IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

IDENTIFYING ILLICIT ACTORS USING TOR NETWORK AND PUBLIC BITCOIN BLOCKCHAIN DATA

[1]VIJAY GIRI, [2]RAJESH PANDEY, [3]SUBODH KANT TIWARI

[1]Student, [2]Assistant Professor, [3]Student

[1] *Dept. of Computer Science and Engg.*,

[1]Shobhit University, Meerut, India

Abstract: A Bitcoin transaction is a transfer of value between Bitcoin wallets. It involves the sender's digital signature, timestamp, transaction hash, recipient's address and the amount being transferred. These transactions are recorded on the blockchain, a decentralized and public ledger, ensuring transparency and security in the Bitcoin network. TOR exit node, on the other hand, is the final point in the TOR network (a decentralized network used for anonymous communication) through which encrypted traffic exits to reach its destination. As cryptocurrencies are being increasingly used facilitate illicit transactions, growing need monitor such transactions has risen. The use of the TOR exit node to monitor and investigate illicit Bitcoin users has been discussed in this paper.

***Index Terms* -** Bitcoin transaction analysis, Tor exit node, traffic sniffing, Bitcoin address extraction, script field analysis, clustering algorithms, darkweb investigations, llicit cryptocurrency activity, law enforcement tools.

I. INTRODUCTION

Privacy and anonymity on the Internet are more important than ever. To address this, numerous methods are being implemented to enhance consumers' anonymity whether transacting online or viewing the web. The Tor anonymity network and decentralized cryptocurrency are the most well-known examples of these solutions. The Bitcoin network [2], which enables users to conduct online transactions "pseudonymously," is one of the early examples. More than 100,000 merchants globally accept Bitcoin payments as a result of its popularity [1]. The alleged anonymity of Bitcoin is one of the factors contributing to its appeal as bitcoin is the most popular choice for accepting donations or selling illegal things [4]. With millions of daily active users, the most popular anonymous communication network is Tor [5] [6].Regulators and law enforcement agencies, however, have also encountered challenges because, as Tor and Bitcoin are the essential elements required to accomplish anonymous online transactions with complete operational security [7].

Bitcoin is the most widely used digital currency on the Dark Web [8], and many users continue to choose to use it despite the several studies [9, 10, 11] that have shown that transactions are not as anonymous as previously believed. As shown by Biryukov et al. [3], users utilising Bitcoin over anonymity networks such as Tor are nevertheless vulnerable to deanonymization and man-in-the-middle attacks at the network level. The blockchain provides transparency and immutability. It is a public ledger that documents Bitcoin transactions. All transactions have timestamps, transaction amounts, and sender and receiver addresses. Pseudonymous addresses mean that these addresses are not associated with any one user's identity. By sending Bitcoin-related traffic via the Tor network, this level of anonymity can be further strengthened.

Tor (The Onion Router), a decentralised anonymization network, uses a multi-layered encryption scheme.

It is difficult to determine the source and destination of communication since it is routed through a network of volunteer-run relays. Tor’s anonymity makes it a popular choice for both lawabiding users seeking online privacy and those who use the dark web for illicit purposes.

The potential for cryptocurrencies to be used improperly for illegal transactions necessitates the development of investigative methods to locate and follow criminal actors. This research study explores the feasibility of employing Tor exit node traffic analysis in conjunction with clustering algorithms to identify the individuals involved in questionable Bitcoin transactions. By strategically locating an exit node, we may collect traffic from the anonymous Bitcoin network. We extract relevant data from transaction packets by using this traffic analysis, with a focus on port 8333, which is commonly used for Bitcoin connection. This data, in particular the sender and recipient BTC addresses encoded in the packets’ script field, can yield investigative leads.

With further investigation and commonly available clustering methods, these retrieved addresses can be categorised according to transaction patterns[11]. Through linkages between transactions that don’t seem related at first, this clustering may help identify the persons involved in illicit conduct. We aim to use this innovative approach to assist law enforcement and regulatory bodies tasked with combating unlawful activity made possible by Bitcoin and the dark web with useful information.

This study delves into the technical aspects of setting up a managed Tor exit node setup, analysing Bitcoin traffic using Wireshark, extracting Bitcoin addresses from transaction packets, and using publically accessible clustering techniques. We share our findings about this approach’s effectiveness in identifying suspicious entities and talk about potential drawbacks and future research prospects. With this analysis, we hope to bridge the gap that exists between the anonymity offered by cryptocurrencies and the need for proactive research in the fight against cybercrime.

II. BACKGROUND

The essential background information about Tor and Bitcoin is now discussed:.

A. Bitcoin

The Bitcoin [2] decentralised digital cryptocurrency system eliminates the need for a central bank to regulate money transfers. The Bitcoin network is maintained by a peer-topeer (P2P) network of miners that verify transactions in a decentralised, trustless manner.Over 100,000 retailers throughout the world accept Bitcoin due to its widespread adoption [1]. One of the key selling aspects of bitcoin is that it is purportedly anonymous. By employing pseudonyms as transaction addresses in Bitcoin, users can hide their true identities. A Bitcoin address is a 160-bit address produced by a digital signature technique. Within a person’s wallet are their private and public keys. Transaction inputs are signed using private keys to establish ownership.

B. Tor and Tor Exit Nodes

TOR network is used for anonymous communication, which routes internet traffic over a decentralized network of relay nodes. Tor [5] is the most widely used anonymous communication network. The origin and destination of communications are anonymous, making it challenging to follow user online activities. To creates a circuit to reach its destination, Tor goes via several relay nodes . Only the address of the subsequent relay in the circuit is known after each relay decrypts its layer of encryption. This procedure ensures that the source and destination IP addresses stay anonymous, providing users with a certain level of anonymity.

The point where traffic leaves the Tor network and enters into the open internet is known as a Tor exit node. When a website is accessed, the IP address of exit node is visible as the source IP address to the website and real IP is not known. This increases user privacy, but this also has few drawbacks. As the last point of Tor communication, the exit node may be used to track the content of the traffic that is not encrypted when it exits the Tor network. Because of this vulnerability, using Tor exit nodes should be done with extreme caution, especially while engaging in sensitive operations. It should be noted that exit nodes can be exploited by bad actors to carry out illicit activities, hence it should be used cautiously.

III. LITERATURE REVIEW

Because of their increasing popularity, cryptocurrencies like Bitcoin have been the subject of much research into their technological and economic implications [9,10,11,12,13]. It is imperative to conduct study on the potential illicit use of Bitcoin on the dark web. This research gap necessitates looking into ways to identify and monitor people involved in such transactions because of the intrinsic anonymity of Bitcoin and the obfuscation offered by anonymization networks like Tor.

This overview of the literature expands on the work that has already been done on Bitcoin transaction analysis, Tor network traffic analysis, and the use of clustering techniques to anomaly identification in cryptocurrency transactions. Critical analysis of these study areas will hopefully yield a framework for our suggested methodology as well as an understanding of the benefits and drawbacks of the current methodologies. We will discuss how current research contributes to our understanding of the anonymity of Bitcoin transactions, the difficulties in searching through Tor network traffic for activity related to Bitcoin, and how well clustering algorithms work to spot patterns that point to illegal activity within the Bitcoin ecosystem. This review will provide a comprehensive understanding of the current state of affairs and pave the way for the development of a state-of-the-art investigative approach that allows for the identification of actors in dubious Bitcoin transactions through the use of traffic analysis from Tor exit nodes. We will examine a number of important research studies that tackle these important topics in the sections that follow. We will examine their research design, results, and constraints in order to obtain important information that will guide the creation of our suggested strategy. This paper will contribute to the progress of our knowledge on the topic and provide a new perspective on the examination of BTC transactions and the research of criminal activities. Researchers examining activity inside the Bitcoin ecosystem are finding that network traffic analysis is a valuable tool. Packet sniffing tools like Wireshark are effective tools for recording and analysing Bitcoin network behaviour, as demonstrated by Turner and Irwin (2017) [24]. Through the use of this technique, anyone can spot patterns associated with Bitcoin addresses and transactions, potentially leading to the discovery of signs of dubious financial activities.

Shirazi et al. (2015)[25] examined how Tor, a well-known anonymity network with over 2 million daily users, is crucial for safeguarding online privacy. In order to create privacyenhancing solutions like Tor, extensive experimentation is required to evaluate the effects of software modifications, test attacks, and analyze network data. Numerous methods have been developed in response to the risks and limitations of conducting research on the active Tor network, such as simulation, emulation, and small-scale private Tor networks.

Tor experimentation heavily relies on simulation and emulation, which allow researchers to test and model various scenarios without affecting the live network. Through the use of simulation tools such as Shadow, TorPS, and COGS, researchers can examine the security and functionality of the Tor network in a controlled setting. These simulators are a useful substitute for real network testing, allowing researchers to investigate a range of topics pertaining to the anonymity and functionality of Tor. Comparing various Tor experimental tools is difficult due to Tor's distinctive features as an anonymity network, particularly when network simulation is taken into consideration. It is particularly challenging to assess simulation tools since accurate statistics on user activity and traffic patterns on the operational Tor network are lacking. Despite these challenges, researchers can choose the tool that best suits their specific research needs by consulting an extensive overview of simulation approaches and Tor experimentation tools. Evaluating simulation tools becomes more challenging when trustworthy statistics on user activity and traffic patterns on the real Tor network are not available. Despite these obstacles, researchers can consult a thorough overview of simulation methodologies and Tor experimentation tools to determine which tool is best fit for their particular study requirements. Since reliable statistics on user activity and traffic patterns on the real Tor network are unavailable, assessing simulation tools becomes more difficult. Despite these difficulties, researchers can choose the best tool depending on the needs of their particular study by checking an extensive list of Tor experimental tools and simulation techniques. The assessment of simulation tools is made more difficult by the absence of precise statistics on user behavior and traffic patterns in the active Tor network. Despite these difficulties, researchers can choose the best tool for their particular research needs with the help of a thorough overview of simulation techniques and tools used for Tor experimentation.

Harlev et al. (2018) [26] proposed a method for reducing anonymity on the Blockchain with the use of supervised machine learning techniques. This study predicts the category of unknown clusters with an accuracy of 77% and an F1score of roughly 0.75 using the Gradient Boosting Classifier. It also disproves the common belief that Bitcoin transactions are extremely anonymous. A unique method for enhancing transparency and understanding the transaction structure of the cryptocurrency ecosystem is the de-anonymization of the Bitcoin Blockchain via the application of Supervised Machine Learning. [26] claims that advanced classification techniques can be applied to Blockchain data to uncover crucial details about the transactions and the involved parties .

It is essential to comprehend transaction patterns in order to analyze trends in Bitcoin usage. Vlahavas et al.(2024) [27] use unsupervised clustering analysis to explore this area. Their investigation looks into:

- Transaction Network Analysis: The study explores hidden patterns and behaviors within Bitcoin transactions by examining the blockchain as a network.
- Clustering Techniques: Algorithms for clustering and dimensionality reduction are used to group transactions with comparable attributes. This approach provides insights into user behavior and evolving patterns of cryptocurrency usage.

For increased accuracy and assessment of unsupervised clustering algorithms, the authors recommend more research into supervised machine learning techniques.

Deanonymizing Bitcoin transactions can also be accomplished by analysing network traffic. To accomplish this, Tian et al. (2022) [28] present a unique method termed NTSSL, which makes advantage of semi-supervised learning. This method uses the following strategies to get over the limitations of traditional unsupervised approaches:

- Improved Performance: NTSSL achieves higher accuracy with less resource needs when compared to unsupervised techniques.
- Transaction Clustering: The technique employs transaction clustering to further boost the effectiveness of deanonymization of actor by grouping transactions originating from the same Bitcoin node.

IV. METHODOLOGY

This strategy discusses a research methodology that utilizes Tor network environment and evaluate potential methods for identifying illicit actors involved in the bitcoin ecosystem operating on dark web. This strategy may be divided into eight stages(see Fig. 1)

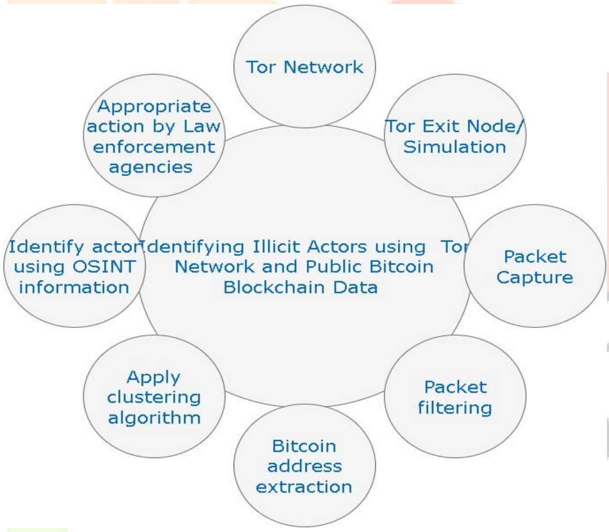

Tor Network
Tor Exit Node/ Simulation
Packet Capture
Packet filtering
Bitcoin address extraction
Apply clustering algorithm
Identify actor using OSINT information
Appropriate action by Law enforcement agencies
Identifying Illicit Actors using Tor Network and Public Bitcoin Blockchain Data

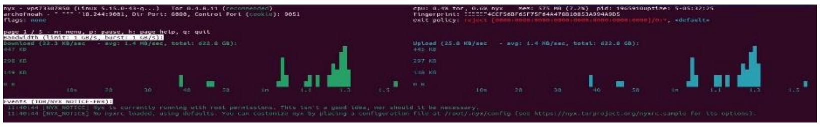
Fig. 1. Stages of this Methodology

1. Tor Network Exit Node or Simulation Setup : We can setup Tor Exit Node on a Linux machine by installing tor, using the ”apt install tor” command, or we can make use of a regulated and authorized Tor network simulation environment[36,37]. After installation, the tor configuration file is modified allow port 8333 traffic to flow through it, a detailed TOR exit node setup process is discussed here[44]. We can use generate this data pragmatically or extract data from port 8333 traffic from Exit Node.
2. Traffic capture and preprocessing: Capture Tool: To record traffic passing via the specified exit node in your exit node/simulated environment, use a network traffic capture tool such as wireshark[38]. Data Filtering: Specifically target port 8333, which is frequently used for Bitcoin connections, by filtering the traffic that has been captured.
3. Transaction packet extraction and analysis: Packet identification is the process of extracting pertinent Bitcoin transaction packets from the traffic data collected. Applications like [35] or any other custom scripts could be useful to parse the packets and obtain sender and recipient bitcoin addresses, IP addresses, etc.
4. Public Blockchain Data Integration: Public Blockchain Access: Use public blockchain explorers with APIs (such as [39,40]) to access historical Bitcoin transaction data. Address Enrichment: To query the public blockchain data, use the anonymized addresses that you retrieved from your traffic data. This could entail comparing addresses or transaction hashes.

5. Clustering and Analysis: Clustering Techniques: Group addresses with comparable transactions from public blockchain data by using clustering algorithms (e.g., as discussed in the literature review) or public algorithms(like [34]) . After generating cluster of target user, check for open source intelligence or publicly available information(like [41,42,43]) and if it matches with any of the given cluster addresses, then we can say we have identified the actor operating on dark web.

V. RESULTS

The results of our study on identifying illegal actors with publicly accessible blockchain data and Tor network are shown in this section. The approach concentrated on recording Bitcoin traffic on a specified exit node(see Fig. 2) in a regulated Tor network setting.

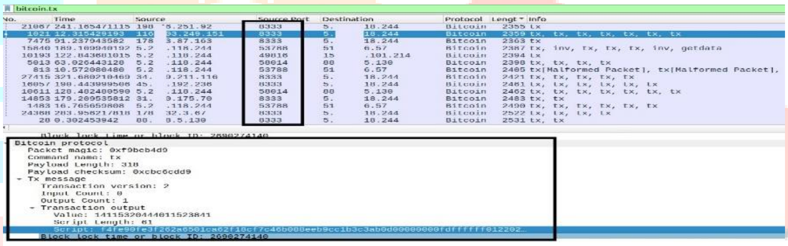


```
nyx - vp673387850 (Linux 5.15.0-43-g...)   Tor 0.4.8.11 (recommended)          cpu: 0.4% tor, 0.0% nyx    mem: 575 MB (7.2%)  pid: 1965910uptime: 5-05:32:25
archefnoah - [illegible]10.244:9001, Dir Port: 0000, Control Port (cookie): 9051   fingerprint: ......74CCF56BF65F75F64A47801085 3A994A9D5
flags: none                                                                      exit policy: reject [0000:0000:0000:0000:0000:0000:0000:0000]/0:*, <default>

page 1 / 5 - m: menu, p: pause, h: page help, q: quit
Bandwidth (limit: 1 GB/s, burst: 1 GB/s):
Download (22.3 KB/sec   - avg: 1.4 MB/sec, total: 622.8 GB):                     Upload (25.8 KB/sec   - avg: 1.4 MB/sec, total: 622.8 GB):

Events (TOR/NYX NOTICE-ERR):
 11:40:44 [NYX_NOTICE] Nyx is currently running with root permissions. This isn't a good idea, nor should it be necessary.
 11:40:44 [NYX_NOTICE] No nyxrc loaded, using defaults. You can customize nyx by placing a configuration file at /root/.nyx/config (see https://nyx.torproject.org/nyxrc.sample for its options).
```

Fig. 2. Tor Exit Node

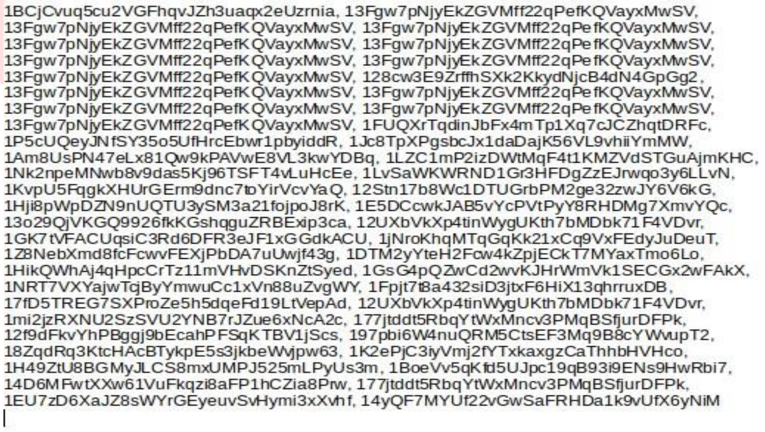
Traffic capture and address extraction:

With Wireshark, we were able to successfully record Bitcoin traffic(Fig. 3) on port 8333 in the Tor network environment that was under control. We were able to extract transaction packets (tx packets) from the traffic that was recorded by using the proper tools. Then, we were able to extract the sender and recipient addresses for the anonymized Bitcoin by examining the script field in these transmitted packets. This proves that pertinent Bitcoin transaction data can be obtained and extracted from Tor network traffic.(Fig. 4)



```
bitcoin.tx
No.    Time              Source          Source Port  Destination       Protocol  Lengt  Info
 21067 241.165471115  198  '5.251.92      8333    5.     18.244      Bitcoin   2355 tx
  1021 12.315429193   116  93.249.151     8333    5.     18.244      Bitcoin   2359 tx, tx, tx, tx, tx, tx, tx
  7475 91.237943582   178  3.87.163       8333    5.     18.244      Bitcoin   2363 tx
 15840 189.109940192  5.2  .118.244       53788   51     6.57        Bitcoin   2387 tx, inv, tx, tx, tx, inv, getdata
 10193 122.843681015  5.2  .118.244       49816   15     .101.214    Bitcoin   2394 tx
  5013 63.026443128   5.2  .118.244       58014   88     5.130       Bitcoin   2398 tx, tx, tx, tx
   813 10.572080480   5.2  .118.244       53788   51     6.57        Bitcoin   2405 tx[Malformed Packet], tx[Malformed Packet],
 27415 321.680210469  34.  9.211.116      8333    5.     18.244      Bitcoin   2421 tx, tx, tx, tx, tx
 16057 190.443999506  45.  .192.236       8333    5.     18.244      Bitcoin   2461 tx, tx, tx, tx, tx, tx
 10611 128.402480590  5.2  .118.244       58014   88     5.130       Bitcoin   2462 tx, tx, tx, tx, tx, tx, tx
 14853 179.209535812  31.  9.175.70       8333    5.     18.244      Bitcoin   2483 tx, tx
  1483 16.765659808   5.2  .118.244       53788   51     6.57        Bitcoin   2490 tx, tx, tx, tx, tx, tx
 24368 283.958217818  178  32.3.67        8333    5.     18.244      Bitcoin   2522 tx, tx, tx, tx
    28 0.302453942    88.  8.5.130        8333    5.     18.244      Bitcoin   2531 tx, tx

    Block lock time or block ID: 2690274140
Bitcoin protocol
  Packet magic: 0xf9beb4d9
  Command name: tx
  Payload Length: 318
  Payload checksum: 0xcbc6cdd9
 - Tx message
    Transaction version: 2
    Input Count: 0
    Output Count: 1
  - Transaction output
      Value: 14115320444011523841
      Script Length: 61
      Script: f4fe90fe3f262a6501ca62f18cf7c46b088eeb9cc1b3c3ab0d00000000fdffffff012202…
    Block lock time or block ID: 2690274140
```

Fig. 3. Bitcoin Traffic Capture on Port 8333



1BCjCvuq5cu2VGFhqvJZh3uaqx2eUzrnia, 13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV,
13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV, 13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV,
13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV, 13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV,
13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV, 13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV,
13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV, 128cw3E9ZrffhSXk2KkydNjcB4dN4GpGg2,
13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV, 13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV,
13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV, 13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV,
13Fgw7pNjyEkZGVMff22qPefKQVayxMwSV, 1FUQXrTqdinJbFx4mTp1Xq7cJCZhqtDRFc,
1P5cUQeyJNfSY35o5UfHrcEbwr1pbyiddR, 1Jc8TpXPgsbcJx1daDajK56VL9vhiiYmMW,
1Am8UsPN47eLx81Qw9kPAVwE8VL3kwYDBq, 1LZC1mP2izDWtMqF4t1KMZVdSTGuAjmKHC,
1Nk2npeMNwb8v9das5Kj96TSFT4vLuHcEe, 1LvSaWKWRND1Gr3HFDgZzEJrwqo3y6LLvN,
1KvpU5FqgkXHUrGErm9dnc7toYirVcvYaQ, 12Stn17b8Wc1DTUGrbPM2ge32zwJY6V6kG,
1Hji8pWpDZN9nUQTU3ySM3a21fojpoJ8rK, 1E5DCcwkJAB5vYcPVtPyY8RHDMg7XmvYQc,
13o29QjVKGQ9926fkKGshqguZRBExip3ca, 12UXbVkXp4tinWygUKth7bMDbk71F4VDvr,
1GK7tVFACUqsiC3Rd6DFR3eJF1xGGdkACU, 1jNroKhqMTqGqKk21xCq9VxFEdyJuDeuT,
1Z8NebXmd8fcFcwvFEXjPbDA7uUwjf43g, 1DTM2yYteH2Fcw4kZpjECkT7MYaxTmo6Lo,
1HikQWhAj4qHpcCrTz11mVHvDSKnZtSyed, 1GsG4pQZwCd2wvKJHrWmVk1SECGx2wFAkX,
1NRT7VXYajwTqByYmwuCc1xVn88uZvgWY, 1Fpjt7t8a432siD3jtxF6HiX13qhrruxDB,
17fD5TREG7SXProZe5h5dqeFd19LtVepAd, 12UXbVkXp4tinWygUKth7bMDbk71F4VDvr,
1mi2jzRXNU2SzSVU2YNB7rJZue6xNcA2c, 177jtddt5RbqYtWxMncv3PMqBSfjurDFPk,
12f9dFkvYhPBggj9bEcahPFSqKTBV1jScs, 197pbi6W4nuQRM5CtsEF3Mq9B8cYWvupT2,
18ZqdRq3KtcHAcBTykpE5s3jkbeWvjpw63, 1K2ePjC3iyVmj2fYTxkaxgzCaThhbHVHco,
1H49ZtU8BGMyJLCS8mxUMPJ525mLPyUs3m, 1BoeVv5qKfd5UJpc19qB93i9ENs9HwRbi7,
14D6MFwtXXw61VuFkqzi8aFP1hCZia8Prw, 177jtddt5RbqYtWxMncv3PMqBSfjurDFPk,
1EU7zD6XaJZ8sWYrGEyeuvSvHymi3xXvhf, 14yQF7MYUf22vGwSaFRHDa1k9vUfX6yNiM

Fig. 4. Bitcoin address extracted from sample traffic

Cluster Analysis :

We looked into using different clustering programs that are available to the public and can be accessed on websites like GitHub and found results for few extracted addresses(see Fig. 5,6). However, within the scope of this study, a thorough assessment of clustering efficacy for detecting illicit actors was not feasible due to the limited data.

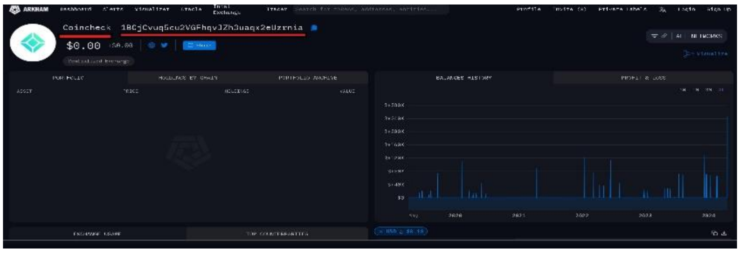
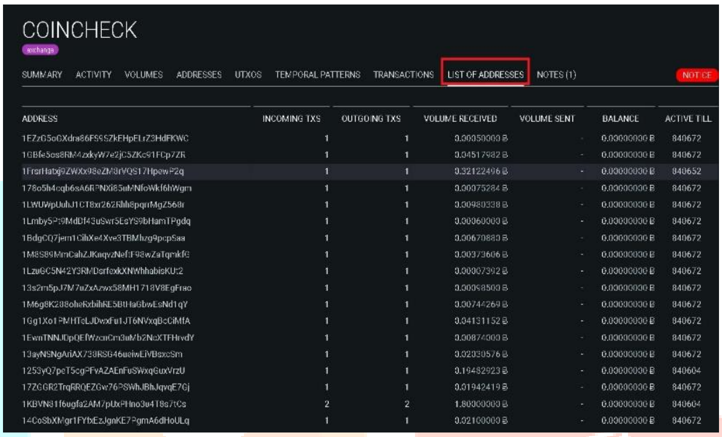


Fig. 5. Correlation of extracted bitcoin traffic with open source information



ADDRESS	INCOMING TXS	OUTGOING TXS	VOLUME RECEIVED	VOLUME SENT	BALANCE	ACTIVE TILL
1EZzG5oGXdra86FS9SZkEHpELrZ3HdFKWC	1	1	0.00050000 B	-	0.00000000 B	840672
1GBfe5os8RM4zxkyW7e2jC5ZKc91FCp7ZR	1	1	0.04517982 B	-	0.00000000 B	840672
1FrsrHatxj9ZWXx98eZM8rVQS17HpewP2q	1	1	0.32122496 B	-	0.00000000 B	840652
178o5h4cqb6sA6RPNXi85uMNfoWkf6hWgm	1	1	0.00075284 B	-	0.00000000 B	840672
1LWUWpUuhJ1CT8xr262Rhh8pqnMgZ568r	1	1	0.00980338 B	-	0.00000000 B	840672
1Lmby5Pt9MdDf43uSwr5EsYS9bHamTPgdq	1	1	0.00060000 B	-	0.00000000 B	840672
1BdgCQ7jem1CihXe4Xve3TBMhzg9pcpSaa	1	1	0.00670880 B	-	0.00000000 B	840672
1M8S89MmCahZJKnqvzNeftF98wZaTqmkfG	1	1	0.00373606 B	-	0.00000000 B	840672
1LzuGC5N42Y3RMDsrfexkXNWhhabisKUt2	1	1	0.00007392 B	-	0.00000000 B	840672
13s2m5pJ7M7uZxAzwx58MH1718V8EgFrao	1	1	0.00098500 B	-	0.00000000 B	840672
1M6g8K288oheRxbihRE5BtHaGbwEsNd1qY	1	1	0.00744269 B	-	0.00000000 B	840672
1Gg1Xo1PMHTcLJDwxFu1JT6NVxqBcCiMfA	1	1	0.04131152 B	-	0.00000000 B	840672
1EwnTNNJDpQEfWzcnCm3uMb2NcXTFHrvdY	1	1	0.00874000 B	-	0.00000000 B	840672
13ayNSNgAriAX738RSG46ueiwEiVBsxcSm	1	1	0.02030576 B	-	0.00000000 B	840672
1253yQ7peT5cgPFvAZAEnFuSWxqGuxVrzU	1	1	0.19482923 B	-	0.00000000 B	840604
17ZGGR2TrqRRQEZGw76PSWhJBhJqvqE7Gj	1	1	0.01942419 B	-	0.00000000 B	840672
1KBVN81f6ugfa2AM7pUxPHno3u4T8s7tCs	2	2	1.80000000 B	-	0.00000000 B	840604
14CoSbXMgr1FYfxEzJgnKE7PgmA6dHoULq	1	1	0.02100000 B	-	0.00000000 B	840672

Fig. 6. More cluster addresses were found

All things considered, the outcomes validate the ability to record and retrieve Bitcoin address data from Tor network traffic and correlate the retrieved BTC address with open source information to identify the entity. To assess the efficacy of this strategy in identifying criminal actors within the cryptocurrency ecosystem, more study using real-world is required.

VI. CONCLUSION

This study investigated the possibility to identify the people who were involved in questionable Bitcoin transactions on dark web by using clustering techniques in combination with traffic analysis from Tor exit nodes.This technique includes capturing of bitcoin traffic on TOR exit node, bitcoin address extraction from the captured traffic and identification of involved actor using clustering techniques. We showed how this method could yield useful leads for investigations by carefully setting up a Tor exit node, recording Bitcoin traffic, and extracting related data from transaction packets(see Flowchart Fig. 7).The extracted bitcoin addresses after clustering, if found to be belonging to a threat actor, this shows the possibility of de-anonymize bitcoin user operating on Tor network. Through concentrating on port 8333 and obtaining bitcoin addresses from transaction packets, we managed to obtain essential information for additional examination. By grouping these addresses according to transaction patterns, easily accessible clustering algorithms helped identify relationships between seemingly unrelated bitcoin addresses. Regulatory and law enforcement organizations may find this capacity to link potentially suspicious bitcoin addresses to one another.

Notwithstanding these drawbacks, the study provides a useful avenue for further research on Bitcoin transactions and dark web activities. This strategy may narrow the gap between the anonymity offered by cryptocurrencies on the dark web and the need for proactive investigation in the fight against cybercrime. With more research and enhancements, law enforcement agencies may find considerable success using this tactic in the battle against illicit activity.

VII. LIMITATIONS AND CHALLENGES

It is imperative to recognise the constraints of this methodology. The quantity and quality of the extracted data have a significant impact on how well clustering algorithms performs. Setting up a reliable and secure TOR exit node requires technical expertise and caution. It is advised to set up the TOR exit node on a VPS server and not on our personal computers as it exposes your IP to the internet. Another limitation of this study is that, it is done using a single TOR exit node and in real world scenario thousands of exit nodes would be required to run which involves cost and human resources. Moreover, after the clustering step we may find the address could belong to a legit crypto user, crypto service or unidentified crypto address as well. Also, capturing hundred percent bitcoin traffic flowing through the TOR network is a big challenge. To evaluate the potential of this technique to identifying illegal actors on darkweb using bitcoin for transactions, more real world implementations are required.

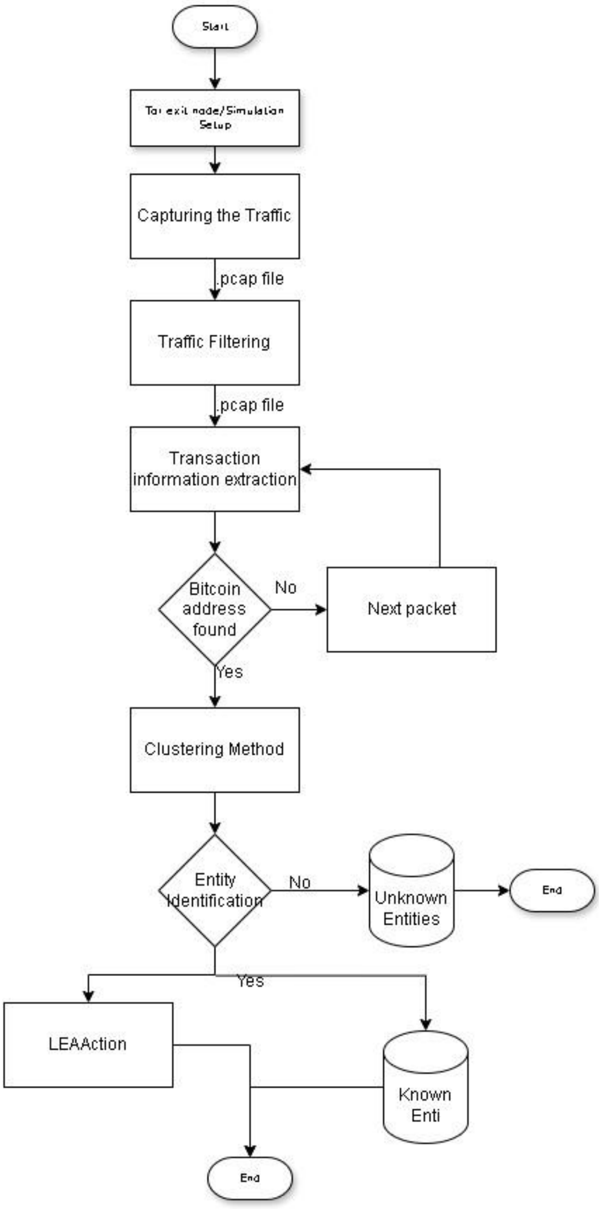


Fig. 7. Methodology Flowchart

VIII. REFERENCE

1) Anthony Cuthbertson. 2015. Bitcoin now accepted by 100,000 merchants worldwide. International Business Times. (2015).

2) Satoshi Nakamoto. 2008. Bitcoin: A peer-to-peer electronic cash system. (2008).

3) Alex Biryukov and Ivan Pustogarov. 2015. Bitcoin over Tor isn't a Good Idea. In 2015 IEEE Symposium on Security and Privacy, SP 2015.

4) Alex Biryukov, Ivan Pustogarov, Fabrice Thill, and

Ralf-Philipp Weinmann. 2014. Content and Popularity Analysis of Tor Hidden Services. In 34th International Conference on Distributed Computing Systems Workshops (ICDCS 2014 Workshops).

5) Roger Dingledine, Nick Mathewson, and Paul Syverson. 2004. Tor: The Second- Generation Onion Router. In Proceedings of the 13th USENIX Security Symposium. 303–320.

6) The Tor Project. 2017. Tor Metrics Portal. https://metrics.torproject.org. (2017).

7) Vincent Van Mieghem and Johan Pouwelse. 2015. Anonymous online purchases with exhaustive operational security. CoRR abs/1505.07370 (2015)

8) Michael del Castillo. 2016. Bitcoin Remains Most Popular Digital Currency on Dark Web. CoinDesk. (2016).

9) Jules DuPont and Anna Cinzia Squicciarini. 2015. Toward De-Anonymizing Bitcoin by Mapping Users Lo-
cation

10) Michael Fleder, Michael S Kester, and Sudeep Pillai. 2015. Bitcoin transaction graph analysis. arXiv (2015).

11) Sarah Meiklejohn, Marjori Pomarole, Grant Jordan, Kirill Levchenko, Damon McCoy, Geoffrey M. Voelker, and Stefan Savage. 2013. A Fistful of Bitcoins: Characterizing Payments Among Men with No Names.

12) Fergal Reid and Martin Harrigan. 2013. Security and Privacy in Social Networks.

13) Al Jawaheri, H., Riley, R., Srivastava, J., et al. (2019). Deanonymizing Tor Hidden Service Users Through Bitcoin Transactions Analysis.

14) Philip Koshy, Diana Koshy, and Patrick Mcdaniel. 2014. An Analysis of Anonymity in Bitcoin Using P2P Network Traffic. Financial

15) Vincent Van Mieghem and Johan Pouwelse. 2015. Anonymous online purchases with exhaustive operational security.

16) Kyle Torpey. 2016. Darknet Customers Are Demanding Bitcoin Alternative Monero.

17) P. Reynolds and A. S.M. Irwin, "Tracking digital footprints: anonymity within the bitcoin system"

18) C. Zhao, "Graph-based forensic investigation of bitcoin transactions," 2014.

19) P. Koshy, D. Koshy, and P. McDaniel, "An analysis of anonymity in bitcoin using p2p network traffic"

20) Chainalysis, "Protecting the integrity of digital assets." www.chainalysis.com

21) Bitcoin.org. n.d.. BitcoinCore. (n.d.). https://bitcoin.org/en/bitcoin-core/

22) Bitpay. 2016. BitcoreNode. (Oct 2016). https://github.com/bitpay/bitcore-node

23) Moser M. 2013. Anonymity of bitcoin transactions: An analysis of mixing services. Munster Bitcoin Conference (2013).

24) Turner, A., & Irwin, A. S. M. (2017). Bitcoin Transactions: A digital discovery of illicit activity on the blockchain. Journal of Digital Forensics.

25) Shirazi, F., Goehring, M., & Diaz, C. (2015). Tor Experimentation Tools.

26) Harlev, M. A., Sun Yin, H., Langenheldt, K. C., Mukkamala, R., and Vatrapu, R. (2018). "Breaking bad: Deanonymising entity types on the bitcoin blockchain using supervised machine learning".

27) Vlahavas, G., Karasavvas, K., & Vakali, A. (2024). Unsupervised clustering of bitcoin transactions. Financial Innovation, 10(25).

28) Tian, C., Ge, Y., Shi, R., Liang, Y., Lan, L., Liu, P., & Peng, Z. (2022). Deanonymization of Bitcoin transactions based on network traffic analysis with semisupervised learning.

29) Biryukov, A. (2015). Crawling for Tor Hidden Services: Detection, Measurement, Deanonymization. In Proceedings of IEEE Symposium on Security and Privacy (SP'13). IEEE Computer Society.

30) Tovanich, N., & Cazabet, R. (2023). Fingerprinting Bitcoin entities using money flow representation learning. Applied Network.

31) Turner, A. B., McCombie, S., & Uhlmann, A. J.
(2020). Analysis Techniques for Illicit Bitcoin Transactions. Frontiers in Computer Science, 2, 600596. doi: 10.3389/fcomp.2020.600596

32) Antonopoulos A (2010) Mastering Bitcoin. USA: O'Reilly Media.

33) Biryukov A, Tikhomirov S. Deanonymization and linkability of cryptocurrency transactions based on network analysis.

34) https://github.com/thomasverweij/bitcoin-addresscluster

35) https://github.com/donutAnees/$sih_{d}eanonymization_{b}itcoin$

36) http://torflow.uncharted.software/

37) https://mininet.org

38) https://www.wireshark.org/

39) https://blockchair.com/

40) https://www.blockcypher.com/

41) https://oxt.me/

42) https://www.arkhamintelligence.com/

43) https://www.bitcoinabuse.com/

44) https://medium.com/@kee$_{9}$4942/*how*-i−*ran*the−*second*-best−*tor*-exit−*node*-f3c20de086e2/