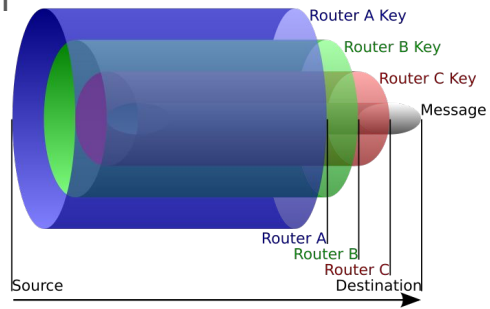
Measurement of Bitcoin in Tor networks

Chungnam National University, Korea
Data Network Lab.
Hyunsu Mun, Jonghyeon Park, Youngseok Lee
{munhyunsu, pjh000901, lee}@cnu.ac.kr

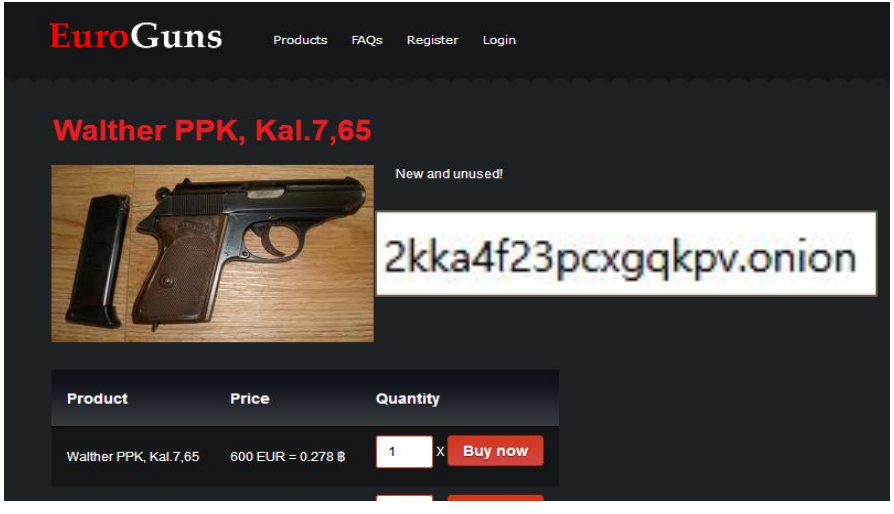
Tor and **Bitcoin** in Hidden Service

- Tor (The onion router, https://www.torproject.org/):
  - Special network that supports anonymous communication using onion routing
    - Onion routing: A series of network nodes that peels a single layer of the encrypted data, uncovering the data's next destination
    - Apply anonymity to clients for websites
  - A kind of Deep web
- Hidden Service
  - Web site or server that only receives inbound connections through Tor
  - Use a special address called "onion address" consisting of 16 alphanumeric characters instead of IP address or common domain



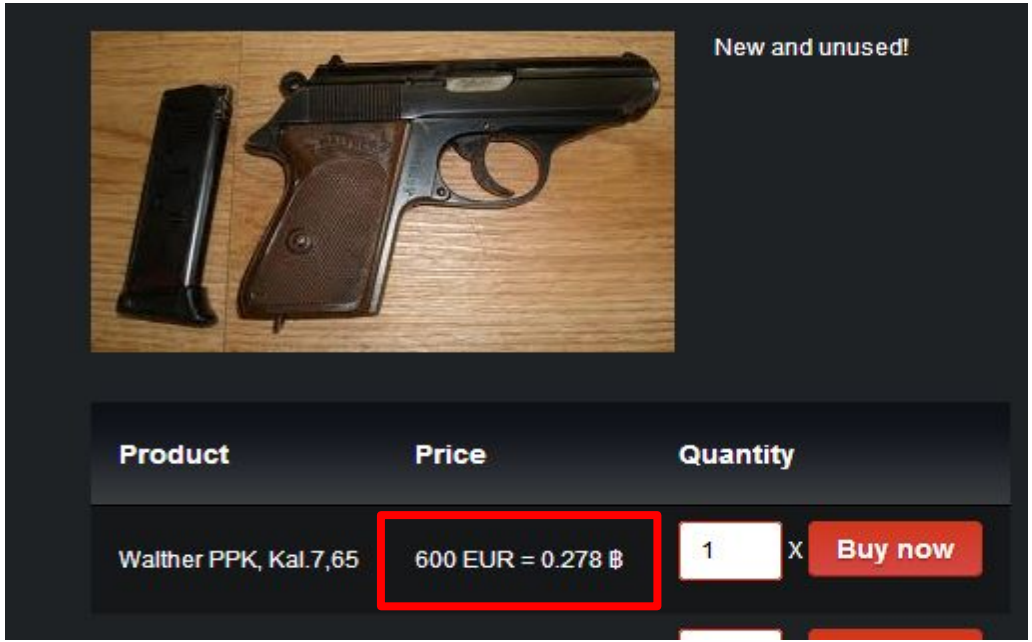
Hidden Services in Tor

- Hacking, Personal tracking, Ransomware creation services
- Sale of counterfeit passport, cloned credit card, ID card
- Many markets where various illegal goods are traded, and communities that deal with dangerous contents.



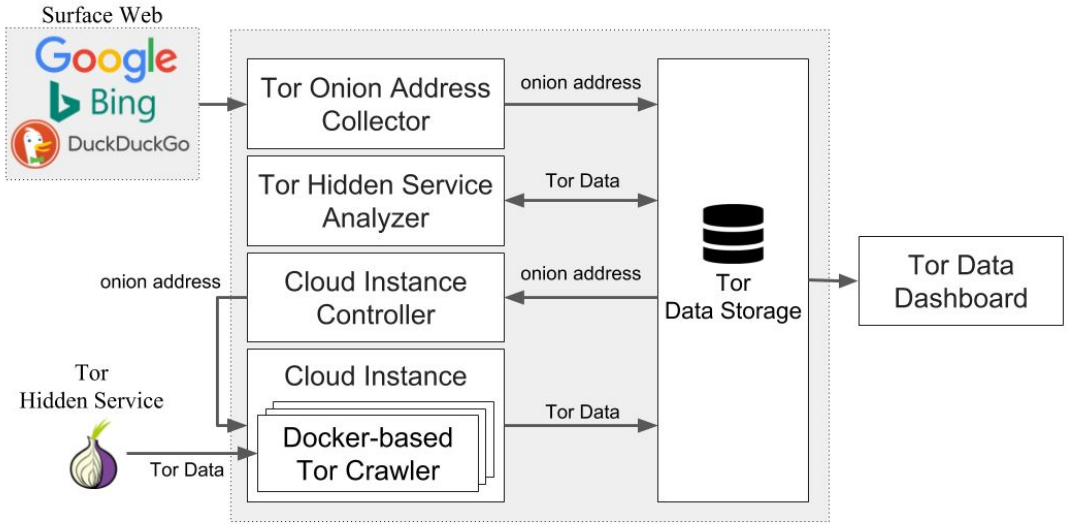
Bitcoin in Hidden Services

- Peer-to-peer transation is available without intervention of the agency
- Difficult to find correlation between trader and Bitcoin address
- Currency of hidden services
    - Anonymity of Tor and Bitcoin



Measurement of Bitcoin in Tor networks

1. Collect Tor onion addresses
2. Crawl deep web pages
3. Extract Bitcoin address in deep web
4. Analysis Bitcoin address and contents of deep web



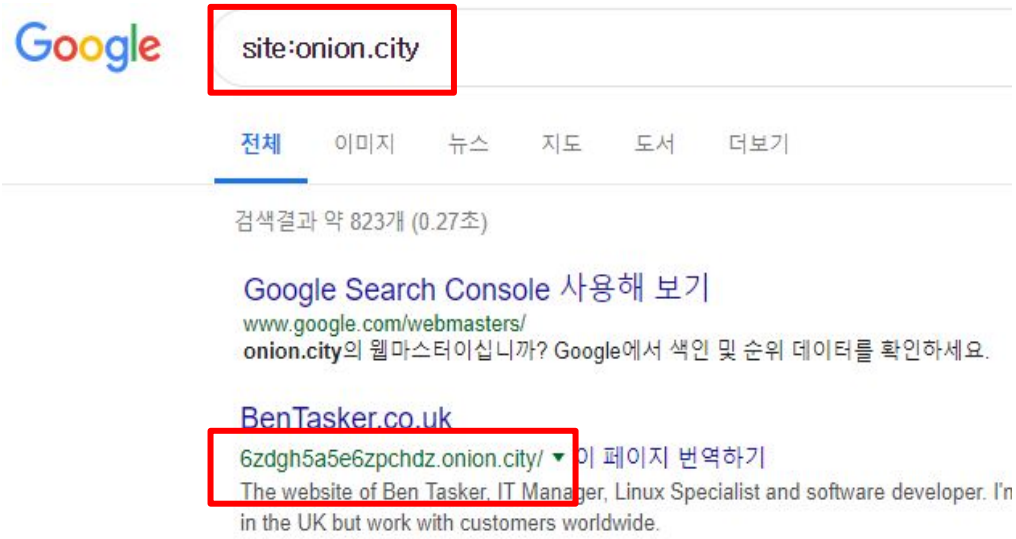
Collect Onion Address using Out-of-Band Search

- A collector of deep web address is required because the onion address contains 16-digit alphanumeric letter, which is different from general URL
- Can find deep web address from the surface web using search engines
  - http://AAAAAAAAAAAAAAAA.onion > http://AAAAAAAAAAAAAAAA.onion.city
- Tor2Web: A tunnelling service to access hidden services from general web

onion.city	onion.nu
onion.to	onion.lt
onion.cab	onion.direct
onion.link	tor2web.org
onion.lu	tor2web.fi
onion.rip	torstorm.org

Collect Onion Address using Out-of-Band Search

- Search for every TLD (ex. site:onion.city)
- 174,252 Onion addresses were collected in 30 days
  - 2,439 Onion addresses after deletion of repeated or subpage addresses

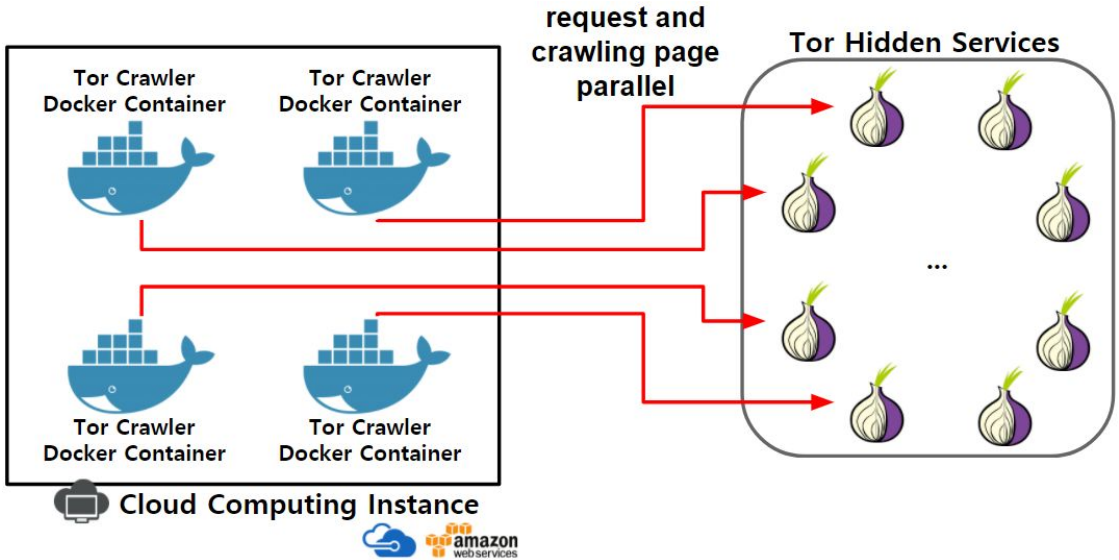


Docker-based Hidden Service Crawler

- Tor Onion connection is slow because it connects to a server through at least 3 nodes
- Slow access to Tor Hidden Services is a challenge to quickly collect and analyze Hidden Services.
- A virtualized crawler on the Cloud Computing Instance
    - Maximize the utilization of computing resources
    - Speed up crawler in parallel

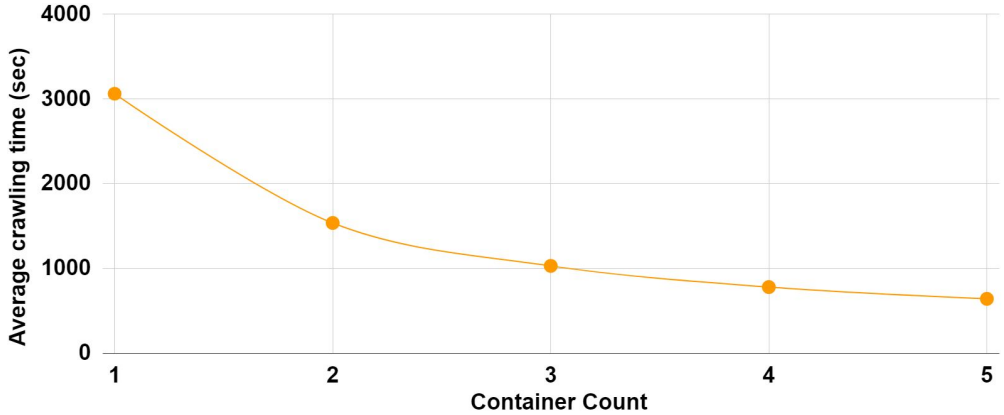
Docker-based Hidden Service Crawler

- Use a docker to analyze hidden services using fewer instances.
- Run one or more docker containers in one instance, such as Microsoft Azure or Amazon EC2.



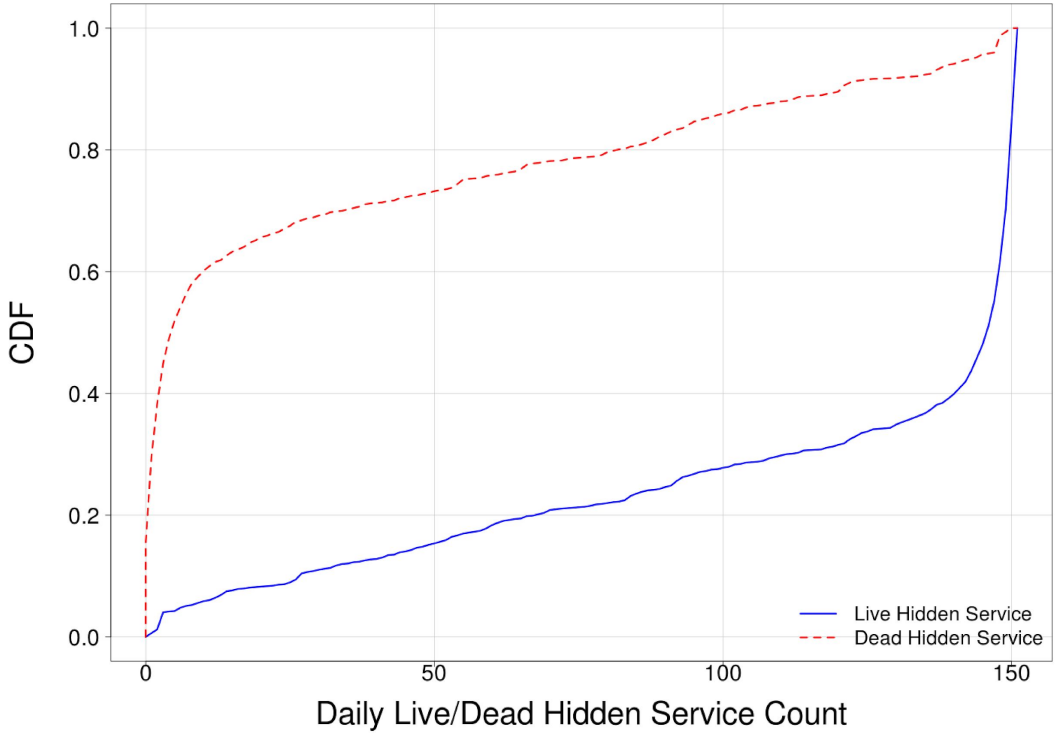
Improved Crawling Time with Docker

- Experiment on Microsoft Azure's 2 core 7GB memory instance
- Measures the crawling time according to the number of containers for 100 onion addresses
- For one container, it takes 3,062 seconds.
- With five Docker containers, it takes average of 640 seconds (4.78 times faster than one container)



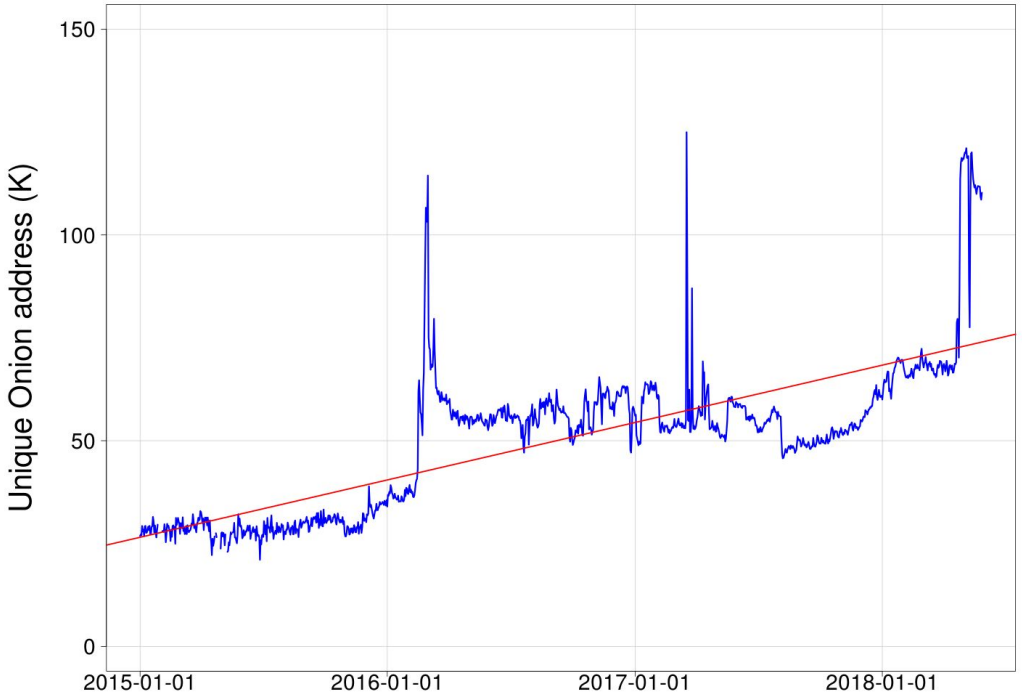
An irregular pattern of Tor Deep Web Service

- 80% of Tor services disappeared in about 80 days
  - Unable to access the address
- New addresses are created consistently
  - Address extraction from Deep Web is required



Onion Address in Deep Web Crawler

- Extract Onion address from Deep web page as well as surface web searching
- Found new address consistently
    - Delete addresses that not accessible for 3 months



Similar Tor Deep Web Page

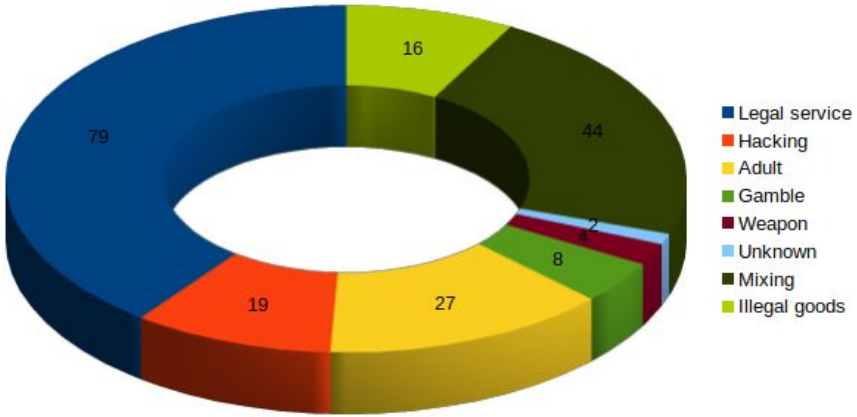
- Several Onion addresses for same hidden service
  - This makes collecting repeated addresses
    - Collecting is delayed due to the slow connection of Tor
  - Grouping to collect only one Onion from the service
- Using Cosine similarity
  - HTML document word vector
  - HTTP Header

Categories of Tor Deep Web contents

- Categorizing the contents to understand the status of Deep Web Page
  - 8 categories: Adult, Illegal Goods, Mixing, Gamble, Legal Service, Hacking, Weapon, Unknown
- Extract feature of document for TF-IDF
- Classification using Naive Bayesian model
  - Accuracy 83%, Recall 82%

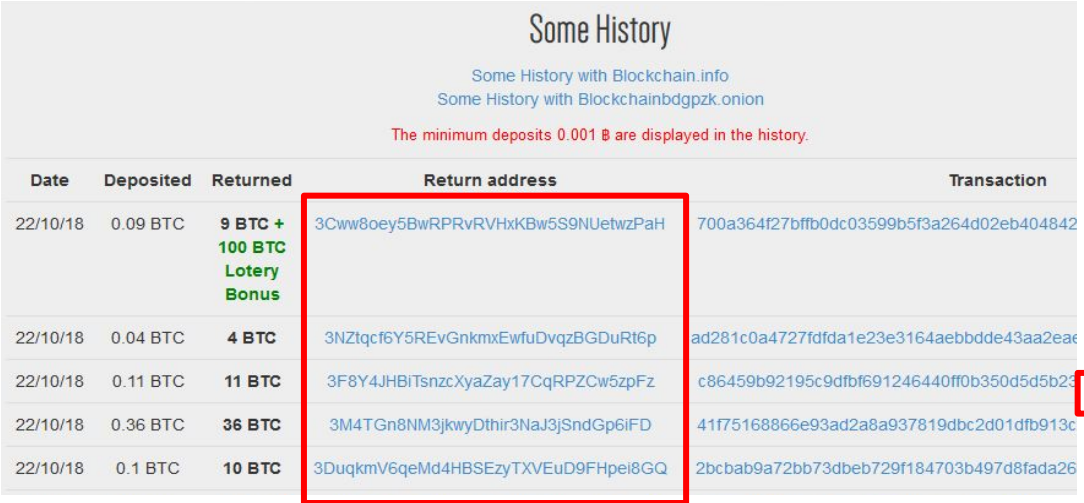
Illegal Services in Tor Dark Web

- Found Bitcoin addresses from 199 hidden service main pages
  - 8% of collected hidden services
- Out of 199, 118 were illegal service (59%)



Extracting Bitcoin address

- Extract Bitcoin address of service provider/user in Deep Web page
  - Able to extract Bitcoin address from the text of collected hidden service page by regular expression
    - [13][A-HJ-NP-Za-km-z1-9]{25,33}
- Extracted 3,917 Bitcoin addresses
  - 2,558 hidden service pages were collected using 12,511 Onion addresses



Some History

Some History with Blockchain.info
Some History with Blockchainbdgpzk.onion

The minimum deposits 0.001 ฿ are displayed in the history.

Date	Deposited	Returned	Return address	Transaction
22/10/18	0.09 BTC	9 BTC + 100 BTC Lotery Bonus	3Cww8oey5BwRPRvRVHxKBw5S9NUetwzPaH	700a364f27bffb0dc03599b5f3a264d02eb404842
22/10/18	0.04 BTC	4 BTC	3NZtqcf6Y5REvGnkmxEwfuDvqzBGDuRt6p	ad281c0a4727fdfda1e23e3164aebbdde43aa2eae
22/10/18	0.11 BTC	11 BTC	3F8Y4JHBiTsnzcXyaZay17CqRPZCw5zpFz	c86459b92195c9dfbf691246440ff0b350d5d5b23
22/10/18	0.36 BTC	36 BTC	3M4TGn8NM3jkwyDthir3NaJ3jSndGp6iFD	41f75168866e93ad2a8a937819dbc2d01dfb913c
22/10/18	0.1 BTC	10 BTC	3DuqkmV6qeMd4HBSEzyTXVEuD9FHpei8GQ	2bcbab9a72bb73dbeb729f184703b497d8fada26

Hack Facebook Account

We hack Facebook accounts and we sell this service.

Price per account: **0.019 BTC**

Bitcoin address for making deposit:

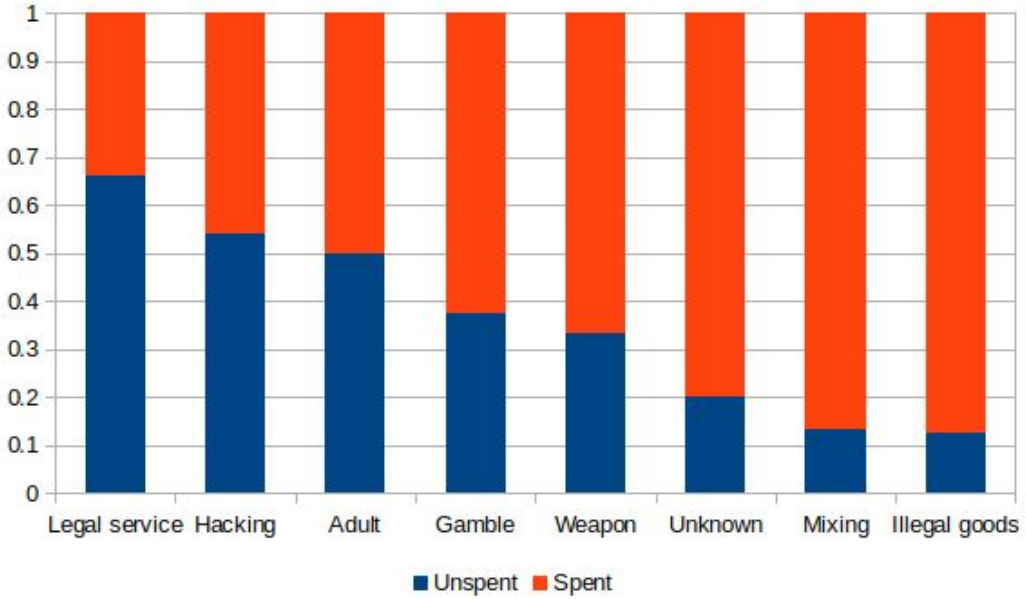
19SjPWrqL9vd9RmtTCGoURbhJXxKLK4d9P

Bitcoin Address: UTXO

- Unspent Transaction Output (UTXO) can be spent as an input in a new transaction
- Most of the Bitcoin addresses were already used: 87% of 3,917 Bitcoin addresses
    - It is because an illegal service moves to another address after Bitcoin is deposited in its representative address

UTXO according to Bitcoin categories

- In legal service, 67% of the addresses were unspent addresses
- A high rate of spented address appears in illegal goods and mixing Bitcoin



Bitcoin transaction amount by content

- Mixing: A large amount of Bitcoin is involved in mixing transactions to improve the anonymity
- Hacking/Cyber attack: Observed in gathering the remittances of victims of ransomware and hacking
- Illegal Goods, Adult/Child Pornography, Weapon services: about 0.1BTC on average

Category	Maximum Tx Value (BTC)	Minimum Tx Value (BTC)	Average Tx Value (BTC)
Mixing	48500.2	0.00000001	45.8486
Hacking	1280.1	0.00000001	8.9076
Adult	15.0	0.00000001	0.1219
Legal Service	15.0	0.00000008	0.3962
Illegal Goods	3.0	0.00001200	0.1183
Gamble	1.0	0.00001159	0.0232
Weapon	0.9	0.00031184	0.1468
Unknown	10.0	0.00000001	0.0786

Conclusions

- Improving Tor Hidden Service crawler performance
  - Reduced crawling time by up to 79%
  - Onion address clustering reduces address set by 21% and crawling time by 39%.
- Tor Hidden services life and operation looks different from normal web page
  - Many illegal Hidden Services
  - Irregular operation that is difficult to access at all times
- Bitcoin address in Tor deep web actual transaction observation
  - Many Bitcoin purses were used in Bitcoin mixing
  - Illegal goods trading uses Bitcoin most actively in the transation: low UTXO