An Investigation into the Relationship Between Cryptocurrencies, Anonymisation Networks, and the Darknet Cybercrime Scene

Angelos Velissaratos

Thesis to obtain the Master of Science Degree in

Information Security and Cyberspace Law

Supervisors:
Prof. Nuno Santos
Eng. Nelson Escravana

Examination Committee

Chairperson:	Prof. Paulo Mateus
Supervisor:	Prof. Nuno Santos
Member of the Committee:	Prof. Miguel Correia

June 2018

Acknowledgements

I would like to thank all the people who helped me completing this thesis.

Lisboa, June 2018
Angelos Velissaratos

A7e8aCYpjW2jGSMrfczVDGe+5g
/oGLygx+o1aWFEqqrGn/p4B77M
uPUIvQQzZp9nIqXXnaAaSho88I
oDe10X5SNX1ex7I/UeT2Pt/3mr
P/BeE7+dl+EufvKR6wvU1nMwWI
7uSijcfbSXoRsITQ1hoUjZH5wN
BQKrIrJJ+7TFDmBGZMVDYIHKOo
bazaH6JLU3l9Ok5PHI+N/tdSLM
B8tbCc7bvf/FlAr7/PJy9tRiru
WevEqRC2ku14P6iKGINjzdABD+
5qdtLKQuHZH4gDycBe/2EnusS1
ld0UtmeHuL+2gNNfxD5CS70VPY
C+T3KhMqRA1UCc4JlDdm0mZ7fr
7b+ncGfPDgQH7cZlLioga5Hh1C
hz9qZ70HNxRs5kgigNVjWOsU4c
khaTQKrUtHTEEA64B2fvvYaMLx
DQ6VYfGnsFADi9QHjnxystT1VX
bcLWl8m30+B0P9XHS0lVup5Vwi
qrY+EPd9LzV2JhwoamzGmuwrhm
A6NfHhS69JvSdW2aNU

... how can you challenge A perfect, immortal Machine?

Resumo

Nos últimos anos, os avanços nas técnicas e tecnologias de anonimato têm permitido que criminosos realizem os seus negócios online e os disponibilizem a todos aqueles que possam instalar software de uso gratuito, permitindo ao mesmo tempo que os ditos criminosos permaneçam ocultos do olhar vigilante das autoridades policiais à escala mundial. O objetivo deste trabalho é examinar a rede de privacidade do Tor e a cripto-moeda Bitcoin e documentar, por um lado, como estes sistemas são usados atualmente para facilitar o crime na dark web e, por outro, como as autoridades legais pode localizar artefatos digitais que poderão ser admissíveis como prova em tribunal caso esses sistemas sejam usados de forma ilegal. Tendo em conta que esta dissertação aborda muitos sistemas diferentes, não é pressuposto que o leitor tenha conhecimento prévio de eles. Portanto, os capítulos que compõem a primeira parte deste documento fornecem uma introdução breve, mas necessária, de todos os principais componentes destes sistemas, seguida de um exame mais aprofundado. Na segunda parte, por forma a auxiliar as agências de segurança pública e os escritórios de procuradoria, é introduzida uma série de diretrizes com base principalmente na literatura existente nos campos das vulnerabilidades Tor e Bitcoin e nas técnicas de ciber-segurança forense que poderiam potencialmente auxiliar agentes de investigação com autorização legal e competência técnica. Em última análise, esta dissertação visa ajudar as autoridades a atingir dois objetivos: identificar as transações do Bitcoin que podem ter feito parte de transações ilegais e ajudá-las a descobrir artefatos forenses nos sistemas em que Bitcoin e Tor terão sido usados.

Abstract

In recent years, advancements in anonymisation techniques and technologies have allowed criminals to move their business online and make it available to everyone who can install and use freely available software, while at the same time they can remain hidden from the watchful gaze of law enforcement on a worldwide scale. The goal of this work is to examine the Tor privacy network and the Bitcoin cryptocurrency, and document on how they are currently used to facilitate crime in the dark web and how law enforcement can locate digital artefacts that could potentially be admissible as evidence in a court of law, if those systems are ever used in an unlawful manner. Taking into account that this thesis deals with many different systems and concepts, it does not take for granted that an individual with knowledge of one will be able to understand all of them. Therefore, the chapters that make up the first part of this document provide prospective readers with a brief but necessary introduction to all the major components followed immediately with a deeper examination. In the second part, in order to assist law enforcement agencies and prosecutorial offices, a series of guidelines is introduced based mostly on existing literature in the fields of Tor and Bitcoin vulnerabilities and cyber forensic techniques that could potentially allow an individual with the legal and technical competence to investigate a case. Ultimately, this dissertation aims to aid law enforcement agencies in achieving two objectives: de-anonymising Bitcoin transactions that may have been part of an unlawful exchange and aid them to uncover forensic artefacts on the systems that Bitcoin and Tor have been used on.

Palavras-Chave
Keywords

Palavras-Chave

Tor

Bitcoin

Darknet

Anonimato

Cibercrime

Ciber-segurança Forense

Keywords

Tor

Bitcoin

Darknet

Anonymity

Cybercrime

Cyber Forensics

Contents

List of Figures

List of Tables

Acronyms

API Application Programming Interface

BTC Bitcoin

DNM Darknet Marketplace

IXP Internet Exchange Point

HTTP HyperText Transfer Protocol

HTTPS HyperText Transfer Protocol (Secure)

LEA Law Enforcement Agencies

OP Onion Proxy

OR Onion Router

OS Operating System

PT Pluggable Transport

SSL Secure Sockets Layer

TCP Transmission Control Protocol

TLS Transport Layer Security

Tor The Onion Router

VPN Virtual Private Network

1 Introduction

The aim of this dissertation is to examine the Bitcoin cryptocurrency and the Tor privacy network and the different methods by which they can be used to facilitate criminal activities in the dark net. This thesis attempts to act as a comprehensive guide as to how an individual with both the legal and technical competence can attempt to pursue a forensics investigation in scenarios where Tor, Bitcoin or both in conjunction have been used by one or many individuals either as a means to commit a crime or as an accidental consequence if those systems were the targets of cybercrime themselves. For that purpose, it presents an up-to-date list of known vulnerabilities that are affecting both Tor and Bitcoin and methods by which one can extract digital artefacts from computer systems related to both technologies that could lead to de-anonymising their users and as a consequence aid in an investigation.

1.1 Motivation

In recent years there has been an increase in the activity in the *dark net*, the section of the Internet that is not accessible normally without some sort of special software, in the form of concealed, digital marketplaces that are designed to privately deal with the selling of drugs and contraband medication, weapons, child pornography, among other things. The latest advancements in methods in which Internet users can conceal their true identities from the world and more importantly from the eyes of Law Enforcement Agencies on a worldwide scale helped in facilitating this trend [31]. Although techniques to provide anonymity while navigating public network are not a new concept and have in fact been around for quite some time, the advent of the Tor anonymity network with its ease of use and simplicity in the ways one can gain access to it and take advantage of its privacy-enabling features has allowed the facilitation of illicit storefronts in the bowels of the deep web. The Tor network and more importantly all the tools to gain access to it, including its official browser, are available to nearly every Internet user and are easily acquired from the so called clear web with little to no effort.

A close partner to Tor is a form of currency that is being used in order to facilitate the buying and selling of illicit goods online and that is the cryptocurrency that goes by the name Bitcoin or BTC for short. Since its release as open-source software back in 2009, this peer-to-peer system of transactions has created a new financial paradigm that nations and international markets are still trying to make sense of. Bitcoin and almost all other kinds of widely-circulated cryptocurrencies have one unique trait in common: they are not being issued by a centralised authority, such as a central bank or (supra) national governmental body, but instead are being created through a process that is referred to as *mining* [36]. Another unique aspect of Bitcoin is that transactions do not have to go through a bank or other financial network, but instead transfer from one party to another. Due to its reliance on cryptographic protocols in order to facilitate those transactions, Bitcoin offers at the very least a basic level of anonymity in

transactions. Despite the fact that Bitcoin is relatively recent, a lot of research went into developing techniques that allow for de-anonymisation of transactions. Some of them, such as Reid and Harrigan's 2013 research into the subject [39] proved that anonymity is not guaranteed and that it is possible to link public keys (meaning users' wallets) with external identifying information, such as databases with recorded transactions.

When it comes to the matter of user anonymity in the Tor and Bitcoin networks the situation is further complicated if privacy-conscious users utilise techniques in order to further conceal their transactions and further remove the links that connect them to their activities. There is a plethora of options available, that can either be used individually or in conjunction with one another: Bitcoin mixers (or tumblers), online services that act as laundering services [33, 35, 1], and of course there is the combination of cryptocurrencies and anonymisation networks such as Tor. The latter is one of the most difficult situations to tackle, due to the fact that it provides Internet users with a way in which they can further remove themselves from their online activities by introducing several intermediate relay stations, that further confuse the digital tracks of a user.

1.2 Goals

This document was written with two very specific purposes in mind:

- As an attempt to complement existing literature in the fields of Tor, Bitcoin, Darknet Cryptomarkets by examining how those aforementioned systems work together in cyber-criminal activities.
- Act as a guide to help forensic investigators to locate digital artefacts generated by Tor and Bitcoin that could potentially be used as evidence in a court of law.

In addition, this dissertation is written for an audience with possibly no prior knowledge of the aforementioned systems and technologies. For this reason, it provides a very basic introduction for multiple aspects surrounding the the Bitcoin and Tor systems, the dark net and cybercrime.

Note that this dissertation was not created with mobile or cloud computing forensic investigations in mind and focuses solely on traditional PC systems running some recent version of Microsoft Windows, Apple macOS, and GNU/Linux distributions such as Ubuntu. While the underlying operating principles of both the Bitcoin and Tor protocols respectively remain the same regardless of the device or operating system they are used on, mobile devices tend to differ significantly from personal computers in how they deal with issues such as file management, application file acquisitions and methods of installation. Perhaps more importantly, there is a significant level of difficulty involved in extracting the contents of mobile devices' storage media for reasons that range from enforced full disk encryption to the fact that getting physical access to their storage requires special tools and different methods of extraction from the ones that an investigator might be familiar with.

1.3 Main Contributions

This dissertation analyses and evaluates the core systems (namely Tor and Bitcoin) and it makes the following three main contributions:

- Perform a survey on existing literature that cover the subjects of Tor and Bitcoin de-anonymisation techniques and forensic artefact acquisition and combine their collective findings into a singular document.
- Analise the currently-known vulnerabilities and shortcomings that are affecting the security and anonymity-providing features of Tor and Bitcoin.
- Deliver a comprehensive and all-inclusive guide into forensically examining computer systems where Tor and Bitcoin were used and an analysis of the meaning of the recovered artefacts.

1.4 Thesis Organisation

The remainder of this dissertation is organised as follows. Chapter 2 presents a quick introduction of all the different aspects covered by this dissertation. The operating principles of the Tor and Bitcoin protocols are presented in Chapters 3 and how they allow cryptomarkets to exist in the darknet. Known vulnerabilities that can lead to compromising users' anonymity is examined in greater detail in Chapter 4. A comprehensive guide to forensically examining all the systems discussed in this dissertation is presented in Chapter 5. Ultimately the dissertation is concluded in Chapter 6.

2 A Primer on Tor and Bitcoin

Tor is a popular anonymization network which in simple terms allows an Internet user to hide their IP online. Bitcoin is the most widely-used cryptocurrency in circulation today. Together they are used by hundreds of thousands of people on a daily basis and they currently are the two most prominent ways by which crime is facilitated on the darknets. This chapter examines the basic components of both systems, followed by a description of the way they operate. But first, we provide an overview of the main cryptographic operations which are used in the design of these systems.

2.1 Foreword on Cryptography

Every single system examined in this dissertation have cryptography and encryption at their core. Therefore it is crucial that several key concepts are discussed, even briefly before looking into each individual system. Next, the basic encryption techniques utilised by both Tor and Bitcoin will be introduced so that the reader will have a better understanding of how those two systems work.

Cryptography and encryption are used interchangeably in everyday discourse but they actually have a different meaning. The term Cryptography refers to a series of techniques that were designed for the purpose of keeping communications secret and concealed from the eyes of unintended recipients. Encryption refers to mathematical functions that can be applied to messages in order to keep them secret.

- **Public-key cryptography (asymmetric cryptography):** Computer scientist Dr. William Stallings gives a very detailed description about the definition of public-key cryptography in his highly-referenced book “Cryptography and Network Security: Principles and Practice” [42]. It is described as an encryption method where two related but not identical keys are used, a public one and a private one. Unlike symmetric cryptography where a single key is used for both encrypting and decrypting: the public key is used to encrypt, and the private to decrypt. While the two keys are mathematically related, it is almost impossible and impractical to discover the private key by analysing the public key. Because of this, public keys are meant to be shared freely. In the case of Bitcoin and similar cryptocurrencies this cryptographic method carries a big significance due to the fact that Bitcoin address creation is utilising the public key cryptography model and it will be explored in more detail in Section 2.4.3 of this thesis.
- **Digital Signatures:** They are mathematical techniques that are used in order to validate the authenticity and integrity of a message and can be considered as the digital equivalent of a handwritten signature. Their purpose is to offer an added layer of security in communications by allowing

the different parties in a communications channel to check if the source and destination of a message is as intended and if the transmitted messages have been tampered with in some way. In the case of Bitcoin it is used to verify the validity of transactions.

- **Hash functions:** Hash functions are mathematical operations that take input data (e.g. bytestrings), perform operations on it, and ultimately return output data of a fixed size. They are used used to ensure the data integrity of files or the safe storing of passwords. In the bitcoin protocol, hash functions are an integral part of the block hashing algorithm which is used to record new transactions into the blockchain.

2.2 Why Focus on Tor and Bitcoin?

Of all major anonymity and privacy networks, Tor remains the most widely used [20]. This could be attributed to many different factors, potentially chief among them being the ease in which a person can access the network and utilise its services. From the end-user perspective, accessing the services offered by the network might be as easy as downloading and installing a modified version of the Firefox web browser[1] from the Tor project's website. For individuals looking to offer private and anonymised services through the Tor network, the setting up process might be as simple as editing a few simple configuration files on their systems.

A typical Internet user has few options available to them if they want to hide their tracks when they are online. They can use Virtual Private Networks to route their network traffic through overlay networks or use intermediary, proxy servers to access information on their behalf. The problem with those solutions is that they generally tend to be centralised and fall under the control of authorities that keep logged entries of users and how those users are making use of them.

In the case of Bitcoin, the cryptocurrency first appeared on the scene with the introduction of a self-published paper by "Satoshi Nakamoto" back in October of 2008 [36] and was immediately followed by the project's launch on sourceforge as an open-source effort. The project's creator (or group of creators) simply known by the online moniker of "Satoshi Nakamoto" described it as a "purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution". Although the reasoning behind Nakamoto's decision to release the protocol specifications were never fully made clear, this paper's author's interpretation is that it has to do with the dismay the Bitcoin creators experienced with the way the international banking system operates and how susceptible it is to external influences that could threaten its very existence. In the *Genesis Block* of the Bitcoin's Blockchain (more on both in the following sections of this chapter), the following somewhat cryptic text is encoded in hexadecimal form:

"The Times 03/Jan/2009 Chancellor on brink of second bailout for banks"

This is a reference to the cover page of the British newspaper "The Times", dated 3rd of January 2009[2], that was promoting a news story in the same issue that was discussing the then Chancellor of

[1] `https://www.torproject.org/docs/faq.html.en#DistributingTor`
[2] `https://bitcointalk.org/index.php?topic=397557.0`

the British Exchequer's Alistair Darling plan to issue another financial aid package to the banks after the worldwide economic crisis of 2007 - 2009. While this by itself does not reveal any actual information about the Bitcoin creators or their modus operandi, if it is taken into account in conjunction with the Bitcoin system's unregulated, de-centralised, peer-to-peer nature it allows a person to make some reasonable assumptions as to the reason why Bitcoin exists in the first place.

Since then, Bitcoin has become a worldwide phenomenon. In their highly cited report which examines Bitcoin's meteoric rise in popularity, Barber et al. [5] go through the reasons why Bitcoin succeeded where other attempts throughout the decades failed in order to create a functional system that could serve as a large-scale e-payment system. Amongst the reasons listed in their paper, arguably the most prominent ones are:

- **No central authorities and points of trust:** As mentioned, the system relies entirely on a network of peer-to-peer nodes in order to function and on the assumed honesty of the majority of participant nodes in the network in order to validate transactions and to eliminate the problem of double spending and forgery.
- **Predictable money supply:** Bitcoins are currently generated in chunks of 12,5 per block and this aspect of the protocol is going to be examined in more detail in subsequent sections of this chapter. The amount of newly-created Bitcoins used to be 50 per block but the protocol has a hard-coded limiter to its supply algorithm [36, 5][3] halving the amount of Bitcoins per block once a certain threshold has been reached. In addition to that, there is also a hard-coded limit to the total number of coins that can ever be minted, at 21 million [36, 5]. For reference, at the time of the writing of this dissertation, there was a total of 16,820,025 Bitcoins in circulation[4]. Bitcoin's capped supply was conceived as a means to mirror the Gold Standard monetary model and to prevent spasmodic fluctuations in the market value of the cryptocurrency, something that traditional, fiat currencies suffer as a result of banking or political decisions.
- **Creation of new businesses and business models:** Being both a new type of currency and a new method of conducting financial transaction online, Bitcoin has enabled the creation of new types of businesses. Some legitimate, some not so much.

2.3 Tor in Detail

Tor offers one solution to the problem described above in Section 2.2. It provides a way by which network traffic is forwarded through a series of somewhat randomly-selected network nodes, that are referred to as relays in the technical language of the project. While for the most part the network works in a de-centralised manner, there exists a series of authority servers situated in different countries around the world that manage the consensus, a list of every Tor nodes available for use within the network. The location and Internet addresses of those authorities and how Tor interacts with them is hard coded in the Tor protocol [19, 44]. The consensus list gets updated periodically and is sent to clients wanting to

[3] https://en.bitcoin.it/wiki/Controlled_supply
[4] https://blockchain.info/total-bitcoins

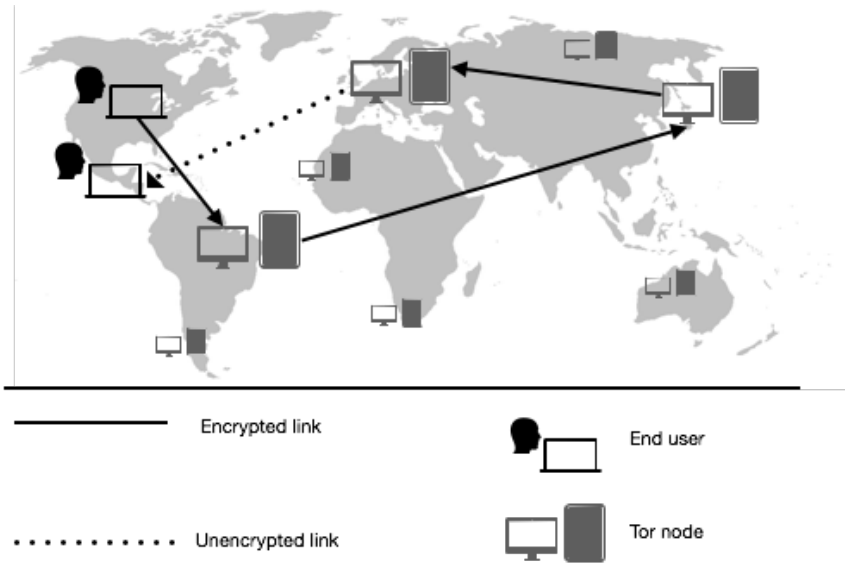

Figure 2.1: Typical Tor Circuit

make use of the network in order to have a "map" of sorts of all the available Tor nodes that they can connect to.

2.3.1 Tor Circuits

After a client acquires a copy of the consensus, it initiates the connection to the network by randomly selecting three Tor relays from that list, see Figure 2.1. The first of which is called the Entry Node, the second one is referred to as Middle and the third one is called Exit node. The three nodes create what is referred to as a Tor Circuit. The information that gets transported through the circuit is safeguarded by becoming encapsulated inside multiple layers of encryption and the origin of the sender gets hard to trace due to the fact that Tor nodes only know the previous and immediate nodes. Once information reaches the final node in the circuit (exit node), all layers of encryption are removed and information is ready to reach their intended destination.

2.3.2 Tor Cells

As mentioned, traffic inside the Tor network gets routed through a circuit composed of randomly-selected nodes while encrypted with TLS connections using ephemeral keys. The basic unit of communication for onion routers is a fixed-size "cell", which like the description implies, always has the same size of 512 bytes and it contains a header and the payload. Enforcing fixed-sized messages is also a way for the Tor protocol to mask traffic. According to the 2014 draft of the Tor design paper, which includes all changes and implementations added the Tor project since 2004 [19], fixed-size cells make the network traffic generated by Tor predictable and susceptible to fingerprinting, therefore a variable-length cell was introduced but it is currently unused. Methods in which tor cells could be examined in order to fingerprint traffic generated from the Tor network is discussed in more detail in Chapter 4.

2.3.3 Tor Bridges

Before connecting to the Tor network, users will need to acquire the (publicly available) list of Tor nodes. If somehow they are being prevented from acquiring that list, they will be incapable of making use of the Tor Network. Bridges are hidden relays that are not listed in the main Tor directory. They are one of possible alternative entryways into the Tor network, implemented by the developers as a means to circumvent attempts at censoring the medium[5].

This does not mean that bridges are a complete foolproof method of circumventing censorship, though. Prior research has shown the limitations of bridges in places such as the People's Republic of China, where the state has developed a highly sophisticated and adaptable system for filtering with high efficiency network traffic going in and out of the country [45].

2.3.4 Pluggable Transports

Pluggable Transports are services whose function is to disguise the traffic generated by the Tor protocol, making it look like something else, for example normal HTTP traffic. A list of the most popular Pluggable Transport services are listed on the Tor website along with a description of their functionality [6].

2.3.5 Tor Hidden Services

Tor allows for users to offer services inside the Tor network such as website hosting / web servers but also offering the ability to hide their location inside the system [8]. Such services are named *hidden services* and have been supported by the Onion protocol since its inception.

In order to provide this level of confidentiality, the Tor protocol utilises what is known as Tor "rendezvous points" in order to for Tor users to connect to those services. Rendezvous points in simple terms are used in order for hidden services to remain truly hidden inside the Tor network, access to them is not facilitated directly but instead happens through what is known as a "rendezvous point" node inside the network. The contact information to reach a hidden service is stored inside the DHT, which is short for Distributed Hash Table and acts as Tor's version of a DNS, in a distributed form and it allows for resolving a .onion hostname into the contact information necessary to establish a connection to the hidden service.

2.3.6 An Insight into Tor Relay Providers

According to the project's internal metrics and analytics, there are well over 6000 relays in operation by the time this report was composed[7]. Relays are spread across different countries across the world and for the most part they are operated by volunteers who are willing to dedicate hardware and bandwidth for the cause. Surprisingly enough, information regarding the name of the relay, their IP address,

[5] https://www.torproject.org/docs/bridges.html.en
[6] https://www.torproject.org/docs/pluggable-transports#list-of-pts
[7] https://metrics.torproject.org/networksize.html

geographical location, operating system and operational status are readily available for everyone to see, thanks to the Tor project's *Atlas* web API[8]. At the time of the writing of this text, "Atlas" was in the process of being renamed "Relay Search"[9], so names might be different.

Currently there are two ways in which an individual or group can volunteer to assist with the allocation of relays on the network[10]: one can either configure their hardware accordingly in order to be able to function as a Tor relay, or alternatively they can provide financial aid to relay operators in the form of monetary donations.

Of the three different types of relays (guard, middle, exit) perhaps the most interesting when it comes to the field of deanonymising communications conducted over the Tor network is the exit node. As was discussed earlier, exit relays are responsible for sending out traffic directly to the user's end destination of choice. Therefore, any potential activity that could be considered to be in violation of some civil law or any other legal system could draw the attention of law enforcement authorities to the operators of the exit node [18]. As far as a hypothetical individual who might be observing the traffic between the exit node and the end destination, the exit node appears as the originating source of all traffic.

2.4 Bitcoin in Detail

This section examines the inner workings of the Bitcoin protocol in more detail. Areas of focus include but are not limited to: cryptocurrences and Bitcoin, definition of Bitcoin, Bitcoin wallets and addresses and how mining, transactions and transaction verification work.

2.4.1 Cryptocurrencies and Bitcoin

Cryptocurrencies are digital assets and units of credit that are designed in order to work as a means of payment similar to typical fiat currencies. They utilise a combination of methods such as cryptographic functions and peer-to-peer networking technologies in order to facilitate their minting, and to keep an unchangeable and tamper-proof record of the transaction history [3].

They are not an entirely contemporary concept. The idea of using cryptographic algorithms in order to facilitate monetary transactions using electronic equipment has been in circulation since the 1980s [13], starting with U.S. computer scientist and cryptographer's David Chaum's idea of using a cryptographic method called *Blind Signatures* in order to create a method to facilitate electronic, untraceable payments. Over the years some of those ideas have been put into practice in the late 1980s up to mid to late 1990s, without much success [16].

The digital currency landscape has changed irreversibly since the introduction of Bitcoin in the late 2000s, though [36]. Bitcoin managed to solve a significant issue that digital, decentralised forms of payment have faced since their conception which is to have a trustworthy way of keeping track of

[8] https://www.torproject.org/getinvolved/relays
[9] https://blog.torproject.org/we-made-big-improvements-searching-relays
[10] https://www.torproject.org/getinvolved/relays

transactions and the prevention of double spending [39]. In recent years, partly thanks to their ability to provide their holders a sense of anonymity and concealment and as a consequence their extensive use in the dark net marketplaces, cryptocurrencies have exploded in popularity, in a relative very short amount of time.

Again, as it was the case with anonymity networks, cryptocurrencies were not developed with malicious or nefarious intents. The unregulated nature of both systems allows for every sort of activity to go unnoticed for the most part, therefore it was only natural that criminals would make use of it.

2.4.2 The Bitcoin Currency

Before explaining how the transferring of funds from person to person happens in the Bitcoin ecosystem, we need to have a look at what exactly is a Bitcoin. In the project's specification paper [36], Nakamoto describes a coin as "a chain of digital signatures" that owners "transfer(s) the coin to the next by digitally signing a hash of the previous transaction and the public key of the next owner and adding these to the end of the coin." What this means in practice is that despite the real-world, skeuomorphic analogies that are used to describe the various different aspects of the Bitcoin project, such as "coin", *transfer*, *mining*, *wallets* and so on, there is no digital or real-world artefact, considered a *Bitcoin*. Therefore, it could be said that Bitcoins are just records of transactions between different Bitcoin addresses, with balances that increase and decrease as people use their accumulated currency in order to make out payments to other Bitcoin addresses in the network.

A more apt analogy for Bitcoin is that it is closer to a bank account than a currency. And as is the case with banking transfers, when a payment is made, the payer does not physically send money to the payee but rather the payer's account is debited and the recipient's account is credited. Alongside the main Bitcoin network, a separate but identical network operates with the name Bitcoin Testnet. The underlying protocol remains the same but it maintains its own transaction history, meaning a blockchain.

2.4.3 Bitcoin Addresses

As mentioned, users inside the Bitcoin network are known only by their addresses. Addresses are unique identifiers that are created from a user's private and public key pairs [36, 22]. A good analogy would be to compare the Bitcoin addresses to email addresses. If a person wants to send an email to another person they would need to know their email address. However, unlike email, Bitcoin provides the ability for its users to have a different Bitcoin address for every transaction they conduct. In fact, it is considered poor practice to re-use the same address for multiple transactions as every transaction gets permanently recorded in Bitcoin's publicly-distributed ledger, the blockchain [6].

A peculiar *quirk* of the Bitcoin system is the concept of *change addresses*. If a person that has in their possession a certain amount of the cryptocurrency, e.g. 10 and they want to pay for something that only costs half as much, they are required to spend all of their Bitcoins and receive the change back in their change address.

2.4.4 Bitcoin Wallets

While there is some confusion regarding the difference between the two, Bitcoin Wallets and Bitcoin Addresses are two separate concepts. A Bitcoin wallet can refer to two entities:

- The collection of a user's private keys.
- The software that manages those keys and also allows a user to conduct transactions inside the Bitcoin network.

So, a Bitcoin wallet can be seen as analogous to a physical wallet. But instead of storing Bitcoins literally, what is stored in them is the information that are used in order to access Bitcoin addresses and carry out transactions. There are four main types of Bitcoin wallets: desktop, mobile, web-based and physical, either in paper or electronic hardware form.

- *Desktop wallets* are applications that are installed on a personal computer and provide the user with complete control over the wallet. Desktop wallets allow users to create Bitcoin addresses for sending and receiving Bitcoins over the network. They also allow the user to download the entire Bitcoin blockchain and participate in the mining and transaction verification processes. One of the most famous implementations of this type of wallet is Bitcoin Core, the Bitcoin project's official reference application.
- *Mobile wallets* are applications that as their name implies, are meant to be installed o mobile devices such as smartphones and tablets. They have several benefits over desktop wallets, including their self-evident ability to be taken everywhere their user goes at since they reside inside their portable devices. In addition, they tend to be more resource-intensive than their desktop counterparts as they don't usually need to download the blockchain but instead synchronise with a remote server.
- *Web wallets* are similar in some aspects to mobile wallets as they allow for the usage of Bitcoins from any location either on a personal computer or mobile. Their main difference compared to the two previous types of wallets exists in the fact that the users' private keys are not stored locally but on a remote server.
- *Physical wallets* exist mainly in two different forms: Papers that include the user's private keys usually in the form of a scannable QR code. And then there's the electronic hardware wallets that store the private keys inside a secure hardware device. This happens by storing the private keys usually in an encrypted form inside the hardware device's microcontroller.

Another way to classify Bitcoin wallets is based on the type of key derivation system they use. Two different types of key derivation techniques are used by Bitcoin wallets today: Deterministic and non-deterministic. In contrast to non-deterministic wallets, deterministic ones derive keys from a single starting point known as a *seed*, which is usually a human-readable mnemonic phrase. They allow a user to easily back up and restore their wallets without the need for more information that is hard to memorise and easy to misplace. An example of a mnemonic phrase is: "vendor shiver expire table salon horse".

Deterministic wallets have the added advantages of being able to generate a seemingly unlimited number of addresses on the fly and users can conveniently create a single backup of the seed in a human readable format that will last the life of the wallet and can be used on multiple wallets.

2.4.5 How Does One Acquire Bitcoin?

When a natural or legal person wants to acquire currency using traditional (and law-abiding) means, they have a few options available to them such as work to earn them or beg for them. In the Bitcoin ecosystem there are two main ways in which a person can acquire currency: mining and exchanging real-life, fiat currency for cryptocurrency at an exchange service.

In their 2013 paper "Quantitative Analysis of the Full Bitcoin Transaction Graph", Ron and Shamir [40] discuss how a large number of all the Bitcoins that have ever been minted remain dormant in addresses which had never participated in any outgoing transactions.

- **Bitcoin mining:** This will be examined in more detail in the following sections of this chapter.
- **Bitcoin exchanges:** An exchange is where buyers and sellers of the cryptocurrency conduct their business. A prospective seller deposits Bitcoin with the exchange's address. They can then use their positive BTC balance in the exchange to sell his BTC for traditional fiat currencies (such as Euros or Dollars). Similarly, a buyer of BTC deposits USD with the exchange and then uses the balance to buy BTC from sellers. According to the bitcoin wiki, the most popular bitcoin exchanges at the time of the writing of this dissertation were[11] and are summed up in Table 2.1.

2.4.6 Risks Associated with Bitcoin

Using Bitcoin carries certain risks for its users. In their paper "Bitcoin Risk Analysis", Kiran and Stannett [27] talk about the risks that Bitcoin presents or will potentially present in the future within economic models. Alongside the risks it also presents risk-management analysis that could be applied in order to mitigate the negative effects presented by said risks. Their definition of a risk is "a risky situation is one which presents potential exposure to danger, and the level of risk can be thought of as a measure of the assets that would be affected as a result of a particular threat being realised through the system under analysis." They employed the use of a numeric score ranging from 1 to 7 in order to classify the different risks associated with Bitcoin usage, with 7 indicating a severe risk.

Among the risks examined, we have social risks such as: bubble formation, the "cool factor", etc. Social and technological risks including deanonymisation. It then discusses the vulnerabilities that could affect the economic landscape, such as Double Spending.

[11] `https://en.bitcoin.it/wiki/Comparison_of_exchanges`

Service	Website	Assets (M)	USD	EURD	Holds BTC	Holds fiat
Kraken	kraken.com	$247,000,000	Yes	Yes	Yes	Yes
GDAX	gdax.com	$235,000,000	Yes	Yes	Yes	Yes
Bitfinex	bitfinex.com	$169,000,000	Yes	No	Yes	Yes
BitMEX	bitmex.com	$72,000,000	Yes	No	Yes	No
BTC-e	btc-e.com	$60,000,000	Yes	Yes	Yes	Yes
Bitstamp	bitstamp.net	$34,000,000	Yes	Yes	Yes	Yes
Gatecoin	gatecoin.com	$19,000,000	Yes	Yes	Yes	Yes
HitBTC	hitbtc.com	$17,000,000	Yes	Yes	Yes	Yes
BitBay	bitbay.net	$6,000,000	Yes	Yes	Yes	Yes
Bitso	bitso.com	$4,000,000	No	No	Yes	Yes
itBit	itbit.com	$3,000,000	Yes	Yes	Yes	Yes
Coinfloor	coinfloor.co.uk	$2,000,000	Yes	Yes	Yes	Yes
TheRockTrading	therocktrading.com	$2,000,000	Yes	Yes	Yes	Yes
C-CEX	c-cex.com	$1,000,000	Yes	No	Yes	Yes
Luno	luno.com	$1,000,000	No	No	Yes	Yes
Kraken	kraken.com	$247,000,000	Yes	Yes	Yes	Yes
GDAX	gdax.com	$235,000,000	Yes	Yes	Yes	Yes
Bitfinex	bitfinex.com	$169,000,000	Yes	No	Yes	Yes
BitMEX	bitmex.com	$72,000,000	Yes	No	Yes	No
BTC-e	btc-e.com	$60,000,000	Yes	Yes	Yes	Yes

Table 2.1: Most popular Bitcoin exchanges ordered by assets held in US$

2.5 Summary

This chapter was dedicated to examining the Tor and Bitcoin protocols in more detail. Short for The Onion Router, Tor traces its roots back to work conducted by researchers from the U.S. Naval Forces for the purpose of keeping U.S. intelligence information secure online. After the code was released under a free licence, the Tor project was created and to this day it helps maintain the largest implementation of the protocol, the Tor network. Bitcoin is a non-centralised payment system that implements a variety of cryptographic methods in order to provide a model for payments either online or in the real world, modeled after traditional methods used but with an added degree of confidentiality. Next chapter discusses how these systems have been used in the context of cryptomarkets.

3 Darknet Cryptomarkets

Cryptomarkets are a byproduct of the tools and technologies made available freely to the general public by Tor and Bitcoin. By using Tor's hidden services, an individual can setup a server inside the Tor network that has its IP address and as a consequence its geographical location entirely hidden from individuals who try to communicate with it. Aided by Bitcoin's ability to provide its users the tools to conduct pseudo-anonymous monetary transactions online, this naturally makes those kinds of servers ideal breeding grounds for illegal activity. This chapter focuses specifically on the problem of cybercrime in the Internet, in particular on darknet cryptomarkets which leverage primarily Tor and Bitcoin as key enabler technologies.

3.1 Cybercrime in the Internet

This section provides an overview of how cyber-criminal activities take place in the Internet. We start by clarifying the concept of cybercrime, and then introduce the darknet as one of the deepest layers of the "deep Web".

3.1.1 Definition of Cybercrime

Cybercrime, or computer-related crime is a form of criminal activity that involves the use of electronic equipment, such as personal computers (e.g. Windows and Linux-based PCs, Apple Macintosh, in both desktop or laptop configurations) and various other personal electronic devices such as mobile phones (smart phones), tablets and in the past Personal Digital Assistants, and potentially a use of a network, either local or wide-area such as the Internet with the ulterior motive of inflicting to a person or group of people physical or metal harm [34].

While in years past, especially since the advent of the personal computer / microprocessor revolution in the mid 1970s, when personal computers started to became a household commodity, cybercrime was either not even a fully conceived idea or it was mostly encountered in cases where isolated computer systems were the victim of theft or the source of data breach. In recent years the definition of cybercrime has evolved in order to cover areas such as online abuse and harassment via email, message boards and chat rooms, financial fraud and unauthorised access to computer systems and networks via means such as phishing and social engineering, etc.

Since the proliferation of Internet access and advancements in the field of computational technology and encryption, a different form of cybercrime has emerged. From the bowels of a certain area of the

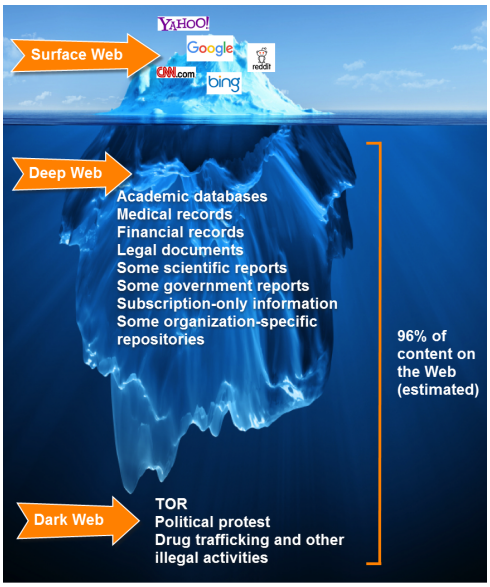

Figure 3.1: The iceberg parallelism.

Internet simply described as the darknet has emerged a new kind of technology for enabling stealth criminal activities.

3.1.2 Surface Web, Deep Web, Darknet

The Internet, and more specifically the World Wide Web is comprised of three distinct, layers that are often visualised in the form of an iceberg[1] as seen in Figure 3.1. The iceberg is separated in three parts each one representing The Surface Web, the Deep Web and the darknet [8]. The analogy is as such: The part of an iceberg that is seen floating on the sea comprises only a small fraction of its true volume. The same can be applied to different layers that make up the Internet and the World Wide Web. The first layer, the Surface Web, is the one that is immediately available for observation and interaction. It is comprised of all the information that are available on the Internet and in the Web that can be indexed by a search engine (e.g. Alphabet's Google) and can be accessed by anyone using the Internet.

The second layer, is the one that is not immediately obvious to a normal Internet user and it is usually referred to as the Deep Web. While the name implies in a sense mystery and potentially hidden and concealed activities, it is actually the part of the Internet and the World Wide Web that exists behind log-in screens, news site paywalls (e.g. The Wall Street Journal), video-on-demand (e.g. Netflix), corporate Intranets, pages behind CAPTCHAs, and so on. Websites and other content on the Deep Web could potentially be accessed by knowledge of specific, hidden URLs via a normal web browsing application.

The third layer of the Internet is the darknet. Depending on the circumstance and situation, the terms Deep Web and Dark Web/Net are used interchangeably, but for the purpose of this report and to avoid a pointless semantic dispute it would suffice to describe it as its own layer in the triad, that also forms a symbiotic relationship with the second one. It is a part of the Internet that is built on top of either

[1] https://anchisesbr.blogspot.pt/2017/10/seguranca-deep-dark-web.html

overlay networks or *darknet networks* and generally require specialised software in order for someone to gain access to them.

While their mere existence and method of acquiring access to them does not necessarily make them illegal in any way, the level of privacy they offer and the means by which one can successfully conceal their activities on them makes them a natural hotbed for the development of criminal activities. With that said, most of the interest in the darknet lies in the activities that are happening or have the potential to happen inside them than their technical aspects.

3.1.3 Darknet Architecture Analysis

As the name of this dissertation implies, a large part of it is dedicated into examining darknets, how they operate and the kinds of illegal activities that either take place on them or can be enabled to take place. While the technical implementation is different for each one of them, for the most part all of the darknet platforms that are in use today, such as I2p (Invisible Internet Project) [4], Freenet and Tor [19], all utilise a decentralised, peer-to-peer architecture in order to function as information exchange channels. What this means, in very generalised and simplistic terms, is that they forgo the traditional model of a centralised client-server architecture that serves as the focal point of a network, providing information to connected client machines and being responsible for the coordination of traffic. Instead, in the peer-to-peer model *peers* (meaning machines participating in the network) have equal privileges in how they tackle the assigned workload.

What is really interesting and noteworthy is that although these systems utilise the Internet in order to operate, they can be considered as different layers, or more specifically, Overlay Networks. That means that they are networking infrastructures that are built on top of other networking systems [**?**]. This gives overlay darknets several advantages, since they don't have to *obey* the rules of existing networks, especially in the way data is encoded when passed through their channels. For example, the Tor privacy network has support for its own top-level domain [19], .onion, which is not formally supported by the Internet's DNS (Domain Name Service).

In their paper *Determining What Characteristics Constitute a Darknet*, Aked et al. [2] provide a very simple and concise description of what constitutes a darknet and what their core differences are compared to the clearnet, all summed up neatly on the following table:

Characteristics	Clearnet	Darknet
Encryption	Sometimes	Always
Anonymity	Users Traceable	Users Anonymous
Routing	Based on Capacity, cost and geography	Few
Applications	Many	Specific
Visibility	Obvious	Hidden

Table 3.1: Differences between clearnets and darknets.

What this means is that on a clearnet, encryption is not mandatory to be implemented by the clearnet service provider, while on the darknets it is an always-enabled feature. Anonymity on the clearweb is not guaranteed while darknets are almost always build in order to hide their contents and the mask the users that access them. The path that network traffic takes on the clearnet is coordinated

by ISP routers until they reach their ultimate destination, while on the darknet, and more specifically on the Tor network, the path that network traffic will take until it reaches its ultimate destination is randomly generated when a user starts a new Tor instance. Content on the clearnet can be accessed by a variety of different software, ranging from web browsers, email, IRC and FTP clients and so on, while access to darknet content is usually done via a very small number of specialised software, in Tor's case the project's official Browser Bundle and system service. Finally, access to clearnets is available to almost anyone with access to the aforementioned applications and the existence of clearnets is well advertised by indexing services. Visibility of darknets on the other hand is limited.

Despite their extensive use in the facilitation of illegal operations on the darknet, they were not originally created with those purposes in mind. Freenet traces its roots back to university work conducted by Ian Clarke as a means of graduation [17]. The Invisible Internet Project on the other hand has its roots in activists' attempts to create a censorship-resistant platform [**?**]. The case of the "Onion Routing", the protocol that Tor was based upon, is perhaps the most interesting of the bunch, due to the fact that it was originally conceived by the United States Naval Research Laboratory as a means of protecting U.S. Intelligence communications online [38]. After the code for protocol was released as free software[2], computer scientists Roger Dingledine, Nick Mathewson and Paul Syverson began work all the way back in 2002 on what is known today as the Tor network. The aforementioned crew of scientists were some of the key people who later on turned the Tor project into a non-profit organisation with financial assistance from several, freedom-of-speech proponent organisations, such as the Electronic Freedom Foundation[3].

As mentioned earlier, one of the main differentiating factors between the Surface Web and the darknet is the fact the access to the the latter requires the use of either special system software configuration, or entirely different software altogether than the ones the general Internet-browsing public uses. For example, in the case of all three major privacy / anonymity networks, when a user who needs to use their services in order to conceal their activities online, they would need to configure their applications in such a way for them to bypass the normal internet communications channels and to use the anonymity networks and the encryption that they provide. Alternatively, if a user's activities are isolated to web browsing they could download Tor's official web browsing bundle that facilitates the connection to the Tor network without any end-user meddling[4].

3.2 Online Trading and Darknet Marketplaces

This section deals with examining the history and background of online selling of goods and services and the recent approach of doing illegal transactions using the darknet marketplaces. A large part of this section will be dedicated to examining the case of *Silk Road*, a darknet marketplace that has cast a huge shadow on the cryptomarket industry and is still discussed to this day, years after its demise.

The first known online, "e-commerce"-like transaction was a 1971 marijuana exchange between students. As described on J. Martin's book [31], students of Stanford University and the Massachusetts

[2] `http://www.torproject.org/docs/faq#DistributingTor`
[3] `https://www.torproject.org/about/sponsors.html.en`
[4] `https://freenetproject.org/pages/help.html`

Institute of Technology were using ARPANET, which was the predecessor to the modern Internet to make arrangements in order to exchange cannabis for money.

In recent times new platforms have emerged that allow buyers and sellers to facilitate the buying and selling of drugs and other illicit goods online. They are called darknet markets or most commonly Cryptomarkets. J. Martin [32] describes a Cryptomarket as a type of website that employs advanced encryption in order to protect the anonymity of users. In addition, Cryptomarkets tend to shy away from accepting fiat currencies as a form of payment and instead rely on Cryptocurrencies such as Bitcoin. Barratt and Aldridge [7] also add to the definition of the term Cryptomarkets that it is a marketplace that can host multiple sellers and that it "provides participants with anonymity via its location on the hidden web" and "aggregates and displays customer feedback ratings and comments". It is mentioned how participation in such markets usually requires a certain level of technical competence, meaning that an individual wanting to buy or sell will need to be able to know how to use specialised software required to access the darknet URLs (i.e. Tor Browser Bundle), and the ability to own and use cryptocurrencies and digital wallets.

Most of these platforms act as the *middleman* of sorts between the two parties, meaning that they do not sell the illicit goods themselves, instead provide a platform in which sellers can reach a wider audience. They might make use of an escrow system, meaning that when a buyer buys something from a vendor, the funds are not transferred immediately to the vendors but the marketplace holds on to them until the transaction is complete. This has been exploited many times in the past with "exit scams".

While they differ in the good and services that they trade, darknet marketplaces have several characteristics in common: Firstly their design structure is reminiscent of legitimate big-name websites such as Amazon and eBay, giving users the ability to search via a search function and the ability to browser according to category. Second, they utilise the encryption methods that the darknet platform they operate provides them with in order to keep communications secret, and third, they forgo the use of fiat currencies and other traditional forms of monetary payment in order to avoid leaving a trace of financial transactions conducted on their systems [21]. Instead they employ the use of cryptocurrencies such as Bitcoin for preserving anonymity. A combination of all three previously-mentioned make darknet marketplaces a lucrative place for individuals who wish to do business of dubious legal nature away from the curious eyes, especially those of law enforcement.

Silk Road and other cryptomarkets that followed used mainly two payment systems: "Finalise Early" and the "Escrow system" and for the most part it was entirely left up to the buyer to choose which payment system they preferred to use. As mentioned, cryptomarkets act as intermediates between buyers and sellers. When a buyer makes a purchase they do not transfer the funds directly to the seller but to the marketplace. After the transaction has been completed and the buyer indicates to the seller that they have indeed received their order, then the marketplace releases the funds to the seller. If the vendor has a good reputation on the market and is well trusted by the community they can ask for an early release of the funds. As a general rule of thumb, it can be considered appropriate to Finalise the transaction early when the seller is entirely confident that the vendor will ship the product. Finalise early is generally avoided in the case there is not enough trust generated by the vendor.

Another benefit of vendors having a good reputation on the darknet is they can charge a premium price for their products. Purchasing illegal drugs from a faceless and anonymous dealer online will

always carry the inherit risk of it being a scam, and in that case a vendor that has a stellar profile and they are reviewed well by members of the community will have their merchandise considered more valuable since it has a much higher probability of actually arriving at the buyer's door.

A major drawback when it comes to cryptomarkets' lack of legitimate operating credentials is that they can also be the targets of scams. Anyone with the means to replicate the design of an existing cryptomarket can do so and have it operational inside the Tor network as a Hidden Service. That would allow them to scam unsuspecting buyers and sellers. It falls onto the users and the community of legitimate cryptomarkets to police the landscape. Typically, in places like Reddit[5] users will report those scam websites to other cryptomarket users. Because users of those services cannot seek legal advice from the police or any law enforcement organisation, it is vital to them and their interests that they participate in the policing and observe news and developments in the cryptomarket world. With no centralised authority in place, users need to rely on each other in order to remain safe and secure while making use of those services. That is the main reason why, as it was clearly demonstrated by Silk Road, that buyers and sellers try to maintain their reputation high. A seller with a bad reputation (usually a low feedback score) will have trouble selling to prospective buyers. The same principle could also be applied to buyers. A buyer with a low score might potentially mean that they have tried to scam sellers by falsely reporting that the merchandise they have ordered has not yet arrived and so on.

3.2.1 The Case of Silk Road

According to the Global Drug Survey of 2016 [**?**], the launch of the Silk Road marketplace can be considered as "Year 1" in the world of darknet cryptomarkets. Recent statistics reveal that while the online drug sales account for only a small share of total drug sales, they are quickly growing, from an estimated US$15m in 2012 to US$180m, in 2015. According to Barratt and Aldridge [7], examining cryptomarkets can lead to a better understanding of drug traffic in real life. It allows us to better understand the supply side of it in its totality, since cryptomarkets are not isolated from the broader drug markets. It can enable us to monitor emerging drug trends.

Celestini, et al. [12] in their research on the Tor-based marketplaces mention how a common trait on most cryptomarket storefronts is that they are using *off-the-shelf* web technologies, in order to imitate familiar websites such as Amazon and eBay, providing sellers with personal profile pages and giving the users the ability to provide feedback on the products and services that they have acquired online. It is argued that this leads to poor quality of data that can be acquired, due to the complexity of extracting data. Therefore they propose methods in which data can be *crawled* from such places.

While Silk Road was not the first attempt to set up a drug exchange either the open or hidden web, it is often credited as being the first successful cryptomarket. To the day of this writing, long after its demise, it remains one of the largest operations setup on the darknet marketplace scene for illicit goods, and has left a long shadow and a long-lasting legacy. According to Barratt and Aldridge [7], it made full use of the anonymisation technologies provided by Tor and Bitcoin and it allowed its users to hide their activities from prying eyes.

[5] https://www.reddit.com/r/DarkNetMarkets/comments/3l515f/dnm_here_is_a_list_of_confirmed_scam_sites_if_you/

Silk Road pioneered the use of a *ratings system* in the darknet marketplace scene. Through that, the quality of products was evaluated by buyers, reducing the likelihood of contamination and dilution among others, overdosing etc. [15]. A seller's reputation on Silk Road was open to all users, which meant that sellers who wanted to keep their business up, they had to provide good quality services.

Trust between all parties involved in the transactions was a major factor on the successful operation of Silk Road. The community that was created around the website's forums displayed characteristics that are usually found in social activists. According to Lacson & Jones [28], the user base of Silk Road used to exhibit three predominant traits of coordination, social cohesion and Extremist Tendency. Trust and camaraderie were paramount to the success of Silk Road not only as a drugs exchange platform, but also as a place were people would go to in order to seek advice of other drug users.

According to the FBI's criminal complaint filed in the trial of Silk Road's alleged founder and chief operator, Ross Ulbright (USA v. Ross Ulbricht, 2013), at its peak that particular cryptomarket had an estimated 4,000 vendors serving 150,000 buyers. While its user base was mostly located in the United States, it included buyers and sellers from different other countries.

According to reports, based on the exchange rates of 2013, the FBI seized approximately USD3.6 million worth of Bitcoins from the Silk Road computers. It was speculated that the site facilitated sales worth up to 9.5 million Bitcoin and the comission it collected on those sales is over 600,000 Bitcoin [25]. At the time that was over US$1.2 billion in sales and many millions in commission fee profit.

As was the case with the closure of Silk Road by law enforcement, in every single case observed so far, resulted in the user base of the about-to-close cryptomarket to jump on to the next big thing in the darknet. Therefore, when it comes to the effectiveness of shutting down one of those online storefronts a question that one might ask is "does it even work?" If we take into consideration two aspects:

- Cryptomarkets do not sell themselves drugs but for the most part act only as the middleman between buyers and vendors.
- There is well documented precedent where users (buyers and sellers) of one cryptomarket jumped to the next competing cryptomarket that offered features similar to their previous one.

It would appear that any attempt at law enforcement to shut down one or multiple cryptomarkets contributes to nothing in the grand scheme of things. In fact it might even have the opposite effect since traditionally it made people a a lot more innovative in their attempts to hide their tracks further. So the real question regarding attempts to shut down cryptomarkets is that if the ecosystem can be disrupted, what steps need to be taken in order to do so. It is an interesting question indeed and probably one with multiple answers and approaches to the subject. But, it is not this dissertation's design goal to answer it but only to make the question available to the reader.

According to darknet news reporting website deepdotweb.com[6], at the time of the writing of this thesis there were approximately 32 operational markets and vendor shops in the dark web. When arranged by the date of their creation, we can see that Dreammarket is the oldest remaining marketplace

[6] https://www.deepdotweb.com/about-deepdotweb/

in the dark web, having been around since 2013 and still going strong[7]. When examined by other notable characteristics, such as user reviews and satisfaction of their services, Dream Market, Libertas Market, Wall Street Market, and CGMC are amongst the highest ranking in the darknet.

3.3 Cryptocurrencies and Anonymisation Networks

Cryptocurrencies, especially Bitcoin, are widely used in darknet marketplaces to facilitate anonymous monetary transactions online [11]. Both Bitcoin and Tor have developed a somewhat symbiotic relationship with one another, in multiple ways. Assuming that neither the buyer and / or the seller avoid significant missteps that would compromise their anonymity while using a cryptocurrency, such as advertising in some capacity their Bitcoin wallet address and associating it with their real life name or their online moniker that could be traced back to their actual identities, or a transaction ID, it is extremely difficult if not outright impossible at times to link transactions to individuals.

Since the peer-to-peer architecture of cryptocurrencies make them a *community-operated* system, there is no central authority that keeps checks and records of real world names and activities that take place inside the system, making it impossible for law enforcement agencies to issue a subpoena to a governing body, when the need arises. That does not mean that there is not a ledger in which transactions inside the system are recorded, but what gets transcribed in there does not directly reference people, locations, Internet addresses or the type of goods or services purchased with the currency. Yet again that does not mean that the system provides full anonymity. As mentioned earlier, if for example a user of the currency advertises their wallet address in conjunction with elements that could lead to their actual identities, that immediately creates a significant opportunity for individuals looking to de-anonymise cryptocurrency users. Therefore, systems such as Bitcoin are considered “pseudo-anonymous” [23] and not fully anonymous.

There is another method in which cryptocurrencies and anonymity networks can work together to create what, in theory, can be described as a more secure method of performing financial transactions online. As mentioned, Bitcoins and similar electronic currencies exist solely as digital assets, therefore they rely on computer networks in order to circulate and be useable as a type of currency. Therefore, it is entirely possible to alter the way in which cryptocurrency wallets are connected to the network, in this case the Internet. There are methods in which a user can configure their wallet software to utilise one of the aforementioned privacy networks in order to further blur their online traces.

While utilising the encrypted and abstractly-conceived communications channels of anonymity networks does indeed sound like a good idea on paper, recent studies into the field of cryptocurrency usage over a privacy network (specifically Tor) has shown that their concurrent use is anything but safe and that users should be very careful about doing so [10, 9].

[7] `https://www.deepdotweb.com/dark-net-market-comparison-chart/`

3.3.1 How to Set Up a Darknet Marketplace

As mentioned, the majority of darknet cryptomarkets today run on the Tor network. More specifically they make use of Tor's Hidden Services feature in order to be accessible only via the Tor Browser or any other browser that is configured to access the Tor network and keep their geographical location secret via the use of rendezvous points on the Tor network.

The process of setting up a darknet cryptomarket has been refined and it now exists in an almost standardised form since the days of Silk Road. As shown in the instructional manual released by TheOnionShop creators titled "Onionshop Installation Guide"[8], it is a fairly simple and straightforward action to create a Hidden Service inside the Tor network. The *steps* one must take in order to do are plenty and it includes:

- **Selecting an operating system:** The authors of the paper recommend an Open Source operating system instead of a closed one, such as Windows.
- **Self host or use a hosting service:** Having access to the server infrastructure gives the user a greater degree of autonomy and potentially privacy.
- **Fake identity setup:** In case a hosting service needs to be utilised in order to host the Hidden Service, it is highly recommended for obvious reasons that one does not provide their real identity. In addition it is recommended that they create an email address that does not draw suspicion.

3.3.2 Harm Reduction and Cryptomarket Culture

While cryptomarkets have the potential to introduce people to drugs that might otherwise be outside of their ability to purchase, it is theorised that purchasing from those online places can have some unintentional positive side-effects:

- The rating system is a way for the community to keep track of vendors who sell "quality" substances.
- Avoidance of "back alley" exchanges that could lead to physical violence.
- Users know that they can't order online too often, so they try to ration their drugs to last them longer. Avoidance of overdosing and similar situations.

3.3.3 User Participation in Cryptomarkets

One of the most well documented case study about user participation in cryptomarkets came from the paper authored by Van Hout and Bingham [43] which was written prior to the closure of Silk Road by law enforcement. It is a single case study regarding a selected individual picked from the Silk Road

[8] https://www.deepdotweb.com/2015/03/27/onionshop-guide-how-to-set-up-a-hidden-service/ http://thehub7gqe43miyc.onion/index.php?topic=7507.0;topicseen

forums that participated in a research which documented user experiences while browsing and shopping for recreational drugs inside the various DNMs that were available in the darknet at the time.

The participant, a male of 25 years old, was in professional employment. He described himself as a “psychonaut” and called his drug usage a “life-enchancing tool” to expand his consciousness within a personal and lifestyle oriented journey. The main reason he started purchasing drugs from darknet markets was to avoid legal issues that usually come with normal drug purchasing. While reviewing his options in the darknet market he took other user experiences into consideration.

As an experienced computer user found no difficulty acquiring and using the software that is required to access the darknet (read: Tor) and he was able to follow instructions online in order to further enhance his anonymity and security as he was a privacy-conscious individual. He also spoke about personal responsibility when dealing with such matters and specifically mentioned that if, for example you leave the Tor software on your computer during a police raid, you are responsible. Acquiring Bitcoins was perhaps the hardest part of his darknet market escapades, as he had to setup a fake bank account in order to exchange his fiat currency for Bitcoin and then transfer those Bitcoins to Silk Road.

3.4 Summary

This chapter provided an in depth analysis of cryptomarkets, which are a byproduct of the tools and technologies made available freely to the general public by Tor and Bitcoin. By using Tor’s *Hidden Services* feature, an individual can setup a server inside the Tor network that has its IP address and as a consequence its geographical location entirely hidden from individuals who try to communicate with it. Aided by Bitcoin’s ability to provide its users the tools to conduct pseudo-anonymous monetary transactions online, this naturally makes those kinds of servers ideal breeding grounds for illegal activity.

Silk Road, which is considered to be the first successful darknet marketplace, by utilising the technologies and anonymity features available through Tor and Bitcoin has created a paradigm shift in the methods by which illegal activities can escape the eye of law enforcement online and even though it has long been shut down, it has left behind a rich legacy by having spawned a variety of imitators and successors. In the following chapter, a more in-depth analysis of Tor and Bitcoin vulnerabilities will be provided. Such vulnerabilities can be used as potential sources of forensic evidence.

4 Tor and Bitcoin Vulnerabilities

This chapter is devoted into examining known Tor and Bitcoin vulnerabilities that could be exploited for the purposes of de-anonymising the users of both systems. Likewise, such vulnerabilities can be used for forensic investigations.

4.1 Tor Vulnerabilities

There is a variety of reasons that define whether the Tor network and the randomly-generated circuits that users create are secure or not. Due to its nature and different methods of being used, Tor vulnerabilities can come in many different forms. This section reviews the literature that revolves around the field of known Tor security vulnerabilities and exploitations. Being a network that is designed to operate on top of infrastructure that is donated by people the world over, Tor is bound to be the target of intense scrutiny in order for potential vulnerabilities to be discovered.

As mentioned previously, Tor is built around a network of volunteer nodes, situated across different countries and different continents around the world. The primary operating principle of the system is to hide the tracks of its users online by re-routing their traffic across a series of nodes before that traffic reaches its ultimate destination. There are a few known methods by which Tor-related traffic can be analysed in order to de-anonymise the network's users, but can for the most part be separated into three main categories:

- **Protocol-level attacks –** A protocol-level attack can occur when an attacker takes advantage of a shortcoming in the design of a specific protocol, or by taking advantage of an outdated or insecure implementation of it in order to be able to carry out an attack.
- **Traffic correlation attacks –** As it currently stands, Tor cannot protect against the monitoring of traffic at the fringes of the network (i.e., traffic entering through guard node and traffic that is leaving the exit nodes to reach its ultimate destination). While Tor employs the use of several techniques in order to mitigate traffic analysis (i.e. splitting the traffic in even-sized cells), it cannot prevent traffic confirmation.
- **Fingerprinting attacks –** Also referred to as "Traffic Fingerprinting" is a technique that can be used to identify web traffic and user behaviour while browsing the web.

Next, we present some relevant attacks from each of these categories.

4.1.1 Protocol-level Attacks against Tor

The Tor network is susceptible to a number of protocol-level attacks that can be broken down into two different categories:

- **Cell manipulation –** Consists of the purposeful manipulation of "cells", which are the transmission units of Tor, in order to correlate traffic as it enters and exits the Tor network.
- **Routing attacks –** Involves the manipulation of the Tor network's circuit building mechanisms in order to build a circuit that will include compromised nodes.

In their paper *Protocol-level attacks against Tor*, Ling et al. [29] propose methods by which a single manipulated Cell could be carried along a Tor circuit and how its recognition errors could be tracked all the way from the point of manipulation to the exit router. Via experimentation, the authors have demonstrated that such an attack can significantly limit the anonymity factor of Tor, especially if the attacker can control even a small number of Tor routers. According to their research, as the manipulated cell will reach the exit node, it will produce an error message that is unique to those types of attacks and as a consequence of that, an attacker with the ability to observe the network, will be able to confirm the relationship between the sender of the manipulated cell and the receiver. The authors conclude their research by providing some potential solutions for this Tor shortcoming, in particular by minimising potential compromised entry nodes. The methods proposed to achieve this include making improvements to the Entry router selection algorithm in order to prefer fully-trusted nodes.

When it comes to routing attacks, in their heavily cited paper *Low-Resource Routing Attacks Against Tor* [**?**], Bauer et al. examine a method by which Tor's circuit-building mechanism can be tampered with in order to build a circuit that contains compromised nodes, therefore compromising a user's anonymity before even actual data have been transmitted over the network. By design, Tor tends to favour nodes with higher bandwidth capacities when constructing its three-hop circuits. A network node's bandwidth is advertised on the network via the consensus and at the same time, the Tor protocol does not contain mechanisms in order to check if the advertised bandwidth capabilities are in fact valid. As such this leaves room for an attacker to compromise a large number of circuit-building requests.

4.1.2 Traffic Correlation Attacks against Tor

With respect to traffic correlation attacks against Tor, we briefly highlight three main techniques:

- **Tor HTTP usage and information leakage –** In a paper published by research students at the University of Technology in Austria [24], it is argued that, while Tor is the most widely used open anonymisation network available today, expected to be used by hundreds of thousands on a daily basis, few people understand on how to use it properly, meaning that sophisticated privacy-enhancing tools like "Tor Button" and "Tor Flow" are generally not widely used. As such, the TOR network is being responsible for a large number of plain HTTP tunnelling. The main argument presented is that there is very little research conducted on the kind of HTTP traffic that is being

tunnelled via Tor. The authors did their research by deploying their own controlled exit nodes, capturing a significant number of HTTP requests in the process. Among their most significant findings was the discovery that that the majority of users were susceptible to sophisticated deanonymization attacks by a compromised exit server. Furthermore, they discovered that an estimated 7 percent of the captured traffic was related to users' social media activities. Among other information, the paper also discusses the amount of "plaintext" information observable even on a passive manner.

- **Architectural exploits –** In a paper that was the result of a collaborative effort between researchers at the US Naval Research Laboratory and students at Georgetown University in Washington DC [26] the methods by which the architectural shortcomings of the Tor protocol could be exploited in order to correlate traffic entering and exiting the network. It presents some realistic (or realistic-like) scenarios in which adversaries might attempt to intercept, alter, drop, add or tamper in other ways with the communications channel.

- **Exit node traffic sniffing –** In Tor, despite the fact that as traffic travels across the network is being encapsulated inside multiple layers of encryption, there is no guarantee that the nodes it is passing through are not being monitored or they are compromised in some capacity[1]. Therefore there is an inherent risk in using Tor for transmitting information. This risk is increased significantly as traffic leaves the exit node and reaches its ultimate destination. Currently and for the foreseeable future, Tor does not offer the ability to mask the traffic generated between the exit node and the final destination.

4.1.3 Fingerprinting Attacks against Tor

Fingerprinting, also referred to as "machine fingerprint" or "browser fingerprint" is the process of collecting information from a remote system for identification purposes. While the collection of information from web browsers by analytic services has been happening since the Web's inception, the advancements in the technologies used to create modern web sites and content-rich interactive web applications have allowed for the collection of far more information than it was possible before. More specifically, HTML5, the fifth revision of the HTML standard, and more specifically its "Canvas Element" which is meant to be used to create interactive 2D elements on a web page, can be exploited to extract information from a user's machine or online actions in order to create a unique fingerprint.

In the case of Tor browser, Internet tracebacks via specific browser states and configurations are still an issue. While Tor provides a solid level of privacy on the architectural level, that is not by any means enough to safeguard a user's privacy. When browsing the web (clear and deep), either by using a traditional web browser or the Tor browser, information related to the user and their activities online, their machine's configuration or their physical geographical location can be leaked to attackers. This can happen via the use of browser security vulnerabilities, insecure browser add-ons and plug-ins, or in some extreme cases by the way a user likes to position their browser window on their screen.

We also add that the Tor Browser Bundle is based on a modified version of the Mozilla Firefox web browser, with included additional security and privacy oriented modifications and add-ons. Thus, the

[1] https://motherboard.vice.com/en_us/article/4x3qnj/how-the-nsa-or-anyone-else-can-crack-tors-anonymity

Firefox and the Tor Browser share a common codebase and as a result strengths and weaknesses and vulnerabilities. Their main differences lie in the fact that Tor is modified in such a way that plugins are disabled by default and that the logging of visited site history and cache are not persistent.

4.2 Bitcoin Vulnerabilities

There has been a plethora of papers that have been published regarding Bitcoin and its inherit security features, or lack thereof. The Bitcoin protocol while providing a high degree of confidentiality in transactions also keeps a record of every single one that has taken place since its activation in 2009. As stated in previous sections, those transactions are recorded in the network's Blockchain, a distributed ledger that is available to all participating nodes in the network. While the information transcribed there exist in an anonymous form (i.e. not tied with information that can trace them back to their users such as real names and IP addresses), by making use of information from outside the Bitcoin network, it has been proven that certain transactions can be linked back to their instigators. Furthermore, while the information inside the Blockchain does not exist in a *plain-text* format, there are utilities and even websites that can allow one to parse the contents of the Blockchain and search for records that contain specific Bitcoin addresses and even get a list of all transactions associated with them.

In addition, besides using external information to correlate users and their transactions, there are methods by which one can for example manipulate the protocol's built-in mechanisms against Denial-of-Service attacks (DoS) in order to route traffic via specific channels on the network. Furthermore, making Bitcoin wallets route their traffic over the Tor network, something that seemed like a good idea not that long ago, was recently discovered to have the potential to be catastrophic when it comes to maintaining Bitcoin users' anonymity, as explained below.

In this section Bitcoin vulnerabilities are classified in two different categories based on the factor that triggers them and allows them to be a liability:

- **User-induced –** Opportunities for exploiting aspects of the Bitcoin network based on actions performed by its users. Someone who can see all of your Internet traffic can easily see when you send a transaction that you didn't receive (which suggests you originated it). Bitcoin-QT has good Tor integration which closes this attack vector if used.
- **Architectural shortcomings –** Flaws or shortcomings presented in the Bitcoin protocol itself that would allow for deanonymisation.

4.2.1 User-induced Bitcoin Privacy Shortcomings

We elaborate on two main limitations of Bitcoin which can compromise the users' privacy:

- **Opportunities for deanonymisation when mixing Bitcoins –** Unless someone can *mine* Bitcoins, a task that is both time consuming and expensive, they would typically need to exchange their fiat currencies for Bitcoin. But, in order to do so, they would have to trust an exchange service

and most of them ask from their users to provide some sort of identification before they allow an exchange process to take place, and this could potentially leave Bitcoin users with their anonymity exposed.

A method by which users can achieve greater anonymity in the Bitcoin network is through *mixing*. A research paper published by Malte Möser [35] deals with examining how the mixing services in the Bitcoin ecosystem work. The paper provides a brief introduction to network mixes and discusses the proposal by Chaum in 1981, which included combining multiple inputs from different users in order to confuse the communications trail. Due to the design of the Bitcoin system though it's not possible to use the Chaum system as it is.

Usually how those mixing services work is by owning a set of addresses to which users can send Bitcoins to. Once a payment has been initiated and confirmed the amount of Bicoins is transferred to the destination address using a different address that is not linked to the the first address. Operators of such services usually charge a small transaction fee. Some services can be accessed from the clear web and some require the use of anonymisation networks, e.g., Tor.

The paper also makes mention of how attackers could monitor the inputs and outputs and manage to create a link and some suggestions are provided in order to avoid that, including but not limited to splitting the outgoing transaction into multiple, smaller parts, especially ones that are widely used (e.g. 0.1 Bitcoin or less). In addition, for some of those services, an account might need to be created, which may potentially give an attacker a window for some sort of attack, in case the mixing services' systems are the target of a data breach, since most of them keep track in their own internal logs of all the transactions, incoming and outgoing, with timestamps, etc. A user typically does not have the ability to know for a fact if the logs get deleted after the transaction or at some other point in the future or if the service itself is not compromised and not run by attackers.

Transactions with a large numbers of Bitcoins associated with them are potentially easier to detect since those are more rare than ones with limited number of coins. Again, as mentioned earlier, there is still the risk that the mixing services could be compromised themselves and the end-user has no way of knowing. Therefore one of the methods proposed is the use of multiple services at once in order to further remove the end user from the transaction trail, although that comes with increased costs since traditionally all services charge a mixing fee.

- **Bitcoin wallet heuristics –** Another method by which Bitcoin transactions could be deanonymised is by categorising the nature of the wallets that their are using. In their paper "A Fistful of Bitcoins: Characterizing Payments Among Men with No Names", Meiklejohn et al. [33] offer an insight into how the unique characteristics of the Bitcoin protocol can offer the opportunity to group Bitcoin wallets based on evidence of shared authority and then using re-identification attacks. The authors developed a new clustering heuristic based on change addresses, allowing them to cluster addresses based on change addresses. Two sets of heuristics are presented for linking addresses controlled by the same user, with the goal of collapsing the many addresses seen in the blockchain into larger entities:

  Heuristic 1: If two (or more) addresses are used as inputs to the same transaction, then they are controlled by the same user. By that heuristic the network was partitioned into 5.5 million clusters of users. Big services such as Mt. Gox appear to be using a large number of "pool" addresses to store BTC in order to minimise the risk in case one gets compromised.

Heuristic 2: Although H1 already yields a useful clustering of users, according to the paper, it does not tell the whole story. As mentioned in Chapter 3, one of the idiosyncracies of the Bitcoin protocol is the way currency must be spent. In very simplistic terms: If a user has 10 Bitcoin and wants to spend 1, he must send all 10 to the destination address and he will receive 9 Bitcoin on his change address. This change address is created internally by the Bitcoin client and never re-used; as such, a user is unlikely to give out this change address to other users (e.g. for accepting payments) and in fact might not even be aware of the address unless they manually inspect the blockchain. Therefore, there is potential for clustering not only the input address but also the change address. Because Heuristic 2 does not take advantage of an inherent property of the Bitcoin protocol it does lack robustness in the face of changing patterns in the network.

4.2.2 Bitcoin Architectural Shortcomings

Lastly, we present four architectural shortcomings of the Bitcoin system that can be used for deanonimization purposes:

- **Deanonymisation of clients in Bitcoin P2P network –** In their paper of the same name, Biryukov et al. [9] present an efficient method to deanonymise Bitcoin users by linking pseudonyms to the original IP addresses where the transactions originate from. It is also one of the few papers that specifically makes mention of using Tor as a means of providing an extra layer of anonymity and ways to circumvent it. The authors claim that their methods have a success rate that goes from 11 to 60 percent while also not requiring a significant number of machines to pull off. Among the conclusions is that the level of network anonymity provided by Bitcoin is quite low and there is a lot of room for exploitation.
- **Network traffic analysis –** The Bitcoin protocol is easily recognised by network traffic analysis utilities. Several characteristics can be easily identify Bitcoin network traffic such as the port numbers and the IP addresses of Bitcoin peers. In fact, recent versions of the popular Wireshark utility will automatically characterise Bitcoin Traffic as “Bitcoin Protocol” by default.
- **“Timejacking” attacks –** This is an attack that is, for the time being, purely theoretical and exploits the way the Bitcoin network manages timestamps [**?**]. All participating Bitcoin nodes maintain a counter that represents network time. This is based on the median time of a node’s peers which is sent in the version message when peers connect. The network time counter reverts to the system time, however, if the median time differs by more than 70 minutes from the system time. An attacker that controls a large number of compromised nodes could theoretically slow down or speed up a node’s time counter by connecting as multiple peers and reporting inaccurate timestamps. For instance, a relatively small number of Tor clients could send enough messages to take over the node’s median time.
- **The 51% attack –** This is arguably one of the biggest shortcomings of Bitcoin and targets directly the mining (transaction verification) process. Theoretically feasible but impractical, it requires a hypothetical group with a large number of users colluding in order to manipulate the records that get stored in the blockchain. Assuming that there is such a large number of colluding nodes on

the network that they exceed the 51% of the computing power in the mining process they could potentially influence the results and provide the blockchain with compromised verification results.

4.3 Bitcoin over Tor

Since it has been demonstrated that Bitcoin transactions are not entirely anonymous but pseudo-anonymous and can be linked back to their issuers and receivers, Bitcoin users naturally have been trying to figure out different ways by which they can augment the protocol's privacy features and potentially build upon them in order to further conceal their anonymity online.

One such way was to connect their Bitcoin wallets online through the Tor network. While on paper it sounds like a great idea, recent research in the field has disproved it. The most widely-cited research was done by University of Luxembourg student and fuculty members Ivan Pustogarov, A Biryukov and D Khovratovich [9, 10]. Among their findings, we highlight the following:

- **Routing Bitcoin traffic through Tor –** The Bitcoin protocol has built-in anti-DoS algorithms that works with a reputation-based system. By making use of that system, when a malformed message is sent to a node on the Bitcoin network, the sender is afflicted with a score that varies depending on the type of message that was sent. When that score reaches a value of 100 then the sender's IP address is banned from the network for a 24-hour period.

  What this means in the context of Bitcoin usage over Tor is that it is theoretically possible for an attacker to ban clean and safe Tor exit nodes by making them send malformed messages to the Bitcoin network and then inject their own compromised exit nodes on the Tor network that Bitcoin users will have no option but to use unwillingly.

  The subject can be even further complicated and dangerous if the attacker also has the ability to ban "good" Bitcoin peers on the Bitcoin network. That would mean that the attacker would not only control the communications path to the Bitcoin network but also peers that can validate transactions. That would give the attackers the methods by which they can drop blocks and transactions which would in turn increase the probability of double spending, therefore compromising one of the Bitcoin's alleged innovations and key components. Traffic confirmation attacks could potentially also happen allowing opportunities for deanonymisation and transaction linking.

- **Fingerprinting attacks –** By exploiting the Bitcoin peer discovery protocol it is possible to fingerprint users on the network. The idea is very simple: In order to get the list with the IPs of known Bitcoin clients on the network, clients send GETADDR messages to known peers on the network. In reply they receive ADDR messages with said list. If enough GETADDR messages are being sent, peers will willingly share the IP addresses of all clients that are stored in their database.

  An attacker can easily manipulate the list of known IP addresses that can be sent to other peers. They can include a variety of combination of IP addresses that do not necessarily belong to the Bitcoin network, for example Tor node addresses, VPN addresses, and so on. This effectively sends a *cookie*, as the researchers describe it that can be used to fingerprint the user. Later that Bitcoin client can be queried with a GETADDR message that will make them divulge its list

of known IP addresses and in the process reveal that combination of selected IP addresses that were passed on to its database in the form of the afforementioned cookie.

This method of fingerprinting has a problem though. Each client can have stored in their database 20,480 addresses at a time. Every time a client will resume operation on the network, they will automatically connect to eight peers simultaneously and will request typically 2500 addresses. What this means, that this cookie will have a limited lifespan before it becomes overwritten by new addresses.

- **Sybil Attacks –** By exploiting the design of the Hidden services feature of Tor, it is relatively easy for an attacker to initiate Sybil attacks. A Sybil attack can occur when a reputation-based system is fooled by forging identities in peer-to-peer networks. In the case of Bitcoin over Tor, something that could provide one with a possible for attack is to fill up all the good nodes' connection slots, so that new nodes can connect only to an attacker's nodes. A way to make this possible would be by broadcasting the IP addresses of legitimate Bitcoin network nodes, but provide fake port numbers, so that any broadcast of those same IP addresses with the real port numbers is rejected because a Bitcoin client, stupidly, only considers the IP address, which it thinks it already knows.

4.4 Summary

In this chapter we observed that the Tor and Bitcoin networks, while providing a certain level of confidentiality to their users, they can have their anonymity features compromised due to a variety of reasons. Chief among them is the fact that they both are decentralised networks that rely on user-provided infrastructure in order to operate. That fact alone could potentially allow the traffic that is being generated by said networks and their users through compromised channels. Although both systems' developers are quick to react to fix newly-discovered vulnerabilities, the ones that are presented in this chapter can not necessarily be patched out as they are inherent to the design of the system.

In the case of Tor, the protocol's vulnerable points are examined and separated into three categories: Attacks done at the protocol-level, attacks that can correlate traffic and fingerprinting. In the case of Bitcoin, two main kinds of vulnerabilities are looked into: ones that can be triggered as a result of the users' actions and underlying flaws or shortcomings in the design of the Bitcoin protocol itself that when exploited could compromise the anonymity of its users. On the subject of Bitcoin usage through Tor network, there is a variety of attacks that have proven to be possible and feasable, mostly through the research done by Pustogarov et al. [9]. An attacker that can a large number of nodes on the Tor network can route Bitcoin traffic due to shortcomings in the design of the latter's built-in, Anti-DoS protection mechanisms.

The ethical aspect of research done on methods by which the privacy and anonymity features of both Tor and Bitcoin has been the subject of many published papers. One thing that is always taken into consideration is how Bitcoin and even more importantly Tor do in fact have a legitimate *raison d'être*. Tor has been used extensively in regimes where freedom of speech and expression is forbidden or otherwise punished. As such, researchers that performing studies in the field of deanonymisation always make special mention of that fact and how they would generally shy away from performing tests on the live networks themselves. In their paper which focuses on Tor statistical data, Loesing et al. [30]

ponder about the consequences of publishing data that are measured from the live Tor network and they propose a set of guiding principles which should be followed when measuring Tor data.

5 Tor and Bitcoin Forensics

This chapter examines the forensics artifacts left behind by the Tor and Bitcoin protocols and the most common front-end applications a typical user might use in order to take advantage of those two protocols. It should be reiterated at this point that this thesis is only concerned with investigations conducted on so-called *traditional* computer systems, i.e. personal computers running recent versions of Operating Systems such as Microsoft's Windows, Apple's macOS and different flavours or distributions of GNU/Linux. As such, the little-to-no mention is provided to new forms of computing equipment, such as tablets and smartphones, devices that are capable of running Tor and Bitcoin software.

The Microsoft Windows family of operating systems is the the most widely-used system software Personal Computers. According to statcounter global stats, as of January of 2018, Windows has 82,6% of the marketshare, followed by Apple inc.'s macOS at 13,06%. The rest 4,26% is spread among various other Operating Systems, including different versions of GNU/Linux[1]. As such, this chapter will focus on those three major desktop operating systems according to their market share.

5.1 Areas of Forensic Focus on Operating Systems

This section is meant to familiarise the reader with the most common areas in which modern Operating Systems might store user files, application files, configuration settings and so on that could potentially be used as evidence-related artifacts. Areas of focus include (but are not limited to):

- application installation directories;
- user home directories and user-specific configurations;
- operating system services that log user activities.

5.1.1 Microsoft Windows

In Windows, there are numerous potential sources for evidence extraction. We highlight six relevant ones: application directory structure, Windows pre-fetch system, Windows Registry, Windows virtual memory and pagefile, Windows Search, and Cortana.

[1] `https://statcounter.com/`

- **Application directory structure –** Typically an application and the corresponding files that are required for it to function properly will be installed inside their own folder somewhere in the system. A user will sometimes be given the option to select the folder in which a specific application and its related files will be installed to, but that is not always the case. In addition, depending on the Operating System, the folder and the path that leads to it will also be different. For example, in the Microsoft Windows family of Operating Systems, during installation applications by default will ask to be installed in a sub-directory inside the `C:\ProgramFiles` directory, or inside the `C:\ProgramFiles(x86)` folder if they are 32-bit applications.

- **Windows pre-fetch system –** Introduced all the way back in 2001 with Windows XP, it is a way to speed up application launch times by keeping track of different statistics regarding application usage. For that reason, the Windows keeps files in a specific folder with the name of the application. Some of the information that can be acquired when examining the prefetch system on Windows are: installation dates, application version, first and last execution dates, timestamps displaying the exact time and date of last few executions, the path from which the application was launched from, and total number of executions and files in the filesystem that are used by the application itself.

- **Windows Registry –** The Registry is an important system component of Windows and it is a centralised database-like list containing system and user settings. NTUSER.DAT is the file that includes settings associate with individual users.

- **Windows virtual memory and pagefile –** The swap file allows an Operating System to act as if the computer it is installed in more available RAM memory that it does in reality. This is also referred to as Virtual Memory. The OS uses the swap file to swap information back and forth as physical RAM becomes populated over the course of a computer's usage. The data in the swap file can belong to the Operating System itself, or belong to different applications and files and can remain there for long periods of time.

- **Windows Search –** Windows Search is a system process that is enabled by default and it builds a full index of the files on a Windows computer by crawling the file system upon the system setup and monitors changes in file operations on the file system as the user keeps on making use of their computer. On modern Windows systems it will record number of files on the `C:\ProgramFiles\Microsoft\Search` folder and the included sub-folders. In previous versions of Windows (such as the still-fashionable Windows XP) the index storage used to be located in the `C:\DocumentsandSettingsApplicationData\AllUsersMicrosoftSearch` directory and included subdirectories.

- **Cortana –** Starting with the 2015 release of version Windows 10, Microsoft includes in their Operating System a Personal Digital Assistant simply referred to as "Cortana", that is named after a fictional AI character from Microsoft's Halo series of video games. It can be used for a variety of different functions, either with keyboard-typed or oral commands. Some of those functions include but are not limited to: setting calendar events and reminders, searching the user's system locally for files, folders and applications, launch applications, performing web queries and answering simple questions such as what is the current time and date and what the meaning of life is (hint: the answer is always 42).

There are a few artifacts of forensic interest that are generated by the use of Cortana and they have to do with the assistant's ability to accept voice commands and launch programs.

Voice commands: Cortana stores the commands orally inputted by the users in plain WAV (Waveform Audio Format) files and stores them in a subdirectory inside their home directory, more specifically in the `C:\Users\[USERNAME]\AppData\Local\Packages\Microsoft.Windows.Cortanacw5n1h2txyewy\LocalState\LocalRecorder\Speech` directory. So, assuming for example that a user asked Cortana to open an application (such as Tor or a Bitcoin Wallet) for them, there would be an audio clip stored on their computer containing their voice asking for said application to be opened. It is worth noting that Cortana keeps a maximum of 8 audio commands stored and then proceeds to overwrite them.

5.1.2 Apple macOS

As mentioned earlier, Apple Inc.'s macOS not only has a sizeable marketshare in the desktop Operating System market but all the tools required to use Bitcoin, Tor and get involved in the dark net cybercrime scene are also available for it as well. It shares many similar traits with Microsoft Windows but it also includes some key differences, such as that there is no centralised database that stores user and application information (i.e. Registry) and the directory structure where application settings are stored also differs from that of Windows. Areas of forensic interest in macOS include:

- **Application Installation Directory –** Applications built for Apple's desktop Operating System macOS will either automatically or ask to be installed inside the "/Applications" folder, regardless of architecture.
- **Apple System Log –** It is a process on macOS computers that logs messages from different parts of the operating system, including applications that run on it. They usually reside in the "/var/log/" directory.
- **Crash Reporter and Diagnostic Messages –** The crash reporter collects information from applications that have crashed due to various reasons. It usually stores information in the following folders:

  /Library/Application Support/CrashReporter/

  /var/log/DiagnosticMessages
- **FSEvents –** macOS has a built in function that allows for applications to register for notifications of changes to a given directory tree. Whenever the filesystem is changed, the kernel passes notifications via the special device file /dev/fsevents to a userspace process called fseventsd. Cached information generated by this service are stored inside the "/.fseventsd/" folder.
- **HFS+, the macOS filesystem –** It used to be the default filesystem for Macintosh computers from the release of Mac OS 8.1 in 1998 until it was succeeded by the Apple File System in 2017 with the release of macOS 10.13 High Sierra. Among other things it includes support for journaling which is a feature that maintains a special file called a journal that is used to repair any inconsistencies

that occur as the result of an improper shutdown of a computer. Areas of interested in a macOS computer that makes use of the HFS+ filesystem are:

/.hotfiles.btree

/.journal

The functionality described above has been altered on Macintosh computers that make use of macOS 10.13 High Sierra in conjunction with the new Apple File System.

- **User preferences and application preferences –** As mentioned earlier, macOS does not include a centralised database like the Windows Registry in order to store application and user settings. Instead applications store their settings in special files that usually exist in subdirectories inside the user's home folder. The main directory of interest here exists in this path:

  /Users/[USERNAME]/Library/Preferences

  The relevant files typically have a filename such as "com.apple.finder" and the file extention .plist

- **Spotlight Search –** The default method of searching for files, folders and applications among other things using the graphical user interface on a Macintosh computer. It works by indexing all files and directory structures on a system and storing references to those items in its own database. The main folder of interest exists at the "/.Spotlight-V100/Store-V2/" path.

- **macOS Swap and Hibernation files –** Like most operating systems do, macOS also relies on using swap files for memory management purposes. On macOS the swapfile exists at "/var/vm/swapfile0". In addition, inside the same directory one can find the OS's hibernation file, aptly titled "sleepimage".

- **Temporary user and system files –** Temporary Data: It is normal behaviour for an application to create a series of temporary files and folders during when it loads and during operation. macOS stores those files inside the "/var/folders/" and "/tmp/" subdirectory.

5.1.3 GNU/Linux Running the GNOME Desktop Environment

On Linux and other Unix-like operating systems the situation is a bit more complex as there is a greater degree of variation in the naming and placement of system files and directories. As such, it would take a disproportionate amount of time to describe each different distribution's quirks compared to their marketshare. As such, this section will focus on the aspects that are common on most of them, especially when they utilise the GNOME desktop as their Graphical User Interface.

- **Application Installation Directory:** Compared to Windows and macOS, on traditional Unix systems there is not a singular directory where application files can be installed (i.e. "Program Files" and "/Applications"). A large part of an application's files may exist in one or many directories across the system, the most popular of which include "/usr/share/", "/usr/local/", and "/opt/". In addition, the scripts that are responsible for launching the applications may be placed in directories such as "/bin/" and "/usr/bin".

- **Bash history**: Bash is typically the default shell interface on GNU/Linux systems and traditionally it logs commands typed by the user by default. The area of interest here is in the file ".bash_history" inside the directory "/home/[USERNAME]/"

5.2 Bitcoin Specific Forensics

This section will focus exclusively on techniques that can be used in order to extract digital artifacts from the Bitcoin system. As was mentioned in chapter 3, in order for a user to begin making transactions inside the network, they will to have a piece of software called "wallet". Therefore, most of the focus will be on that area.

5.2.1 Private Keys, Addresses, and the Blockchain

As stated before, the Bitcoin network has a method by which it keeps track (in a pseudo-anonymous fashion) of every transaction that has taken place inside the network since its inception in 2009. That method is a publicly-available database simply referred to as “The Blockchain”. The information inscribed in it is not encoded in plaintext form, meaning that if one tried to see the contents of a blockchain-related file with a text editor such as the Windows Notepad they would see only indecipherable text.

In this article [14] that first appeared in the Forensic Focus magazine, what is presented is opportunities for investigators to look into Bitcoin transactions. It also acts as a mini-guide for a freely-available Python tool that accompanies it article that can carve out Bitcoin artifacts from a computer[2]. It begins by mentioning an alleged rule of so-called “advanced Bitcoin users” which is that if one does not control the private keys they do not control the Bitcoins attached with them also.

On the subject of de-anonymising Bitcoin users it heavily references the research done by Pustogarov et. al [9] and it talks about de-anonymising up to 60 percent of Bitcoin clients on the network. The method is based by "fingerprinting" users based on the connections they have to other nodes on the Bitcoin network. If an attacker is connected to enough nodes, these announcements can be watched and fingerprinting can be done. If the user connects to the Bitcoin network through Tor or a VPN the IP of the exit node would be advertised to the Bitcoin Network but connections could still be made. In addition, using Tor to hide the identity of oneself, could actually expose them to man-in-the-middle attacks and attackers have the ability to potentially ban all Tor exit nodes from connecting to the Bitcoin network by abusing Bitcoin’s spam and anti-DoS protection mechanisms

Although highly unlikely, due to the additional security precautions required to keep Bitcoin storage safe, the best case scenario for a forensics investigator would be to seize a computer and find the private keys in plain text form. Some users might hide their private keys inside small storage devices such as a raspberry pi, a USB pen drive with a Linux installation, or even in paper form with QR codes. The paper concludes with some discussion about file and memory forensics on a local level. It talks about the format in which Bitcoin software stores various artifacts on a machine where a wallet is installed and

[2] https://gist.github.com/chriswcohen/7e28c95ba7354a986c34

how forensics tools could be used in order to extract those artifacts from a machine. It goes into more detail as to how the python tool that accompanies this report can be used.

5.2.2 Bitcoin Core Wallet artifacts

The Bitcoin Core client (formerly known as Bitcoin-Qt[3]) is the reference implementation of the Bitcoin software based on the code originally authored by "Satoshi Nakamoto" [36]. Being a front-end application that allows one to connect to the Bitcoin network, it only natural to assume that it will leave digital artifacts on a system. Bitcoin core usually stores files inside the User's home directory. On Windows systems, the path is `C:\Users\USERNAME\AppData\Roaming\Bitcoin`.

Inside the directory, one can observe multiple files, but the most important ones are wallet.dat and debug.log. As the names may imply, the former file contains information that describes a user's wallet data, including addresses while the latter contains debugging information that include crucial information such as timestamps that mark the software's communication with the Bitcoin network.

A user could move their wallet data, either to another computer or for backing up purposes simply by copying the wallet.dat file and moving it to another place. They can try to keep the contents of the file ambiguous by altering the name to something else. Still, if one is suspicious of its contents they can parse the file with a hex editor and check the inscriptions at the offset 0x12 for traces of the string "b1", which could potentially mean that it is a Bitcoin wallet file [4] Information inside the wallet.dat file is not stored in plaintext, so opening the file with a text editor such as Notepad or Vi will not reveal any information that can be understood with a simple visual observation.

Support for the discovery and extraction of Bitcoin artifacts is slowly but steadily being introduced to popular forensics utilities, such as Magnet Software's popular, internet-related evidence utility Internet Evidence Finder, or IEF[5]. Starting with version 6.1, Bitcoin addresses can be recovered from a wallet, as well as queries on the Bitcoin network from log files created by the Bitcoin wallet software. This can automate the search for traces of Bitcoin software and wallets on a local machine. By parsing files in a system, IEF can locate Bitcoin-related files and from them extract artifacts organised in a report-like form that can contain information such: addresses and transaction labels and timestamps.

5.3 Tor-specific Forensics

As stated before, the Tor Browser Bundle is the easiest way one can gain access to the Tor network. The Tor Browser is based on a modified version of the Mozilla Firefox web browser with custom modifications to allow a user to access Tor and some additional plugins for extra security, such as forced HTTPS connections and the option to disable javascript entirely.

[3] `https://bitcoin.org/en/version-history`
[4] `https://www.garykessler.net/library/file_sigs.html`
[5] `https://www.magnetforensics.com/computer-forensics/bitcoin-forensics-a-journey-into-the-dark-web/`

This section is going to be based on the papers by Runa A. Sandvik and Aron Warren entitled "Forensic Analysis of the Tor Browser Bundle on OS X, Linux, and Windows"[6] and "Tor Browser Artifacts in Windows 10" [25] respectively. These two papers go through artifacts left behind by Tor on the three main operating systems and the authors present some investigation scenarios based on some assumptions as to how the user might make use of the browser bundle. Next, we present some of the most relevant artifacts in their own sections.

5.3.1 Application Installation Files

When the Tor browser gets installed on a system, be it Windows, macOS or GNU/Linux, it will as a consequence leave behind a trail of files that are scattered throughout the computer's storage medium. The most common way to download the Browser bundle is to visit the Tor Project's website and download the installation file. Typically, the Windows version of the installer will have a name like "torbrowser-install-7.5_en-US.exe" where 7.5 is the currently available version of the bundle and en-US will differ depending on the language that the user chooses. If for example a user chooses to download the Brazilian Portuguese version of the browser bundle, the file name will look something like this: "torbrowser-install-7.5_pt-BR.exe". Under Apple's macOS the installation file will be called "TorBrowser-7.5-osx64_en-US.dmg" where osx64 signifies the software's Macintosh-specific, 64-bit architecture. And finally users of GNU/Linux will be presented with a compressed file with a filename like "tor-browser-linux32-7.5_en-US.tar.xz" or "tor-browser-linux64-7.5_en-US.tar.xz".

A regular web user will most likely use their existing browser of choice (i.e. Mozilla Firefox, Google Chrome etc) in order to get the installation file for the Tor Browser onto their computers. Typically web browsers tend to store files downloaded from the web inside a dedicated folder that modern Operating Systems maintain inside the user's home directory aptly called "Downloads". Depending on the users habits of how they manage their system's Downloads folder, the aforementioned installation files might still exist there and it would be a good first place to search for traces of Tor during an investigation.

It should be mentioned here that in the case of GNU/Linux distributions, users have the option of installing the Tor browser via what is called a *repository*. This allows for the automation of the installation process, where the user only needs to input a handful of commands on their Terminal application of choice, and the system will take care of downloading and installing the selected application. In addition, the Tor Browser Bundle's curators advertise it as being able to run in a self-contained, *portable* form. That means that a user could potentially have the browser's application files on a storage medium that exists outside of their computer system (e.g. USB flash drive) with no need for installing the application on their system. This would further complicate the work of an investigator as it would require physical access to the portable drive and potentially scanning a Windows PC's Registry for instances of said drive having being used in the past on that PC[7].

[6] https://www.cnet.com/news/bitcoin-wont-be-the-dark-webs-top-cryptocurrency-for-long/?ftag=COS-05-10aaa0g&utm_campaign=trueAnthem:+Trending+Content&utm_content=5a7d0b349ebbef00076e256f&utm_medium=trueAnthem&utm_source=twitter

[7] https://docs.microsoft.com/en-us/windows-hardware/drivers/usbcon/usb-device-specific-registry-settings

Operating System	Installation File	Downloads Folder Location
Windows	torbrowser-install-7.5_en-US.exe	C:/Users/[USERNAME]/Downloads
macOS	TorBrowser-7.5-osx64_en-US.dmg	/Users/[USERNAME]/Downloads
GNU/Linux	tor-browser-linux64-7.5_en-US.tar.xz	/home/[USERNAME]/Downloads

Table 5.1: Download directory names and locations

5.3.2 Tor Browser Bundle Directory Structure

Upon launching Tor's installation application under Microsoft Windows, the default installation directory that will be recommended for the Browser Bundle will be on the current user's Desktop. The reason for that default installation location lies on the fact that the Tor Browser is meant to be a *portable application*, meaning that upon installation it should not, in theory, leave any traces on the system that can advertise its existence, such as adding registry values and that it can be moved freely to any other destination on a system without it affecting its ability to function as intended. So, depending on how careful (or not) a user is, the best place to begin an investigation for traces of the Tor Browser Bundle would be said user's Desktop folder, which itself is typically a sub-directory inside the user's home directory. For users of macOS, upon opening the downloaded file, they will be presented with the option of transferring Tor's application file inside the system's "/Applications" folder 5.1. Users of GNU/Linux systems will need to unpack and unzip the downloaded file and use the browser from the extracted folder.

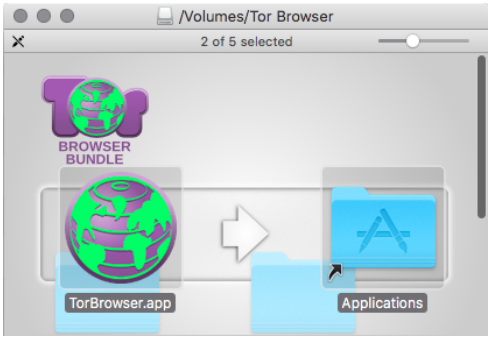

Figure 5.1: Tor installation under macOS

Operating under the assumption that on their Windows system, the user has chosen to install the browser on the default, recommended directory on their desktop. That would mean that the path to the Browse bundle will most likely look something along the lines of:

```
"C:\Users\USERNAME\Desktop\TorBrowser"
```

Upon examining that folder one can see multiple sub-directories but from a forensics perspective only a select few are actually relevant. Starting with the root of the installation directory, one can immediately observe a sub-directory named *Browser* and a shortcut to launch the application, called "Start Tor Browser" (Figure 5.2).

Going deeper inside that first *Browser* directory, we can observe a list of files with the suspicious-looking day time stamp of "1/1/2000 12:00 AM". This was done on purpose by the program's curators[8].

[8] https://www.torproject.org/docs/faq.html.en#Timestamps

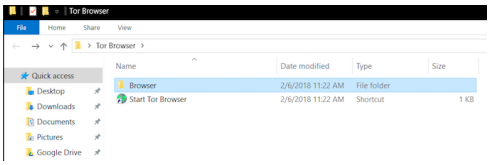

Figure 5.2: Main Tor Directory in a typical Windows 10 installation

What is also immediately obvious is a few other directories present with a different timestamp. The most interesting artifacts that could be extracted from this search exist inside the "/Tor Browser/Browser/Tor Browser/Data/" sub-directory. In it there are two more sub-directories called "Tor" and "profile.default". Inside the first, "Tor", we could find two very important files when it comes to forensics: "state" and "torrc". The state file contains the "last execution date" which is a time stamp that indicates the last time Tor was opened on the system. The torrc file contains the path from which the executable binary was launched from. The other folder, "profile.default", includes even more interesting artifacts. It follows the same structure as a typical Firefox installation. The most interesting files here are compatibility.ini and extention.ini.

5.3.3 Tor Browser artifacts in Windows Prefetch

As mentioned earlier, Prefetch files are a method utilised by Microsoft Windows in order to speed up application launch times based on the frequency that they are opened by the user. For that reason a series of files are stored insider the subdirectory *Prefetch* which is located inside the Windows installation directory, normally `C:\Windows`. As shown in Figure 5.3, the most recent version of the Tor browser bundle as of February 2018 is still susceptible to this known "bug"[9].

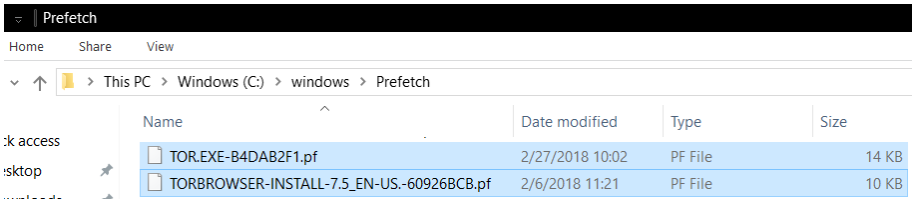

Figure 5.3: Tor Browser artifacts in Windows Pre-fetch

When using a tool such as WinPrefetchView, that can read and decode the contents of the .pf files, a lot of valuable information can be extracted, including the installation folder of the Tor browser bundle, and a list of the most frequent files used by the application.

As shown in Figure 5.4, when viewing the contents of the Tor prefetch file, TOR.EXE, we can observe a list of files closely associated with the Tor browser executable. Besides the obvious sign that the Tor browser was at some point installed and used on a machine, of significant importance are the files with names starting with "CACHED-". They are files that the Tor browser uses in order to store cached information related to the available nodes in the Tor network and their certificates.

[9] `https://trac.torproject.org/projects/tor/ticket/8916`

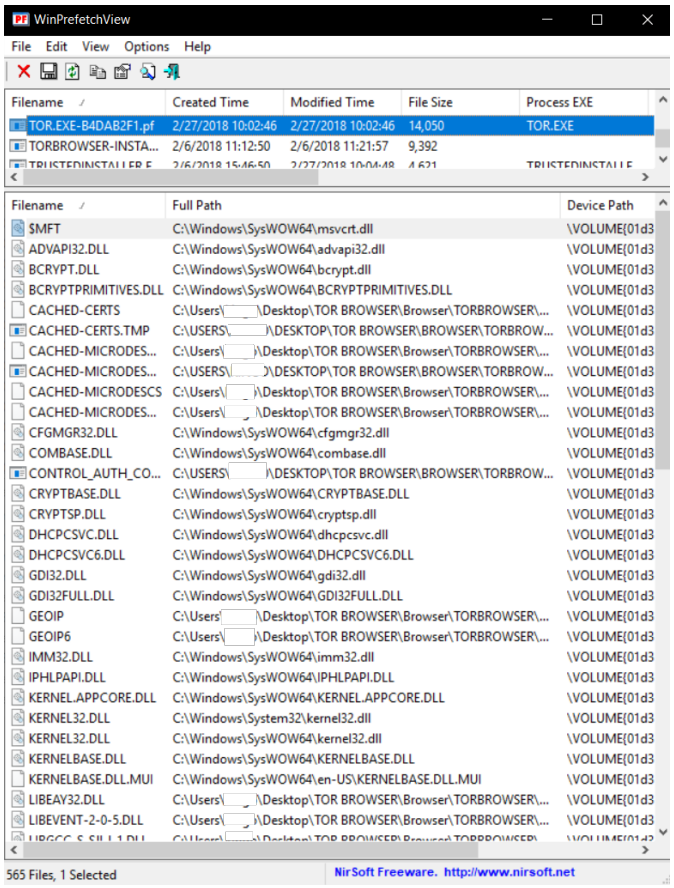

Figure 5.4: Contents of the Tor Prefetch File

5.3.4 Tor Artifacts in Windows Registry

As mentioned, the Windows Registry is a hierarchical database that is used by the Operating System in order to store information about itself, installed applications, user information and devices attached to a computer. If one were to use the tool called *Registry editor* that comes bundled with Windows they would see a contiguous, hierarchical tree that is composed by artifacts that appear as folders. They are called hives and they are the storage space of keys and subkeys and values that dictate different system and application settings. But while the appearance of the singular, hierarchical tree might give one the impression that the registry is comprised on a single file, it's anything but. For the purpose of this chapter, only the one called NTUSER.DAT will be examined, which exists inside the user's home directory and contains personalised settings.

Among those user-specific settings are Most Recently Used lists (MRU), which contain the last activities performed by a user. This also applies to several other Registry keys that have values named "MRUList". A modern Windows version will keep track of activities such as "Recent Programs Opened" and "Recent Files opened". Figure 5.5, contains a list with the locations in the registry that also tie with Tor usage.

Item	Location in the Registry
Recent Programs Opened	HKEY_USERS\USER\Software\Microsoft\Windows\CurrentVersion\Explorer\ComDlg32\LastVisitedMRU
Save Locations by Filetype	HKEY_USERS\USER\Software\Microsoft\Windows\CurrentVersion\Explorer\ComDlg32\OpenSaveMRU
Recent Files Opened, by Filetype	USERNAME\software\microsoft\windows\currentversion\explorer\recentdocs

Figure 5.5: Windows registry keys and hives

5.3.5 Pagefile and Hibernation File Artifacts

While serving a different purpose, both system pagefile and hibernation file have a unique relationship with a system's volatile memory, RAM. In the case of the pagefile, in order to not run out of addressable RAM, the operating system, on occasion and depending on the circumstance will move data from RAM to a file on the computer's storage medium. The process is called interchangeably paging and swapping. Although arguing the semantics is not important for this investigation, the principle remains the same: the OS will move data into a file on the hard disk when the computer is about to run out of addressable memory.

The Hibernation file serves a different purpose. Modern systems have the ability to fall into a state of hibernation when a user shuts them down in order to be up and running in a very short amount of time compared to the normal boot process. For that reason, during hibernation the Operating System will record the entirety of RAM memory in a single, hidden, system file aptly called the hibernation file. Since both operations deal with writing the contents of RAM into a file on the hard disk, it is obvious that operations regarding user activity can be recorded inside those files.

5.3.6 Tor-related Artifacts Generated by Windows Explorer

Another source of semi-hidden forensic artifacts regarding files and applications a user may have opened is the Thumbnail Cache that is generated by the Windows file system-browsing utility Windows Explorer. Windows stores thumbnails of image files and icons for quick access or via random crawling while indexing. The folder `"C:\Users\USERNAME\AppData\Local\Microsoft\Windows\Explorer"` contains various files with the names "thumbcache_32.db", "thumbcache_96.db" and "thumbcache_256.db" that may contain traces of icons of applications used.

5.4 Network Traffic Analysis and Tor & Bitcoin Forensics

If an investigator is tasked with observing a live system for traces of it being involved in some capacity in a Tor and Bitcoin activities, they would most likely need to resort to using an application called a packet analyser (or packet sniffer). For brevity in this section only two applications of this kind will be examined: Wireshark and TCPDump. Their main and most obvious difference is that the former is a graphical user interface application while the latter is meant for use from the console. At their core,

both applications perform a similar task which is to capture traffics flowing inside a network and store them in a file for future analysis.

Among the information that can be observed inside a network capture file are timestamps containing information about when each individual packets were sent, source and destination IP addresses and ports. Recent versions of Wireshark have built-in support for recognising the Bitcoin protocol when they encounter it on a network. During the Bitcoin client startup process clients will try different ways to connect to other peers. The Bitcoin QT client has some seeds hard coded into the client. This process is documented fairly well in the web article *Analyzing Bitcoin Network Traffic Using Wireshark* by Sam Kear [41]. Upon boot, Bitcoins peers will search for DNS seeds that are hard coded into the client, and if the client fails to contact the DNS seeds, it will fall back to a list of hard coded IP addresses.

5.5 Existing Applied Forensic Techniques and Tutorials

The subject of this section is the paper *Bitcoin Forensics: A Tutorial* [37]. It dates all the way back from 2011 and predates the rise and fall of the Silk Road cryptomarket. It is an outline of the methods used by the Middlesex University in London, UK in order to teach its students methods to investigate Bitcoin transactions online. The Bitcoin core software suite was used due to its ability to store locally the entirety of the blockchain. The forensic tool NUMISIGHT was also used in order The crime had been setup in such a way where only a limited amount of clues were available to the students and they had to piece the whole story together by analysing Bitcoin transactions and online activities in the Dark Web.

The assessment was split into a series of tutorials. In the first tutorial, the students were familiarised with the technologies utilised by Bitcoin, mostly the cryptographic protocols. The exercise here was to setup all the relevant accounts and wallets that the learners will use to carry out transactions. A copy of the log sheet was given out for them to record details of their actions.

In the second and third tutorials, the storage and acquisition of Bitcoins was tackled as well as the creation of addresses. The learners are shown how the bitcoin addresses are converted from the public key form through hashing and finally encoded as a Base58 string. Various methods of acquiring Bitcoin were examined and also judged on their ability to provide some sort of privacy to the person seeking to acquire them. An examination of addresses also takes place. The learners are shown how each transaction is broadcast as a message to the network, which is then propagated to each and every node for validation.

In the fourth tutorial, the learners look at the peer-to-peer network and examine its ability to reach consensus in the absence of a central control authority. In the fifth tutorial, the blockchain is examined in much more detail than before. The learners are shown how the hash of the previous block header field is the essential piece of data in maintaining the chain of blocks, and prevents modification of the data within the blockchain.

In the sixth tutorial, the learners partake in the hypothetical crime outlined earlier and they are also introduced to methods in which the Bitcoin network might become compromised. Topics discussed include: *forking*, the concept of the 51% attack, denial of service, and the ability to detect the IP address of the user who makes a transaction. The seventh and final tutorial deals with blockchain analysis and

how to make sense of transaction history and chaining. The learners are reminded of the benefits that the blockchain provides to investigators, in particular, the transparency and auditability.

The authors argue that while the paper is not trying to develop a framework of Bitcoin investigations, the suggestions could be applicable for future investigations.

5.6 Summary

The aim of this chapter was to provide a technical guide as to how an investigator could This chapter we went through all the possible forensics places on a Windows, Mac or Linux personal Computer.

Security incidents can occur in the confidence that a system, either hardware, software or a combination of the two will function as intended (or as advertised) and would not fail under certain conditions. Both Tor and Bitcoin promise to hide the traces of user activities but as further examination on later parts of this dissertation will attempt to show, traces can be left behind unless paranoia on the part of the user makes them take additional steps in order to further hide their tracks.

In computers running Microsoft Windows as their Operating System a great deal of information related to user activity is recorded in the system Registry that upon examination could provide an investigator of clues as to what that activity was and even when it was conducted. Additionally, newer technologies introduced in recent versions of Windows that were meant to either improve the system's performance or aid users in their day-to-day tasks, can be a great source of information regarding user activities when examined from a forensic perspective.

6 Conclusions

The main contribution that this thesis attempted to offer was the creation of a survey of existing literature in the field of Tor and Bitcoin vulnerabilities and then how a forensics investigator could make use of that information in order to examine those two systems in the event that they were used in some form of electronic crime. Bitcoin is the most widely-used alternative payment system, and Tor is the most widely-used anonymity network, used in many countries around the world, including those that prohibit freedom of speech.

Recent advancements in the fields of computer technology and cryptography have made the actions that take place in the darkest corners of the web almost invisible to prying eyes and more importantly when it comes to crime invisible to the eyes of law enforcement.

The examination of the literature in the field of Tor and Bitcoin forensics, reveals that the front-end software that is used to facilitate the connection to both networks can in certain cases leave behind digital artefacts that can be used to either deanonymise Bitcoin transactions and expose wallet addresses. The Tor Browser can under certain conditions leave behind a plethora of trails related to user activity in areas such as the Windows Registry, Page File and RAM.

While Tor and Bitcoin have shown to perform according to the intend of their respective designers, research conducted on examining the operating effectiveness of both systems has shown that their ability to provide anonymity to users while being used can be compromised. And while both systems promise to hide the traces of user activities, traces of their usage can be left behind on devices that their users install the necessary front-end software to make use of them.

6.1 Future Developments

According to a research conducted by Recorded Future, a company that specialises in real-time threat intelligence, Bitcoin is set to lose its place as the dominant payment system in the dark web in the near future[1]. The report entitled “Litecoin Emerges as the Next Dominant Dark Web Currency” argues that Bitcoin has simply become too popular in order to be effectively used as an efficient method of payment in the dark web, and that it will probably be replaced by other cryptocurrencies, such as Litecoin and Dash. The transition period from Bitcoin to other Cryptocurrencies will take place was estimated to be six to twelve months (as of February 2018).

[1] https://www.cnet.com/news/bitcoin-wont-be-the-dark-webs-top-cryptocurrency-for-long/?ftag=COS-05-10aaa0g&utm_campaign=trueAnthem:+Trending+Content&utm_content=5a7d0b349ebbef00076e256f&utm_medium=trueAnthem&utm_source=twitter

The more traditional forms of computing, meaning desktop and laptop computers are declining in sales. If we were to count smartphones and tablets and potentially wearable devices such as smartwatches as “computers”, Android is be the number one used operating system worldwide. Apple’s iOS also has a small but measurable and consistent marketshare. iOS and Android are similar and different in how they handle things such as: file operations, filesystem journaling, temporary file creation, and most importantly “full disk” encryption on the devices that are installed on. As such, most of the architectural flaws and exploits and the forensic techniques discussed in this thesis do not apply to those operating systems and devices.

Bibliography

[1] ABRAMOVA, S., SCHÖTTLE, P., AND BÖHME, R. Mixing Coins of Different Quality: A Game-Theoretic Approach. In *International Conference on Financial Cryptography and Data Security* (2017), Springer, pp. 280–297.

[2] AKED, S., BOLAN, C., AND BRAND, M. Determining what characteristics constitute a Darknet.

[3] ANTONOPOULOS, A. M. *Mastering Bitcoin: unlocking digital cryptocurrencies*. " O'Reilly Media, Inc.", 2014.

[4] ASTOLFI, F., KROESE, J., AND VAN OORSCHOT, J. I2p - The Invisible Internet Project. Web Technology Report, Media Technology, Leiden University, 2015.

[5] BARBER, S., BOYEN, X., SHI, E., AND UZUN, E. Bitter to better—how to make bitcoin a better currency. In *International Conference on Financial Cryptography and Data Security* (2012), Springer, pp. 399–414.

[6] BARCELO, J. "User Privacy in the Public Bitcoin Blockchain. *URL: http://www. dtic. upf. edu/ jbarcelo/papers/20140704_User_Privacy_in_the_Public_Bitcoin_Blockc hain/paper. pdf (Accessed 09/05/2016)* (2014).

[7] BARRATT, M. J., AND ALDRIDGE, J. Everything you always wanted to know about drug cryptomarkets*(* but were afraid to ask). *International Journal of Drug Policy 35* (2016), 1–6.

[8] BERGMAN, M. K. White paper: the deep web: surfacing hidden value. *Journal of electronic publishing 7*, 1 (2001).

[9] BIRYUKOV, A., KHOVRATOVICH, D., AND PUSTOGAROV, I. Deanonymisation of clients in Bitcoin P2p network. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security* (2014), ACM, pp. 15–29.

[10] BIRYUKOV, A., AND PUSTOGAROV, I. Bitcoin over Tor isn't a good idea. In *Security and Privacy (SP), 2015 IEEE Symposium on* (2015), IEEE, pp. 122–134.

[11] BUXTON, J., AND BINGHAM, T. The rise and challenge of dark net drug markets. *Policy Brief 7* (2015).

[12] CELESTINI, A., ME, G., AND MIGNONE, M. Tor Marketplaces Exploratory Data Analysis: The Drugs Case. In *International Conference on Global Security, Safety, and Sustainability* (2017), Springer, pp. 218–229.

[13] CHAUM, D. Blind signatures for untraceable payments. In *Advances in cryptology* (1983), Springer, pp. 199–203.

[14] CHRIS W COHEN. Forensics and Bitcoin, Jan. 2015.

[15] CHRISTIN, N. Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In *Proceedings of the 22nd international conference on World Wide Web* (2013), ACM, pp. 213–224.

[16] CLARK, T. Digicash files chapter 11. *Cnet News. com, Nov 4* (1998).

[17] CLARKE, I., AND OTHERS. *A distributed decentralised information storage and retrieval system*. PhD Thesis, Master's thesis, University of Edinburgh, 1999.

[18] COX, J. The People Who Risk Jail to Maintain the Tor Network, Apr. 2015.

[19] DINGLEDINE, R., MATHEWSON, N., AND SYVERSON, P. Tor: The second-generation onion router. Tech. rep., Naval Research Lab Washington DC, 2004.

[20] FORCE, A. C. . S. P. A. T. Anticounterfeiting on the Dark Web. Tech. rep., International Trademark Association, Apr. 2015.

[21] GAYLE, D. Online market 'is turning drug dealers from goons to geeks'. *The Guardian* (Feb. 2016).

[22] GERVAIS, A., KARAME, G., CAPKUN, S., AND CAPKUN, V. Is Bitcoin a decentralized currency? *IEEE security & privacy 12*, 3 (2014), 54–60.

[23] HARDY, R. A., AND NORGAARD, J. R. Reputation in the Internet black market: an empirical and theoretical analysis of the Deep Web. *Journal of Institutional Economics 12*, 3 (2016), 515–539.

[24] HUBER, M., MULAZZANI, M., AND WEIPPL, E. R. Tor HTTP Usage and Information Leakage. In *Communications and Multimedia Security* (2010), vol. 6109, Springer, pp. 245–255.

[25] JEFFRIES, A. FBI seizes underground drug market Silk Road, owner indicted in New York, Oct. 2013.

[26] JOHNSON, A., WACEK, C., JANSEN, R., SHERR, M., AND SYVERSON, P. Users get routed: Traffic correlation on Tor by realistic adversaries. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security* (2013), ACM, pp. 337–348.

[27] KIRAN, M., AND STANETT, M. Bitcoin risk analysis. *NEMODE Policy Paper* (2015).

[28] LACSON, W., AND JONES, B. The 21st century darkNet market: Lessons from the fall of silk road. *International Journal of Cyber Criminology 10*, 1 (2016), 40.

[29] LING, Z., LUO, J., YU, W., FU, X., JIA, W., AND ZHAO, W. Protocol-level attacks against Tor. *Computer Networks 57*, 4 (2013), 869–886.

[30] LOESING, K., MURDOCH, S. J., AND DINGLEDINE, R. A case study on measuring statistical data in the Tor anonymity network. In *International Conference on Financial Cryptography and Data Security* (2010), Springer, pp. 203–215.

[31] MARTIN, J. *Drugs on the dark net: How cryptomarkets are transforming the global trade in illicit drugs*. Springer, 2014.

[32] MARTIN, J. Lost on the Silk Road: Online drug distribution and the 'cryptomarket'. *Criminology & Criminal Justice 14*, 3 (2014), 351–367.

[33] MEIKLEJOHN, S., POMAROLE, M., JORDAN, G., LEVCHENKO, K., MCCOY, D., VOELKER, G. M., AND SAVAGE, S. A fistful of bitcoins: characterizing payments among men with no names. In *Proceedings of the 2013 conference on Internet measurement conference* (2013), ACM, pp. 127–140.

[34] MOORE, R. *Cybercrime: Investigating high-technology computer crime*. Routledge, 2010.

[35] MOSER, M., BOHME, R., AND BREUKER, D. An inquiry into money laundering tools in the Bitcoin ecosystem. In *eCrime Researchers Summit (eCRS), 2013* (2013), IEEE, pp. 1–14.

[36] NAKAMOTO, S. *Bitcoin: A peer-to-peer electronic cash system*. 2008.

[37] NEILSON, D., HARA, S., AND MITCHELL, I. Bitcoin forensics: a tutorial. In *International Conference on Global Security, Safety, and Sustainability* (2017), Springer, pp. 12–26.

[38] REED, M. G., SYVERSON, P. F., AND GOLDSCHLAG, D. M. Anonymous connections and onion routing. *IEEE Journal on Selected areas in Communications 16*, 4 (1998), 482–494.

[39] REID, F., AND HARRIGAN, M. An analysis of anonymity in the bitcoin system. In *Security and privacy in social networks*. Springer, 2013, pp. 197–223.

[40] RON, D., AND SHAMIR, A. Quantitative analysis of the full bitcoin transaction graph. In *International Conference on Financial Cryptography and Data Security* (2013), Springer, pp. 6–24.

[41] SAMKEARDOTCOM. Analyzing Bitcoin Network Traffic Using Wireshark, May 2013.

[42] STALLINGS, W. *Cryptography and network security: principles and practices*. Pearson Education India, 2006.

[43] VAN HOUT, M. C., AND BINGHAM, T. 'Silk Road', the virtual drug marketplace: A single case study of user experiences. *International Journal of Drug Policy 24*, 5 (2013), 385–391.

[44] WESTERMANN, B., PANCHENKO, A., AND PIMENIDIS, L. A Kademlia-Based Node Lookup System for Anonymization Networks. In *ISA* (2009), Springer, pp. 179–189.

[45] WINTER, P., AND LINDSKOG, S. How china is blocking Tor. *arXiv preprint arXiv:1204.0447* (2012).