t). For example, as illustrated in Figure 3, although there are two transactions ($TX\ E$ and $TX\ F$) taking $addr\ k$ as their inputs, the module only follows $TX\ E$ as a next transaction, because it spends the output $TxOut(b)$, which is originated from the illicit address, as its input $TxIn(e)$. TXF is abandoned because the Bitcoins processed in this transaction are from $TXOut(c)$, which does not carry the Bitcoins from the illicit address.

As described above, starting from an illicit address (root) node, the Flow Analysis module first traverses all transactions having the illicit address as input and adds the recipient addresses to the graph. While traversing subsequent transactions, it repeatedly appends TX and $addr$ nodes to the graph until the last address node on each path from the root node is identified by a $TxOut$ edge with a $UTXO$ tag.

Furthermore, when the Flow Analysis module adds a new $addr$ node to the graph, it attempts to identify the owner of the address by querying *WalletExplorer* [24], which provides ownership information about Bitcoin addresses. If the address is owned by a well-known service provider, it labels the node with the service's name and stops following the next transactions because it reaches the real world destination of this particular money flow. Table VII enlists the categories of the destination services observed.

[6]UTXO stands for unspent transaction output.

Category	Description
Exchange	A digital marketplace to buy, sell and exchange coins, or provide wallet services.
Gambling	An online gambling site (e.g., Poker, casino, etc.).
Market	A marketplace to sell and buy illegal products.
Mixing	A service to shuffle coins to improve anonymity.
Others	Other services (e.g., faucet, legal market, pools, etc.)

Table VII: Categories of the illicit financial flow destinations.

The Flow Analysis module builds Bitcoin transaction graphs for each illicit (MI-only) Bitcoin address obtained from the previous section[7]. However, there were several cases where the money flows[8] from the illicit root nodes of the graphs are too long. We limit the Flow Analysis module to stop tracing each money flow that lasts more than 10 transactions. It then labels the last $addr$ nodes of the flows as *unidentified*. The max length of a money flow is a tunable parameter of MFScope. We choose 10 transactions to build a transaction graph on a Linux workstation with an Intel Xeon E5-2620 2.40GHz CPU and 128 GB of RAM.

B. Quantifying illicit financial flows

For each Bitcoin transaction graph constructed above, we perform the *taint-based Bitcoin flow analysis* to quantify the illicit money flows. Our analysis is inspired by the taint analysis service that *Blockchain.com* offered in the past [57], [56]. They focused on identifying how much BTCs come to a given destination address and what the source addresses sending the BTCs are, whereas our *taint analysis* models *how much BTCs* flows into each destination Bitcoin address from a given Bitcoin address.

We emphasize that quantifying the transferred Bitcoin volume from one address to another is a key factor for tracking illicit money flows. Since perpetrators are able to diversify their money flow paths to transfer the money from one address to another, auditors could be overwhelmed by the vast volume of spurious money flows (e.g., Bitcoin laundry service [37]). Identifying the transferred Bitcoin volume helps the auditors to prioritize money flows in investigation. Furthermore, our analysis identifies how much portion of the money is aggregated at specific Bitcoin addresses from a given Bitcoin address. Such information is a key clue in identifying points where diversified funds are integrated or money is exchanged.

$$taint_{b,t} = \sum_{j} \prod_{pt \in N_j^{t,b}} \frac{output_{pt,next}}{\sum_i output_{pt,i}} \quad (1)$$

$$ratio_t = \frac{\sum_i input_{t,i}}{\sum_{k \in T} \sum_i input_{k,i}} \quad (2)$$

$$taint_b = \sum_{t \in T} ratio_t * taint_{b,t} \quad (3)$$

[7]We exclude Bitcoin addresses clustered by the change-address heuristic to avoid the possibility of falsely linking change addresses, which can create a cluster of Bitcoin addresses that are not controlled by a single entity [55].

[8]The paths from the root node to end $addr$ nodes.

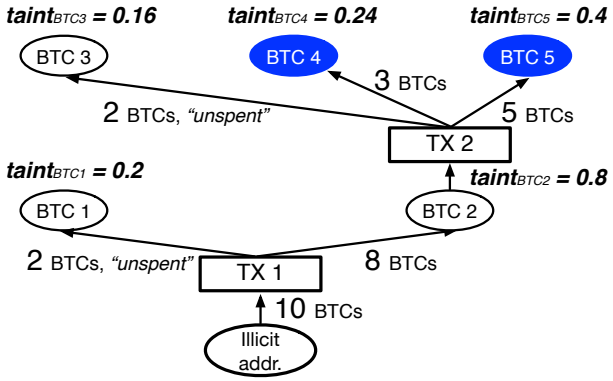

Figure 4: An example of taint-based flow analysis for the destination Bitcoin addresses (leaf nodes) starting with a given illicit Bitcoin address. The blue address nodes are owned by a Bitcoin exchange service.

We define *taint* to be the percentage of transferred BTCs from an input Bitcoin address to each destination Bitcoin address. With the withdrawal transactions T in which one or more unspent transaction outputs (UTXO) linked to the input Bitcoin address a are spent, we calculate the taint value $taint_{b,t}$ for each withdrawal transaction t in T with Equation (1). $N_j^{t,b}$ is the jth set of transactions including the withdrawal transaction t and the next transactions[9] reaching the destination Bitcoin address b, and pt is a transaction in $N_j^{t,b}$. $output_{pt,i}$ is the BTC amount of an output index i of the transaction (pt), and $output_{pt,next}$ is the BTC amount of the subsequent output index spent in the following next transaction in $N_j^{t,b}$.

To reflect the ratio of the flow of funds from the input Bitcoin address a to each withdrawal transaction, we define $ratio_t$, as a normalization function representing the portion of the sum of input values of a in a transaction t divided by the sum of input values of a in all the withdrawal transactions in T, as described in Equation (2). For all transaction inputs with the Bitcoin address a, $input_{t,i}$ is the BTC amount of the input index i in a transaction t, and $input_{k,i}$ is the BTC amount of an input index i of a transaction k for all withdrawal transactions in T. Finally, the final taint value, $taint_b$, is obtained by multiplying the sum of the values of $taint_{b,t}$ with $ratio_t$ for each withdrawal transaction t in T, as described in Equation (3).

Figure 4 illustrates the computation of the taint values of each destination Bitcoin address in a transaction graph. This example assumes one input is held by the illicit Bitcoin address and is involved in a withdrawal transaction, $TX1$, while the leaf Bitcoin address nodes are destinations. As shown, the amount of illicit funds in the graph is 10 BTC, and, via $TX1$, the funds are transferred to two other Bitcoin addresses, $BTC1$ and $BTC2$, on the graph. Here, 20% of the illicit funds and the remaining 80% are transferred to $BTC1$ and $BTC2$, respectively. The final taint value of $BTC1$ is 20% because its output in $TX1$ has the $UTXO$ tag. By traversing $TX2$,

[9]To obtain each transaction set, we compute the paths in a transaction graph from a withdrawal TX node to a destination $addr$ node by traversing the next transactions.

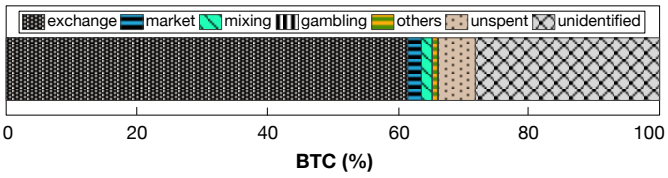

Figure 5: Distribution of the illicit Bitcoins flown into different service categories.

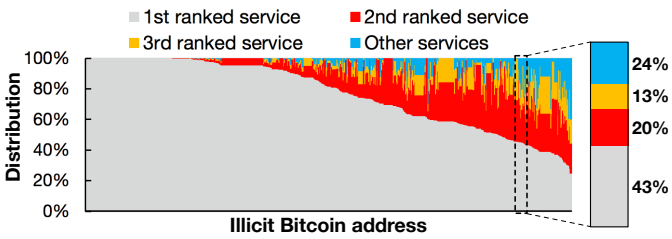

Figure 6: A bar chart for service usage of the illicit Bitcoin address. X-axis represents a Bitcoin address, Y-axis shows distribution of the illicit money transferred to the services that have been identified. A dotted box magnifies one of the bars as shown on the right.

the final taint values based on the fraction of the total output amount transferred to each address are 16% ($0.8 \cdot 0.2 = 0.16$), 24% ($0.8 \cdot 0.3 = 0.24$), and 40% ($0.8 \cdot 0.5 = 0.4$) for $BTC3$, $BTC4$, and $BTC5$, respectively. The Flow Analysis module cannot follow additional next transactions because the outputs toward $BTC1$ and $BTC3$ in $TX1$ and $TX2$ remain unspent and $BTC4$ and $BTC5$ are owned by a well-known Bitcoin service provider (i.e., an exchange service).

In summary, we finally estimate that 36% and 64% of 10 illicit BTCs are either remained *unspent* and transferred into the addresses in the *exchange* category respectively. In spite of the pseudonymity of Bitcoin, the taint-based Bitcoin flow analysis helps to identify how much illicit funds have flown into. In the rest of this section, we characterize illicit Bitcoin addresses on the basis of our taint analysis.

C. Service usage characteristics of the perpetrators

Using the Bitcoin transaction graphs labeled with taint values, we analyze the illicit financial flows and investigate their service usage characteristics. Note that the transaction graphs are built for the illicit Bitcoin addresses of the perpetrators' clusters, classified using the MI-only heuristic from Section V-A.

Categorical popularity of services used by the perpetrators: In order to understand the usage of the illicit funds that the perpetrators on the Dark Web have earned, we measure how many of their Bitcoins have been transferred to different service categories (Table VII) in total, as shown in Figure 5. About 61.4% of the total illicit funds have been deposited into exchange services, while only a small portion of the funds have been transferred to mixing services. This implies that the perpetrators have exchanged more than the half their

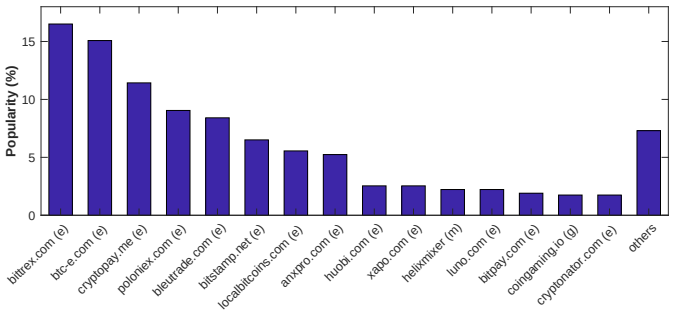

Figure 7: The top 15 most popular Bitcoin services in the financial activities of the perpetrators.

Bitcoins for cash or alt-coins, rather than going through even more complex money laundering processes. Furthermore, only a small amount of the total illicit funds have been sent to black markets. This indicates that the perpetrators spent unlawfully earned cryptocurrencies to purchase illegal goods or services.

Service usage characteristic: We also investigate how much of the Bitcoins the perpetrators have used for different services. Figure 6 illustrates the distribution of the perpetrators' Bitcoin expenditure on different services, and each vertical bar (the dotted box) of the stacked bar chart shows the distribution of the Bitcoins transferred from one illicit Bitcoin address to different services — the gray inner bar represents a service that received the most Bitcoins, red the second most, yellow the third most, and blue for the rest.

As shown in Figure 6, most of the illicit Bitcoin addresses have spent the most Bitcoins on one service. To be more specific, about 84 percent of the illicit Bitcoins have transferred more than 50 percent of their funds to one particular service. This implies that *the perpetrators on the Dark Web tend to transfer a large sum of their money to one particular service rather than diversifying their expenditure*. As a side note, 82 percent of the illicit Bitcoin addresses have sent more than 90 percent of their funds to their top three services.

Popularity of services primarily used by the perpetrators: As we learn from the above, the perpetrators on the Dark Web have used different services, and they tend to transfer more money to one particular service than the other services. Such a service that has received the most Bitcoins from a perpetrator can be understood as the primary or the most preferred service that the perpetrator has used. Therefore, to measure the popularity of services used by the perpetrators, we count the number of the primary destination services of the illicit Bitcoin addresses.

We find 126 distinct primary services and Figure 7 depicts the popularity of the top 15, which account for 93% of the population. Further analysis of these 15 services ascertains their popularity among the perpetrators. *Bittrex* [5] is a long-standing company based in the U.S. *BTC-e* [8][10] is a popular coin exchange with headquarters in Russia. *BTC-e* was seized by the U.S. Justice Department on July 26, 2017 because of alleged money laundering, including the hacking of *Mt.*

[10]The front page of the main domain (btc-e.com) indicates that the domain has been seized.

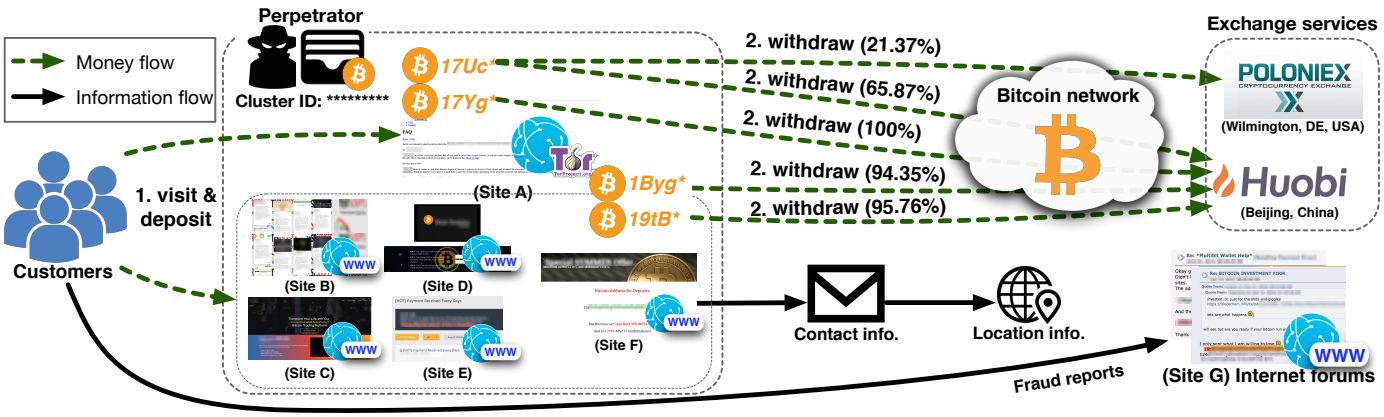

Figure 8: An investment fraud scheme discovered in our analysis: The green arrows illustrate the flow of Bitcoins and the black arrows show how each piece of information was revealed.

Gox [67]. The cryptocurrency exchange services, *Localbitcoins* [16] and *Poloniex* [21], have not enforced the KYC verification until recently [36], [35]. *Helixmixer* [56] is a coin shuffling service, *Coingaming* [9] operates several online gambling sites [3], [23], and *Bleutrade* [6] is a cryptocurrency exchange.

Of the Bitcoin addresses in the Card Dump category in Table V, approximately 44.5% have about 20% taint values to this service equally, and these addresses have the same cluster ID. For the Counterfeit, Dump, and Account-selling categories, 28% of Bitcoin addresses sent their money to *Cryptopay* [10] primarily, and 16.5% of the addresses in these categories are owned by one person who deposited 99% of BTCs to this service. Such trends again show that the perpetrators tend to deposit a large portion of their funds into one Bitcoin service.

VII. Case Studies and Discussion

We present two illegal value chains unveiled through our analysis (Section VII-A and VII-B) and characterize the 85 illegal seed addresses on the Dark Web to unveil hidden financial hubs (Section VII-C).

A. Bitcoin investment scam

One case of cryptocurrency abuse on the Dark Web is a Bitcoin investment fraud. This fraud scheme is particularly interesting, because the perpetrator has (i) leveraged multiple channels (six dark and surface websites) to lure the victims and (ii) transferred most of the embezzled Bitcoins to two Bitcoin exchanges.

As illustrated in Figure 8, the perpetrator has been hosting a Bitcoin investment site on the Dark Web (Site A). This type of website posts their Bitcoin addresses and lures the visitors into investing their BTCs for big returns. During our data collection period (15 months), Site A has updated their Bitcoin deposit address once, and hence, we are able to capture two different Bitcoin addresses (17Uc* and 17Yg*) that belong to the perpetrator. Using these two Bitcoin addresses as seed addresses, MFScope's Address Clustering module (Section V) further discovers two more Bitcoin addresses (1Byg* and 19tB*) that the perpetrator owns; all four Bitcoin addresses, including the seed addresses, belong to a single cluster.

Our system then performs cross-domain analysis (Section V-B) to discover any other footprints that the perpetrator has left on the Surface Web, and it notes that the perpetrator has been operating at least five other Bitcoin investment sites (Site B, C, D, E, and F), which look completely different from each other, on the Surface Web. In addition, from a Surface Web forum (Site G), multiple fraud reports were also discovered, assuring us that those sites have actually been fraudulent by specifically mentioning the perpetrator's Bitcoin addresses. Knowing that the perpetrator has been using four different Bitcoin addresses on six different websites for their Bitcoin scam business, we are able to further trace and gain insights into the financial activities of the perpetrator using MFScope. It performs the taint-based flow analysis as described in Section VI and determines what the perpetrator has done to the embezzled BTCs.

As shown in Figure 8, we learn that the majority of the perpetrator's Bitcoins have been transferred to two different Bitcoin exchanges, Poloniex and Huobi. According to our analysis, about 21% of the BTCs from the Bitcoin address 17Uc* have been transferred to Poloniex and about 66% to Huobi. In the case of the other three Bitcoin addresses (17Yg*, 1Byg* and 19tB*), most (100, 94 and 96 percent) of the BTCs have been transferred to Huobi. The fact that the perpetrator has transferred most of the unlawfully earned BTCs to Bitcoin exchanges is a crucial piece of information, because this implies that they have cashed out the Bitcoins, and thus those exchanges will help investigators detect perpetrators if they follow KYC (know your customer) policy [63].

In addition, our Cross-domain Analysis module also reveals direct information that may lead to the perpetrator. One of the e-mail addresses posted as contact information on the investment site (Site F) was associated with a personal mail server, which leads us to the perpetrator's personal information (e.g., SNS account and magazine subscription receipt with a full name and a billing address). However, to comply with the ethical research standards, we stop our analysis at this stage.

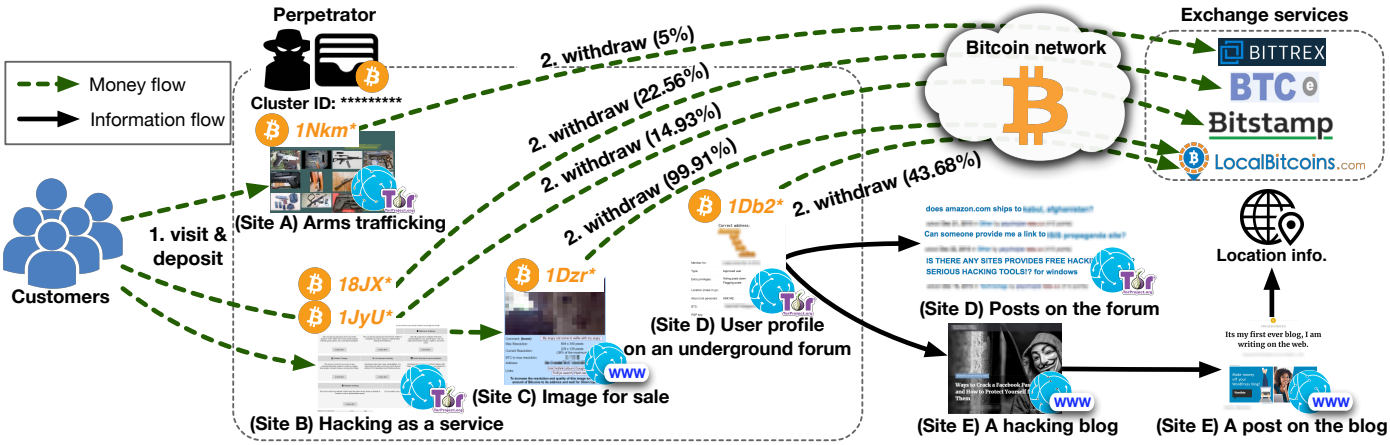

Figure 9: A trafficking scheme discovered in our analysis: The green arrows illustrate the flow of Bitcoins and the black arrows show how each piece of information was revealed.

B. *Trafficking*

Another interesting crime scheme involves the trafficking in firearms and provision of hacking services, as illustrated in Figure 9. Unlike the perpetrator from the previous scheme, this one leverages two completely different dark websites with different Bitcoin deposit addresses; one is a firearms trafficking site (Site A) and the other is a hacking service trafficking site (Site B).

At a glimpse, these sites look like they are owned by two different entities; however, MFScope concludes that they are operated by a single perpetrator. As shown in Figure 9, Site A has been using a Bitcoin address starting with "1Nkm*", and Site B has been using two Bitcoin addresses, each starting with "18JX*" and "1JyU*". MFScope has analyzed these seed addresses with our clustering method (described in Section V-A) and determined that five Bitcoin addresses (1Nkm*, 18JX* 1JyU*, 1Dzr* and 1Db2*) including the three seed addresses belong to a single cluster, implying that Site A and B are operated by the same perpetrator, as illustrated in Figure 9.

Knowing that the perpetrator owns at least these five Bitcoin addresses, MFScope discovers other traces of the perpetrator on the Surface Web, as described in Section V-B. The traces include the fact that one of the Bitcoin address (1Dzr*) and the other address (1Db2*) are mentioned on Site C (sell images for Bitcoins) and D (question-and-answer site), respectively. On Site C, the perpetrator has posted an image for sale with 1Dzr*, and on Site D, he has used 1Db2* as a Bitcoin address for his user profile.

In the case of Site D, although it is a dark website, it has been exposed to Google search engine via a Tor proxy service and thus detected by the Cross-domain Analysis module. This site, a question-and-answer website, allows us to grasp the perpetrator's interests and activities. For example, he asks the following questions: *if Amazon would ship **** to **** (anonymized)*, *how to contact ****[11] (anonymized)*, and *if there is a popular dark website that sells hacking tools*. The username that the perpetrator has been using on Site D has further led us to a personal blog site (Site E) about unethical hacking, and one of the posts has been geotagged to the location that also appeared in the perpetrator's questions asked on Site D.

In addition to the information derived from the investigative analysis performed above, MFScope also investigates the perpetrator's financial activities by performing the taint-based Bitcoin flow analysis (Section VI) with those five Bitcoin addresses. As a result, we observe that the perpetrator has been leveraging at least four different Bitcoin exchange sites to cash out the BTCs deposited into their five Bitcoin addresses. As shown in Figure 9, about five percent of the Bitcoins the perpetrator has gained from trafficking firearms (Site A), has been transferred to *Bittrex*, about 23% of the Bitcoins deposited to the Bitcoin address 18JX* used for hacking service sales (Site B) to *BTC-e*, about 15% of 1JyU* (Site B) to *Bitstamp*, about 44% of 1Db2* (Site D) and almost 100% of 1Dzr* (Site C) to *LocalBitcoins*. Based on these findings, we could infer that the perpetrator may have exchanged some of the unlawfully earned BTCs for cash or alt-coins via the exchange sites.

C. *Revealing hidden financial hubs*

Aggregation addresses are often referred to as the Bitcoin addresses that ransomware actors use to collected ransom fees. For example, the Locky and Cerber ransomware actors moved the ransom Bitcoins from many addresses to a small number of aggregate addresses for easier management of the funds[45]. To demystify and understand the ransomware businesses, such addresses that play crucial role must be revealed and analyzed because they perform as *financial hubs* that are monetarily influential and significant.

In our work, we try to reveal the financial hubs of illegal businesses identified on the Dark Web by measuring *betweenness centrality* of every Bitcoin address associated with the illicit Bitcoin addresses. Betweenness centrality is a measure of a node's influence in a graph and, in a Bitcoin transaction graph, an address node with a high centrality value is considered influential.

[11] A militant organization.

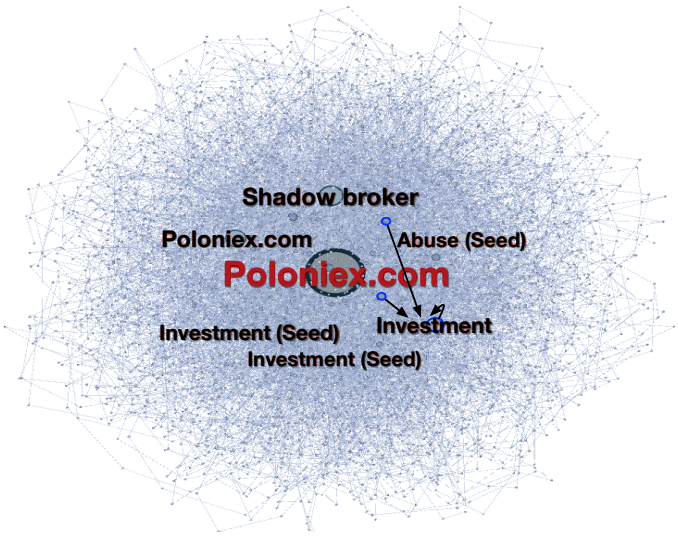

Figure 10: A transaction graph with the betweenness centrality values based on the shortest paths among the seed Bitcoin addresses.

We obtain the shortest transaction paths among the 85 seed Bitcoin addresses from *Learn Me a Bitcoin* [15], which provides a tool for searching the shortest path between Bitcoin addresses. From this data, we construct a Bitcoin transaction graph and calculate the betweenness centrality for all Bitcoin addresses appearing in the graph. Figure 10 illustrates this transaction graph based on the shortest paths among the seed Bitcoin addresses, and the node size denotes the node centrality (e.g., the larger a node, the larger its betweenness centrality value.). Then, for each address with a high centrality value, we search for more information about the address on *WalletExplorer.com* [24] and Google with the following findings:

• One Bitcoin address with the highest centrality value is owned by *Poloniex* [21], which is one of the largest cryptocurrency exchange services. *Poloniex* had not required their customers to undergo identity verification (e.g., no KYC [71]) until recently [35], and the perpetrators may have taken advantage.

• About 18.5% of the Bitcoin addresses appearing in the transaction graph are associated with well-known Bitcoin exchange services, such as *Bittrex* [5], *Xapo* [25], *Mt. Gox* [20], *BTC-e* [8], and *Bitstamp* [4]. In addition, *MoonBit* [19], a popular Bitcoin faucet, features a high centrality value.

• The seed Bitcoin addresses may have been involved in financial activities of the Shadow Broker, which is an infamous hacking group known for selling the confidential information exfiltrated from the NSA [62]. We find that one Bitcoin address (`3CD1QW6fjgTwKq3Pj97nty28WZAVkziNom` [46]) with the second highest centrality value in our graph is one of the Shadow Broker's addresses. This connection implies that the Shadow Broker may have been involved in illicit activities on the Dark Web.

• Two Bitcoin addresses with high centrality values are identified in unknown investment scam sites from the Surface Web. We find their contents and structures to look similar to the investment scam sites where the seed Bitcoin addresses of the Investment category are also identified on the Dark Web.

VIII. Discussion

The following discusses the ethical considerations while conducting the research along with possible solutions to prevent and mitigate illegal transactions behind current anonymity techniques.

A. Ethical concerns

We avoided possible legal compliance issues under the supervision of our government agency, which guided us not to track personally identifiable information and not to share the information without approval. For ethical and respectable research, we set the internal guidelines of (i) collect only publicly accessible data, (ii) do not track any personally identifiable information, such as email addresses and SNS accounts, (iii) store only textual data (e.g., no image or multimedia files) in a private database to which only the four authors have access, and (iv) release data under the supervision of the agency. Furthermore, we reported our findings directly to law enforcement agencies.

B. KYC regulation for preventing illegal activities

In this study, we analyzed illicit Bitcoin addresses from the corpus of the Dark Web, revealed other addresses perpetrators have owned, and traced money flows from these addresses to their destinations. Although we have shown that it is possible to reveal to where the perpetrators have moved funds, it is difficult to investigate further and identify the perpetrators. We also observed that many of the perpetrators sent their unlawfully earned Bitcoins to Bitcoin exchanges, and if these exchanges maintain user record of users, then law enforcement may be able to apprehend the perpetrators.

Government authorities around the world have recently begun to regulate Bitcoin exchanges to comply with KYC (Know Your Customer) policies [63]. Such movements are expected to reduce cybercrimes occurring in the Dark Web gradually. On the other hand, since KYC policies break pseudonymity of cryptocurrencies, a feasible, scientific, and political compromise is required.

IX. Related work

Criminal activity on the Dark Web: The Dark Web is considered to enable perpetrators to perform illegal operations stealthily, and several pioneering researchers have tried to verify this claim [31], [30], [28], [64], [34], [39]. Biryukov *et al.* [31], [30] present an empirical analysis of hidden services hosted over Tor, and they identify many hidden services are maintained for illegal trafficking (e.g., adults, drugs, counterfeits, and weapons). Barratt *et al.* [28] present a global drug survey to determine the reason why drug purchasers prefer (not) to use drug markets on the Dark Web from the perspective of participants. For domain-specific measurements, several works [64], [34], [39] focus on the analysis of popular Dark Web marketplaces. They characterize illegal trafficking on the marketplaces (e.g., transaction patterns, geographical distributions of sellers, and popular items)

and estimate the time-series of their volumes. Although these studies can show the severity of crime schemes on the Dark Web, they have only performed a targeted analysis of specific Dark Web marketplaces. Moreover, they have not investigated the cryptocurrency value chain in the Dark Web, which is a key contribution of this paper.

Analyzing the usage of cryptocurrency for illicit activities on the Dark Web is not an easy problem, and there exist few previous studies within this context [43], [47]. Foley *et al.* [43] propose several features to identify illegal Bitcoin addresses to estimate the volume of illegal activities. While they try to understand the behaviors of Bitcoin usage on the Dark Web, their analysis results are focused on several Dark Web markets and provide only an overall characteristic of Bitcoin usage for those market sites. With respect to understanding illegal value chains in the Dark Web, the recent work [47] attempts to uncover the identities of anonymized users. They use cryptocurrency addresses as a hard identifier that can be linked to real identities and measure possible economic activities through Bitcoin transaction analysis. Our work differs in that we conduct a large-scale analysis with the recently collected data (i.e., March 2018) from diverse dark websites (i.e., more than 23 million pages). In addition, we provide the financial characteristics of cryptocurrency on the Dark Web, such as the dominant cryptocurrency services used by perpetrators, and trace money flows through our taint-based financial analysis. Our case studies also reveal the complete illegal value chains in the Dark Web ecosystem.

Cybercrime exposure: Several previous projects explored various ways to expose domain-specific cybercrimes [51], [45], [60]. Levchenko *et al.* [51] perform an empirical study on advertising spams to determine the end-to-end value chain of spam networks and identify bottlenecks for spam campaigns. Huang *et al.* [45] reveal ransomware value chains by exploiting pseudonyms of Bitcoin and estimate over $16 million ransom payments for nearly 20,000 victims. Rebecca *et al.* [60] present methodologies to cluster sex advertisements by owner based on the algorithms that the Backpage enables premium features for Bitcoin transactions. These studies successfully identify illegal value chains through actively participating in each campaign. Unlike these approaches, we cannot participate in live deals on the Dark Web since even a simple payment can be regarded as an illegal operation (e.g., child pornography[12]). While we have limited strategies restricted by ethical research issues, our work also identifies illegal value chains on the Dark Web.

Several researchers focus on a specific type of crime scheme in cryptocurrencies [68], [69], [29], [57], [37]. Vasek *et al.* [68] present an empirical study on Bitcoin scams to understand their scale and severity. Two studies analyze Bitcoin Ponzi schemes to derive their features based on information collected from public forums [69] or transactions [29]. A money laundry (i.e., mixing) is one illegal service in cryptocurrency and exploits the pseudonymity of cryptocurrencies to avoid tracking financial flows. Möser *et al.* [57] examine several Bitcoin laundry services to expose the limitations of the anti-money laundering (AML) policy as it applies to Bitcoin. Balthasar *et al.* [37] also perform a similar analysis and then estimate the volume of each laundry service through financial analysis of mixing Bitcoin addresses. While these studies provide specialized measurements on a dedicated dataset to each type of cryptocurrency scheme, our work covers not only the larger dataset but also many types of cybercrime schemes. In addition, we analyze how illegal users and activities are related through the heterogeneous analysis over the Surface Web, the Dark Web, and cryptocurrencies.

[12]Some dealers send passcodes through the Bitcoin accounts where buyers have deposited.

X. Conclusion

While the Dark Web and cryptocurrencies are proposed to offer benefits for our communities, it is also known they are leveraged for malicious purposes. However, no previous studies have rigorously investigated the claim — *the Dark Web and cryptocurrencies are misused for malicious operations*. We believe our work is the first significant step toward exposing illicit activities involving the Dark Web and cryptocurrency. Starting from collecting large volumes of dark websites and cryptocurrency usage through these sites, our work provides an in-depth analysis and provides evidence of abuse for malicious purposes. Also, we reveal illegal value chains, Bitcoin investment scams and trafficking, that clearly explain how perpetrators employ cryptocurrency in the Dark Web and how money is traded. Our findings and discussions in our work shed light on the Dark Web black market, which has been minimally evaluated to date.

Acknowledgment

The authors would like to thank the anonymous reviewers for their feedback, which improved the paper. We sincerely appreciate our shepherd Xiaojing Liao for guiding us in reflecting on important comments from the reviewers.

References

[1] *Ahmia*, https://ahmia.fi/.
[2] *Bcoin-cli*, https://github.com/bcoin-org/bcoin/wiki/CLI.
[3] *bitcasino*, https://bitcasino.io/.
[4] *Bitstamp*, https://www.bitstamp.net.
[5] *Bittrex*, https://bittrex.com/.
[6] *Bleutrade*, https://bleutrade.com.
[7] *Blockchain.com*, https://www.blockchain.com.
[8] *BTC-e*, [seized] https://btc-e.com.
[9] *Coingaming*, http://coingaming.io/.
[10] *Cryptopay*, https://cryptopay.me.
[11] *Dream Market*, http://n3mvkmkqb3ry4rbb.onion.
[12] *Fresh Onions*, http://zlal32teyptf4tvi.onion.
[13] *Hansa Market*, [seized] http://hansamkt2rr6nfg3.onion/affiliate/110.
[14] *Haystak*, http://haystakvxad7wbk5.onion/.
[15] *Learn Me A Bitcoin*, http://learnmeabitcoin.com/.
[16] *LocalBitcoins*, https://localbitcoins.com.
[17] *Market Capitalization*, https://coinmarketcap.com/charts/.
[18] *Monero: Private Digital Currency*, https://getmonero.org/.
[19] *MoonBit*, moonbit.co.in.
[20] *mt-gox*, http://mtgox.com/.
[21] *Poloniex*, https://poloniex.com.
[22] *Silk Road Market*, http://silkroad7rn2puhj.onion/.
[23] *Sportsbet*, https://sportsbet.io/.
[24] *WalletExplorer*, https://www.walletexplorer.com/.
[25] *Xapo*, https://xapo.com.

[26] Y. Akdeniz, "Anonymity, democracy, and cyberspace," *Social Research: An International Quarterly*, vol. 69, no. 1.

[27] E. Androulaki, G. O. Karame, M. Roeschlin, T. Scherer, and S. Capkun, "Evaluating user privacy in bitcoin," in *International Conference on Financial Cryptography and Data Security (ICFCDS 2013)*.

[28] M. J. Barratt, J. A. Ferris, and A. R. Winstock, "Use of silk road, the online drug marketplace, in the united kingdom, a ustralia and the united states," *Addiction*, vol. 109, no. 5, pp. 774–783, 2014.

[29] M. Bartoletti, B. Pes, and S. Serusi, "Data mining for detecting bitcoin ponzi schemes," *CoRR*, vol. abs/1803.00646, 2018.

[30] A. Biryukov, I. Pustogarov, F. Thill, and R.-P. Weinmann, "Content and popularity analysis of tor hidden services," in *Distributed Computing Systems Workshops (ICDCSW 2014)*.

[31] A. Biryukov, I. Pustogarov, and R.-P. Weinmann, "Trawling for tor hidden services: Detection, measurement, deanonymization," in *Security and Privacy (S&P 2013)*.

[32] BitcoinWiki, *Address reuse*. [Online]. Available: https://bit.ly/2LRWVCS

[33] M. Burgess, "Hackers took more than 10,000 dark web sites offline," WIRED, Feb 2017. [Online]. Available: https://bit.ly/2LTuPXS

[34] N. Christin, "Traveling the silk road: A measurement analysis of a large anonymous online marketplace," in *Proceedings of the 22nd International Conference on World Wide Web (WWW 2013)*.

[35] *Crypto Exchange Poloniex to Impose Customer ID Requirements*, Coindesk, Dec 2017. [Online]. Available: https://bit.ly/2BV2Uhf

[36] *Users freak out as LocalBitcoins forces users to submit KYC*, Crypto NEWS, April 2018. [Online]. Available: https://bit.ly/2KwuViD

[37] T. de Balthasar and J. Hernandez-Castro, "An analysis of bitcoin laundry services," in *Secure IT Systems*. Springer International Publishing, 2017.

[38] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," Naval Research Lab Washington DC, Tech. Rep., 2004.

[39] D. S. Dolliver, "Evaluating drug trafficking on the Tor Network: Silk Road 2, the sequel," *International Journal of Drug Policy*, vol. 26, no. 11, pp. 1113–1123, 2015.

[40] D. Ermilov, M. Panov, and Y. Yanovich, "Automatic bitcoin address clustering," in *Machine Learning and Applications (ICMLA), 2017 16th IEEE International Conference on*.

[41] Ethereum, "Go-ethereum," https://github.com/ethereum/go-ethereum.

[42] *Bitcoin virtual currency: Unique features present distinct challenges for deterring illicit activity.*, FBI, 2012, https://bit.ly/2nuMpTl.

[43] S. Foley, J. Karlsen, and T. J. Putniņš, "Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies?" 2018. [Online]. Available: https://bit.ly/2nb6kGW

[44] S. Goldfeder, H. Kalodner, D. Reisman, and A. Narayanan, "When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies," *arXiv preprint arXiv:1708.04748*, 2017.

[45] D. Y. Huang, D. McCoy, M. M. Aliapoulios, V. G. Li, L. Invernizzi, E. Bursztein, K. McRoberts, J. Levin, K. Levchenko, and A. C. Snoeren, "Tracking ransomware end-to-end," in *Symposium on Security and Privacy (S&P 2018)*, 2018.

[46] T. R. Inc., *Shadow Broker Group Main Wallets: $ 4,189,786,240 Billion Dollars*. [Online]. Available: https://bit.ly/2OfahpL

[47] H. A. Jawaheri, M. A. Sabah, Y. Boshmaf, and A. Erbad, "When a small leak sinks a great ship: Deanonymizing tor hidden service users through bitcoin transactions analysis," *arXiv preprint arXiv:1801.07501*, 2018.

[48] H. Kalodner, S. Goldfeder, A. Chator, M. Moser, and A. Narayanan, "Blocksci: Design and applications of a blockchain analysis platform," *arXiv preprint arXiv:1709.02489*, 2017.

[49] R. Kang, S. Brown, and S. Kiesler, "Why do people seek anonymity on the internet?: Informing policy and design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2013)*.

[50] D. Y. Kao and S. C. Hsiao, "The dynamic analysis of wannacry ransomware," in *20th International Conference on Advanced Communication Technology (ICACT 2018)*, 2018.

[51] K. Levchenko, A. Pitsillidis, N. Chachra, B. Enright, M. Félegyházi, C. Grier, T. Halvorson, C. Kanich, C. Kreibich, H. Liu *et al.*, "Click trajectories: End-to-end analysis of the spam value chain," in *Security and Privacy (S&P 2011)*.

[52] luigi1111, "xmr.llcoins.net," https://github.com/luigi1111/xmr.llcoins.net.

[53] F. K. Maurer, T. Neudecker, and M. Florian, "Anonymous coinjoin transactions with arbitrary values," in *Trustcom/BigDataSE/ICESS 2017*.

[54] D. MCQUAID, *Bitcoin warning: Criminals turning to other cryptocurrencies on Dark Web*, EXPRESS, 2018.

[55] S. Meiklejohn, M. Pomarole, G. Jordan, K. Levchenko, D. McCoy, G. M. Voelker, and S. Savage, "A fistful of bitcoins: characterizing payments among men with no names," in *Proceedings of the conference on Internet Measurement Conference (IMC 2013)*.

[56] M. Moser, "Anonymity of bitcoin transactions: An analysis of mixing services." 2013.

[57] M. Moser, R. Bohme, and D. Breuker, "An inquiry into money laundering tools in the bitcoin ecosystem," in *eCrime Researchers Summit (eCRS 2013)*.

[58] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system." [Online]. Available: http://bitcoin.org/bitcoin.pdf

[59] M. Paquet-Clouston, B. Haslhofer, and B. Dupont, "Ransomware payments in the bitcoin ecosystem," *CoRR*, vol. abs/1804.04080, 2018.

[60] R. S. Portnoff, D. Y. Huang, P. Doerfler, S. Afroz, and D. McCoy, "Backpage and bitcoin: Uncovering human traffickers," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2017)*.

[61] T. Ruffing, P. Moreno-Sanchez, and A. Kate, "Coinshuffle: Practical decentralized coin mixing for bitcoin," in *European Symposium on Research in Computer Security*.

[62] B. Schneier, *Who Are the Shadow Brokers?*, The Atlantic.

[63] E. Seyi, *The Emerging Role of KYC and AML in Cryptocurrencies*, SmilePass, 2018. [Online]. Available: https://bit.ly/2KBSGpD

[64] K. Soska and N. Christin, "Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem," in *Proceedings of the 24th USENIX Security Symposium*.

[65] P. Syverson and G. Boyce, "Bake in. onion for tear-free and stronger website authentication," *IEEE Security & Privacy (S&P 2016)*.

[66] Tor2web, "Browse the tor onion services." [Online]. Available: https://www.tor2web.org/

[67] *Russian National And Bitcoin Exchange Charged In 21-Count Indictment For Operating Alleged International Money Laundering Scheme And Allegedly Laundering Funds From Hack Of Mt. Gox*, United States Department of Justice, July 2017.

[68] M. Vasek and T. Moore, "There's no free lunch, even using bitcoin: Tracking the popularity and profits of virtual currency scams," in *Financial Cryptography and Data Security*. Springer, 2015.

[69] ——, "Analyzing the bitcoin ponzi scheme ecosystem," in *Bitcoin Workshop*, 2018.

[70] B. Weiser and D. Carvajal, "International raids target sites selling contraband on the 'dark web'." [Online]. Available: https://nyti.ms/2vnpT3j

[71] *Know your customer*, Wikipedia, April 2018, https://bit.ly/S0LGa9.

[72] G. Wood, "Ethereum: A secure decentralized transaction ledger," 2014. [Online]. Available: https://bit.ly/2hhPViV