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†‡

†S2W LAB Inc.
‡KAIST
Anonymity services

- Tor hidden services
  - Strong anonymity guarantees via Tor
  - Anonymous and untraceable network
  - Special software to access (i.e., Tor browser)

Cryptocurrency

- A blockchain-based digital currency
  - Cryptographic currency
  - Strong pseudonymity (almost anonymous)
  - Mostly unregulated and unverifiable
Illicit trades on the Dark Web

- Firearms
- Malware
- Rent-a-hacker
- Hacker recruitment
- Ransomware service
- Fake ID (Passport)
- Hitman services
The procedures of an underground illegal trade

1. Post ads
2. Discover
3. Order & negotiation
4. Receive payment
5. Fulfill the order

1. Post ads/deals
2. Discover & visit
3. Order & negotiation

The Surface/Deep Web
The Dark Web
Cryptocurrency

Trafficker
Customers
How badly is cryptocurrency being abused on the Dark Web?

- Perform large-scale analysis on the Dark Web.

How can we investigate and analyze cryptocurrency abuses on the Dark Web?

- Investigate the perpetrators’ online & financial activities.
Challenge: Limited Dark Web data accessibility

The SURFACE WEB
- Searchable (via Google, Bing, etc.) data
- Most of the known websites

The DARK WEB
- No Dark Web search engines with extensive coverage (i.e., Poor-indexed websites)
- Highly volatile contents
Challenge: Lack of evidence

I will conceal evidence that possibly reveals our entire illegal business!

Hacked account selling

Hacked Netflix accounts

I have gained access to several legit, unlimited Netflix accounts. All accounts are corporate accounts so they will not be deactivated. All accounts are guaranteed as long as the service is offered. Send 0.025 BTC to the wallet below.

hen, email <username>@sigaint.org with your transaction ID. Once confirmed, I will reply with a username and password.

Bitcoin wallet: 1FBSBDp9dKw4SxruWAp7amWb1CNZcK

Using privacy-oriented services (e.g., SIGAINT*)

*SIGAINT is a Tor hidden service offering secure email services
Challenge: Obscure cryptocurrency money flows

1. Lack of explicit links
   ➔ How much BTCs are transferred to each output?

2. Lack of ownership information
   ➔ Who receives funds?

- **Addr_A**: 1.5 BTCs
- **Addr_B**: 3 BTCs
- **Addr_C**: 2 BTCs
- **Addr_D**: 2.5 BTCs
- **Addr_E**: 4 BTCs
**MFScope**: Dark Web and cryptocurrency analysis framework

Data collection part

- Dark Web Search Engines
  - Seed .onion domains
- Dark Web
  - .onion links

Data analysis part

- Crawlers
- D-DB
- Texts
- Address Extraction
- Address Classification
- Cross-domain Analysis
- Address Clustering
- Financial Flow Analysis
Crawling the Dark Web

- Seed 10k .onion addresses from
  - Ahmia*
  - FreshOnions**

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td># of dark websites</td>
<td>36,864</td>
</tr>
<tr>
<td># of dark webpages</td>
<td>27,665,572</td>
</tr>
<tr>
<td>Data collection period</td>
<td>Jan 2017 ~ March 2018 (15 months)</td>
</tr>
</tbody>
</table>

Our data collection

* Ahmia, https://ahmia.fi
** FreshOnions, http://zlal32teyptf4tvi.onion
Extracting cryptocurrency addresses

- Extract addresses with each **regex** from dark webpages
  - Bitcoin, Ethereum, and Monero
- Filter invalid addresses are
  - Cryptographically invalid
  - Having no transactions
  - Extracted from Blockchain mirror sites

<table>
<thead>
<tr>
<th></th>
<th>Bitcoin</th>
<th>Ethereum</th>
<th>Monero</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td># dark websites</td>
<td>2,886</td>
<td>180</td>
<td>121</td>
<td>3,187</td>
</tr>
<tr>
<td># dark webpages</td>
<td>1,579,047</td>
<td>4,743</td>
<td>4,410</td>
<td>1,588,200</td>
</tr>
<tr>
<td># addresses</td>
<td>9,906,129</td>
<td>649</td>
<td>38,440</td>
<td>9,945,218</td>
</tr>
<tr>
<td># final addresses</td>
<td>5,440</td>
<td>50</td>
<td>61</td>
<td>5,551</td>
</tr>
</tbody>
</table>
Classifying Bitcoin addresses

Is this Bitcoin address for promoting illicit goods or services?

<table>
<thead>
<tr>
<th>Vote (%)</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>~ 20%</td>
<td>Legitimate</td>
</tr>
<tr>
<td>20% ~ 70%</td>
<td>Possible illicit</td>
</tr>
<tr>
<td>70% ~</td>
<td>Illicit</td>
</tr>
</tbody>
</table>

Hacked Netflix accounts

I have gained access to several legit, unlimited Netflix accounts. All accounts are corporate accounts so they will not be deactivated. All accounts are guaranteed as long as the service is offered. Send 0.025 BTC to the wallet below. Then, email nixplay@sigaint.org with your transaction id. Once confirmed, I will reply with a username and password.

Hosting service

24 hours FREE hosting time

<table>
<thead>
<tr>
<th>3 months</th>
<th>6 months</th>
<th>1 year</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.02 BTC</td>
<td>0.034 BTC</td>
<td>0.060 BTC</td>
</tr>
</tbody>
</table>

Hosting only!

+ Free specific .onion domain with your 7 first letters if you buy 6 months or with 8 first letters if you buy 12 months hosting

Copyright 2019 Real Hosting. All rights reserved.
Classifying Bitcoin addresses

<table>
<thead>
<tr>
<th>Category</th>
<th>Count</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legitimate addresses</td>
<td>884</td>
<td>16.25%</td>
</tr>
<tr>
<td>Possible illicit addresses</td>
<td>4,471</td>
<td>83.75%</td>
</tr>
<tr>
<td>Illicit addresses</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>5,440</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Cryptocurrency distribution over the Dark Web

- **Legitimate addresses**
  - Donation, escrow, identification, etc.

- **85 seed illicit addresses**
  - Abuse, counterfeit, drug, weapons, etc.
Demystifying perpetrators: Address clustering

**MI (Multi-input) heuristic**

A sender should have corresponding private keys ➔ $\text{Addr}_A, \text{Addr}_B, \text{Addr}_C \subset \text{sender’s addresses}$

**CA (change address) heuristic**

A wallet software generates a new address to receive the remainder ➔ $\text{Addr}_A, \text{Addr}_A' \subset \text{sender’s addresses}$
## The volumes of perpetrator’s Bitcoin

<table>
<thead>
<tr>
<th>Category</th>
<th># addr</th>
<th>BTC (USD) received**</th>
<th>BTC (USD) sent**</th>
<th>Lifetime (TX\textsubscript{first} – TX\textsubscript{last})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Seed</td>
<td>MI</td>
<td>MI + CA*</td>
<td>[MI heuristic only]</td>
</tr>
<tr>
<td>Abuse</td>
<td>15</td>
<td>486</td>
<td>539</td>
<td>3,416 ($3,862,983)</td>
</tr>
<tr>
<td>Account selling</td>
<td>6</td>
<td>60</td>
<td>201</td>
<td>2 ($1,811)</td>
</tr>
<tr>
<td>Card dumps</td>
<td>6</td>
<td>205</td>
<td>833</td>
<td>2,323 ($9,935,313)</td>
</tr>
<tr>
<td>Counterfeit</td>
<td>2</td>
<td>23</td>
<td>27</td>
<td>0.49 ($1,129)</td>
</tr>
<tr>
<td>Drug</td>
<td>4</td>
<td>18</td>
<td>26</td>
<td>5,245 ($14,124,499)</td>
</tr>
<tr>
<td>Investment scam</td>
<td>29</td>
<td>2,025</td>
<td>204</td>
<td>32,428 ($151,438,331)</td>
</tr>
<tr>
<td>Membership</td>
<td>8</td>
<td>95</td>
<td>247</td>
<td>29 ($85,481)</td>
</tr>
<tr>
<td>Service</td>
<td>8</td>
<td>113</td>
<td>861</td>
<td>59 ($60,141)</td>
</tr>
<tr>
<td>Weapon</td>
<td>1</td>
<td>42</td>
<td>754</td>
<td>46 ($32,964)</td>
</tr>
<tr>
<td>Others</td>
<td>6</td>
<td>9</td>
<td>22</td>
<td>65 ($32,043)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>85</td>
<td>3,029</td>
<td>2,044</td>
<td>43,422 ($179,317,131)</td>
</tr>
</tbody>
</table>

* The super clusters are excluded
** The amount is based on the trading currency of BTC to USD
Demystifying perpetrators: Cross-domain analysis

- Cryptocurrency address as a keyword
- Google search API

<table>
<thead>
<tr>
<th>Category</th>
<th>Seed</th>
<th>MI</th>
<th>MI+CA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tor proxy</td>
<td>38</td>
<td>38</td>
<td>45</td>
<td>121</td>
</tr>
<tr>
<td>Community</td>
<td>35</td>
<td>59</td>
<td>20</td>
<td>114</td>
</tr>
<tr>
<td>Sales</td>
<td>17</td>
<td>27</td>
<td>9</td>
<td>53</td>
</tr>
<tr>
<td>Media</td>
<td>10</td>
<td>17</td>
<td>5</td>
<td>32</td>
</tr>
<tr>
<td>Archive</td>
<td>4</td>
<td>12</td>
<td>6</td>
<td>22</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Unavailable</td>
<td>8</td>
<td>17</td>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>113</td>
<td>173</td>
<td>95</td>
<td>381</td>
</tr>
</tbody>
</table>

The search results include:
- real-world identities (profile)
- personal interests
- complaints (fraud reports)
- service feedback
- another illicit business

Help us understand the perpetrators

The statistics of cross-referencing results (# pages)
Taint-based cryptocurrency flow analysis

Stop conditions
1. Unspent output
2. Bitcoin service address*
3. Threshold (i.e., # of transaction)

* WalletExplorer provides ownership information about known Bitcoin services
Perpetrator’s black money operation

61.4 % has flown into "exchange"

Distribution of the entire illicit Bitcoins

Perpetrators prefer to exchange illicit funds into

*alt-coins or traditional currencies*
Service distribution per cryptocurrency account

Service usage per each illicit address

A large sum of their money to **one particular service** rather than **diversifying their expenditure**
Case study: trafficking

Cluster ID: ********

1. Armes trafficking (Site A)
2. Hacking as a service (Site B)
3. Image for sale (Site C)
4. User profile on an underground forum (Site D)
5. Posts on the forum (Site D)
6. A hacking blog (Site E)
7. A post on the blog (Site E)

Withdrawal percentages:
- 5%
- 14.93%
- 22.56%
- 99.91%
- 43.68%

Location info.

Bitcoin exchange services:
- BITTREX
- BITSTAMP
- LocalBitcoins.com

Perpetrator

Bitcoin network

Cluster ID: ********

Withdraw (5%)
Withdraw (14.93%)
Withdraw (22.56%)
Withdraw (99.91%)
Withdraw (43.68%)

Perpetrator (Site D)

Perpetrator (Site D) User profile on an underground forum
Conclusion

- Cryptocurrency abuses in the Dark Web are pervasive.
- We should think about the dark side of anonymity services.

Don’t be relieved;

few evidence could reveal what you did in the Dark Web.