Tampere University

JUHA NURMI

Understanding the Usage of Anonymous Onion Services

Empirical Experiments to Study Criminal Activities in the Tor Network

Tampere University Dissertations 62

Tampere University Dissertations 62

JUHA NURMI

Understanding the Usage of Anonymous Onion Services

Empirical Experiments to Study Criminal Activities in the Tor Network

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty Council of the Faculty of Information Technology and
Communication Sciences of Tampere University,
for public discussion in the Auditorium TB109
of the Tietotalo, Korkeakoulunkatu 1, Tampere,
on 24th of May, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Information Technology and Communication Sciences
Finland

Responsible supervisor and Custos	Associate Professor Billy Brumley Tampere University Finland	
Pre-examiners	Assistant Professor Diana Dolliver University of Alabama United States	Assistant Professor Damon McCoy New York University United States
Opponents	Ph.D. Tobias Pulls Karlstad University Sweden	

Cover design: Roihu Inc.

ISBN 978-952-03-1090-5 (print)
ISBN 978-952-03-1091-2 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-1091-2

PunaMusta Oy – Yliopistopaino

PREFACE

To tell You the truth, I would love to borrow the flying whip of words of Ralph Waldo Emerson, playfulness of scientific text from Carl Sagan, and from the ancient stoics their powerful touch in textual matter. It is generally considered a mark of inexperience to write if a book begins with a few lines about ancient Rome. Nevertheless, as an inexperienced writer, I am delighted to do my mistake by referring to emperor Marcus Aurelius, who ruled Rome from 161 to 180 AD. He has a reputation for being the only wise philosopher king in the history of humans. Essentially in his nightly diary, Meditations, he wrote that

”Nothing has such power to broaden the mind as the ability to investigate systematically and truly all that comes under thy observation in life.”

This is the real preface for a scientific inquiry. To me, during 2010, it meant that I wanted to study what kind of websites there are on Tor. But there was no search engine available for anonymous onion websites in the Tor network. What could I do? I registered ahmia.fi domain and started to create my own search engine for the Tor network.

It was the starting point of my journey to understand how people use these anonymous onion services. I guess we can see the path only afterwards and understand that we were meant to walk it. It is time to summarise this research and thank everyone on this path.

The dissertation was written with the help of Laboratory of Pervasive Computing at Tampere University of Technology. Information security company Cyber Intelligence House greatly supported my work. Furthermore, people behind the Tor Project supported me.

Professor Billy Bob Brumley served as the supervisor of this dissertation. I am deeply grateful to Billy. When the student is ready, the wise supervisor will appear.

Feedback of pre-examiners assistant professor Diana Dolliver and assistant professor Damon McCoy allowed me to take the precision and scientific quality of this dissertation to the next level.

I would like to express my sincere gratitude to Mikko Niemelä, Joona Kannisto, Markku Vajaranta, professor Atte Oksanen, Teemu Kaskela and Jussi Perälä, who all wrote the research papers with me.

Warm thanks to my few friends, for all the good times. Special thanks to my wife who understands me better than anyone else.

Teamwork is essential for success. So I would like to take this opportunity to thank my colleagues with whom I enjoyed working with in my path in different organizations.

Feel free to send your feedback on these thesis and other aspects of this work to juha.nurmi@ahmia.fi. Furthermore, `https://ahmia.fi/` - my very own search engine - enables you to search onion websites.

Wishing You a safe journey inside the Tor anonymity network!

ABSTRACT

Technology is the new host of life, and with each passing year, developments in digitalization make it easier to destroy our understanding of authenticity. A man is more than his distorted shadow on Facebook wall. Another essential shadow dwells under anonymity.

The aim of this thesis is to understand the usage of onion services in the Tor anonymity network. To be more precise the aim is to discover and measure human activities on Tor and on anonymous onion websites. We establish novel facts in the anonymous online environment. We solve technical problems, such as web-crawling and scraping to gather data. We represent new findings on how onion services hide illegal activities. The results are merged with wider range of anonymous onion services usage.

We selected to cast light to the criminal dark side of the Tor network, mainly black marketplaces and hacking. This is a somewhat factitious selection from the wide range of Tor use. However, an archetype villain is found in nearly every story so naturally, for the sake of being interesting, we selected criminal phenomenon to study. To be clear, the Tor network is developed and utilised for legal online privacy and several other essential ways.

The first finding is that as the Tor network becomes more popular also illegal activities become wide spread. Tor and virtual currencies are already transforming drug trade. Anonymous high-class marketplaces are difficult for the law enforcement to interrupt.

On the other hand, now illegal activities are paradoxically more public than ever: everyone can access these onion sites and browse the product listings. The illegal trade is transparent to be followed. For example, by the means of web-crawling and scraping, we produced nearly real-time picture of the trade in Finland following one of the marketplaces on Tor. As a result, statistics shed light on substance consumption habits: the second study estimates that sales totalled over two million euros

between Finnish buyers and sellers.

Due to the network's anonymity and nature of illegal sales, reputation systems have replaced the rule of law: a buyer trusts the seller's reputation because the law is not guaranteeing the delivery. The only available information is the seller's reputation and capacity which were both associated with drug sales as we prove.

Finally, we will identify the limits of online anonymity ranging from technical limitations to operation security dangers. Technology is merely a communication channel and major criminal activities still happen in the physical world. For instance, a drug trade requires that the seller sends the products using post service to the buyer's address. Before that the seller has acquired enormous amounts of illegal drugs. The buyer has to give away his address to the seller who could later be placed under arrest with a list of customers addresses. Furthermore, we show case by case how criminals reveal and leak their critical identity information. The law enforcement agencies are experienced to investigate all of these aspects even if the Tor network itself is secure.

CONTENTS

ABBREVIATIONS

AES	Advanced Encryption Standard
DARPA	Defense Advanced Research Projects Agency
DDoS attack	Distributed Denial-of-Service attack
DHCP	Dynamic Host Configuration Protocol
DHT	Distributed Hash Table
DNS	Domain Name System
DPI	Deep packet inspection
FBI	Federal Bureau of Investigation
GCHQ	British Government Communications Headquarters
GDPR	General Data Protection Regulation
HSDir	Hidden Service Directory
HTTP	Hypertext Transfer Protocol
HTTPS	HTTP Secure
IP address	Internet Protocol address
MITM attack	Man-in-the-middle attack
NAT	Network Address Translators
NSA	National Security Agency
OPSEC	Operational Security
PESTEL	Political, Economic, Social, Technological, Environmental and Legal factor analysis
PGP	Pretty Good Privacy

SMTP	Simple Mail Transfer Protocol
SSH	Secure Shell
TCP	Transmission Control Protocol
TLS	Transport Layer Security
Tor	The Onion Routing
UDP	User Datagram Protocol

LIST OF PUBLICATIONS

This dissertation is a compilation of five publications, of which two (publications I and III) are journal articles, three (II IV and V) appear in conference proceedings.

Finnish Publication Forum[1] ratings indicate the quality of the publication channels. The classification has three levels: 1 is basic; 2 is leading; 3 is top. Other identified publication channels which have not received rating are marked with 0.

The publications are reproduced with kind permissions from the publishers (THL, IEEE, Springer and Elsevier).

I Nurmi, J., Kaskela, T. (2015). "Silkkitie. Päihteiden suomalaista nappikauppaa." *Yhteiskuntapolitiikka*, vol 80, no. 4, pp. 387–394. [1]
Publication Forum level 2.

II Nurmi, J., Kannisto, J., Vajaranta, M. (2016). "Observing Hidden Service Directory Spying with a Private Hidden Service Honeynet." *11th Asia Joint Conference on Information Security (AsiaJCIS)* pp. 55–59. [2]
Publication Forum level 0.

III Nurmi, J., Kaskela, T., Perälä, J., Oksanen, A. (2017). "Seller's reputation and capacity on the illicit drug markets: 11-month study on the Finnish version of the Silk Road." *Drug & Alcohol Dependence* vol 178, pp. 201–207. [3]
Publication Forum level 2.

IV Nurmi, J., Niemelä, M. S. (2017). "Tor De-anonymisation Techniques." *11th International Conference on Network and System Security* pp. 657–671. [4]
Publication Forum level 1.

V Nurmi, J., Niemelä, M. S. (2018). "PESTEL Analysis of Hacktivism Campaign Motivations." *Nordic Conference on Secure IT Systems* pp. 323–335. [5]
Publication Forum level 1.

[1] Finnish Publication Forum, `http://www.tsv.fi/julkaisufoorumi/english.php`

The author of this thesis has the corresponding role in each of the publications and the research behind them. The author was the corresponding author in writing all the publications. He planned the experiments and developed all the technical systems needed to perform the researches. Together, with the help of the co-authors, we prepared and analysed the results.

1 INTRODUCTION

The Internet is growing exponentially: the number of connected devices, the amount of shared data, the amount of saved personal data, and the amount of Internet traffic booms [6]. Thus, it is unsurprising that security problems grow at the same rate. Internet services gather vast amount of personal data. It is the business model of the largest online services to save data for commercial use. Unfortunately, data tends to leak to wrong hands and it is easy to misuse private information.

The Internet is not designed to provide anonymity and privacy: the TCP/IP protocol suite does not protect the privacy of Internet users [7, 8]. After the initial start of the Internet, in 1969, it took two decades before the concrete development of online privacy solutions began [9]. As the Internet grows exponentially the need for online privacy solutions grows as well [10].

Indeed, there are a number of deep motivations for the development of an anonymous online applications. To fully understand further theoretical foundations of online privacy and anonymity the author highlights and recommends ten high quality Ph.D. thesis. The following theses were valuable references to this thesis.

1. *A Pseudonymous Communications Infrastructure for the Internet* by Ian Goldberg. Ph.D. thesis, UC Berkeley, December 2000. [11]

2. *On the Anonymity of Anonymity Systems* by Andrei Serjantov. Ph.D. thesis, University of Cambridge, June 2004. [12]

3. *Better Anonymous Communications* by George Danezis. Ph.D. thesis, University of Cambridge, July 2004. [9]

4. *Anonymity and Privacy in Electronic Services* by Claudia Diaz. Ph.D. thesis, Katholieke Universiteit Leuven, December 2005. [8]

5. *Probabilistic and Information-Theoretic Approaches to Anonymity* by Konstantinos Chatzikokolakis. Ph.D. thesis, École Polytechnique, October 2007. [13]

6. *Covert channel vulnerabilities in anonymity systems* by Steven J. Murdoch. Ph.D. thesis, University of Cambridge, December 2007. [14]
7. *A taxonomy for and analysis of anonymous communications networks* by Douglas Kelly. Ph.D. thesis, Air Force Institute of Technology, March 2009. [15]
8. *Privacy-enhancing Technologies for Private Services* by Karsten Loesing. Ph.D. thesis, University of Bamberg, May 2009. [16]
9. *Improving Security and Performance in Low Latency Anonymity Networks* by Kevin Bauer. Ph.D. thesis, University of Colorado, May 2011. [17]
10. *Privacy Preserving Performance Enhancements for Anonymous Communication Systems* by Rob Jansen. Ph.D. thesis, University of Minnesota, October 2012. [18]

Unfortunately, Internet users need to take constant effort to prevail their privacy. On the other hand, fortunately, there are practical privacy tools available.

In 1996, the design of Onion Routing was published to provide anonymity for low-latency communication systems [19]. In 2004, the final technical implementation of the routing network was ready: Syverson, Dingledine, and Mathewson published their article *Tor: The Second-Generation Onion Router* and the source code of The Onion Routing (Tor) [20]. The Tor network started to provide anonymous TCP/IP connections for everyone.

In addition, Tor enables anonymous Internet services. Onion services are Internet services that are only available through the Tor network and conceal the real IP address and location of the server [16, 17, 20]. Onion services have onion addresses and it is possible to deploy a TCP service on the address. For example, `msydqstlz2kzerdg.onion` is a valid onion address and can be browsed using the Tor Browser.

Information security tools which provide privacy and anonymity create tension: people have a right to be secure in their information and a government's duty is to protect its citizens from harm. For example, in the EU law we have both protections and limitations for the right to privacy [21]. In this dissertation we refer *legal* and *illegal* as it is understood in the EU.

When the use of anonymous technology increases both legal and illegal activities become wide spread [21]. In 2011, the first black marketplace, Silk Road, was founded: it was an onion website providing a platform for selling and buying illegal products, mostly drugs [22]. Silk Road adapted Bitcoin as its payment method and hid the location of the web servers behind the Tor network. It was the first combination of these technologies to enable an enormous scale of online markets for illegal products.

Illegal markets are anonymous but also the trade is transparent. Now everyone can access these onion sites and browse the product listings. As a result, illegal activities are publicly available for research.

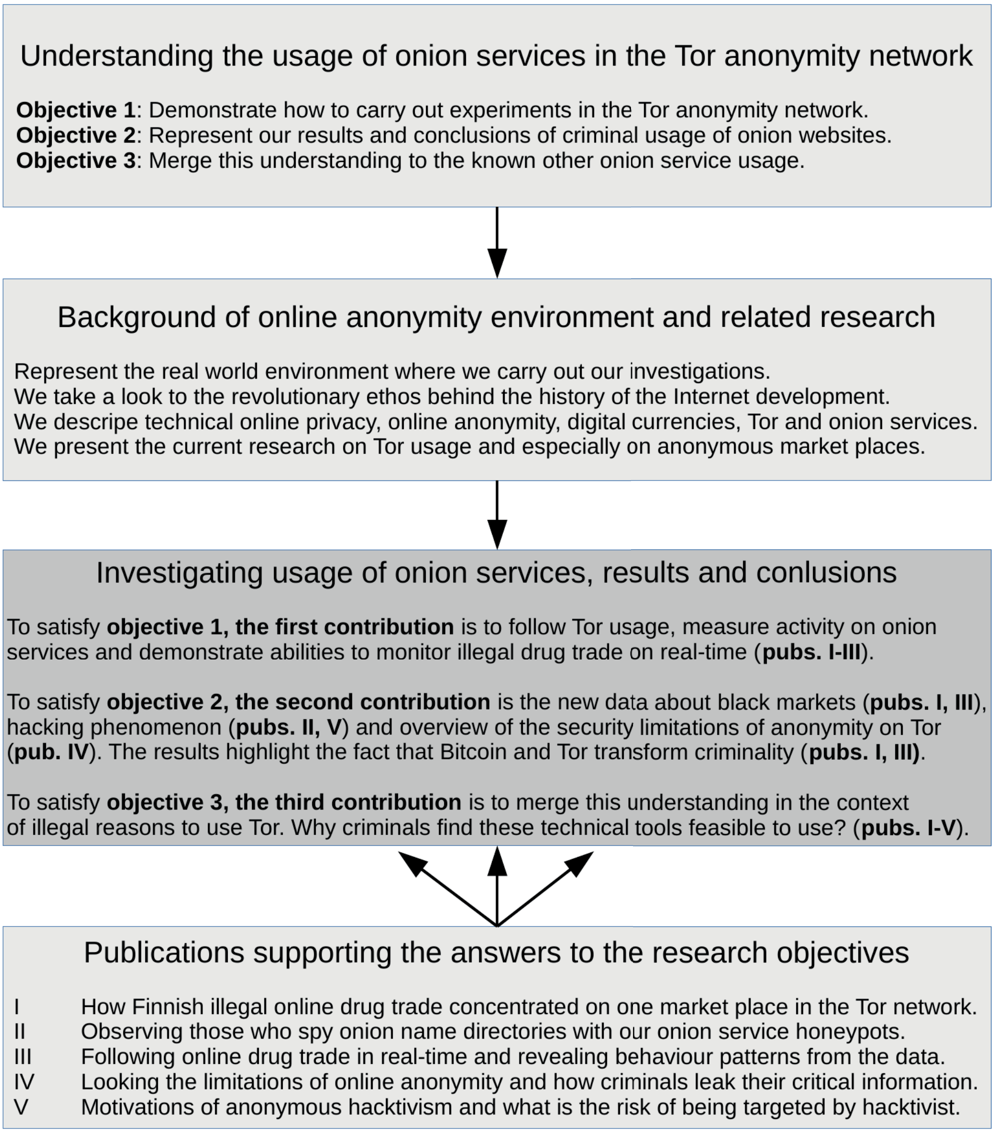
Silk Road was a very ideal place to study how online communication technologies transform crime [22]. Silk Road was the first marketplace that made illegal trade transparent and this trade can be studied with crawling and scraping. As a result, several academic scholars published research regarding different aspects of Silk Road marketplace [22, 23, 24, 25, 26].

In 2013, the marketplace known as Silkkitie was established in the Tor network (Finnish word for Silk Road, also operates under name Valhalla) (publication I). Language and country centric marketplace was unique and one of the first marketplaces on Tor.

Accordingly, in 2014, the author used web-scraping tools to study Finnish illegal drug trade on Silkkitie and described the function of the marketplace. This started the author's investigations of onion services in the Tor network for this thesis.

1.1 Objectives and contributions of the dissertation

This introduction is an overview of the topics of this thesis. A detailed map below is for the reader to quickly visualise the structure of this thesis and logical connections.



Understanding the usage of onion services in the Tor anonymity network

Objective 1: Demonstrate how to carry out experiments in the Tor anonymity network.
Objective 2: Represent our results and conclusions of criminal usage of onion websites.
Objective 3: Merge this understanding to the known other onion service usage.

Background of online anonymity environment and related research

Represent the real world environment where we carry out our investigations.
We take a look to the revolutionary ethos behind the history of the Internet development.
We descripe technical online privacy, online anonymity, digital currencies, Tor and onion services.
We present the current research on Tor usage and especially on anonymous market places.

Investigating usage of onion services, results and conlusions

To satisfy **objective 1, the first contribution** is to follow Tor usage, measure activity on onion services and demonstrate abilities to monitor illegal drug trade on real-time (**pubs. I-III**).

To satisfy **objective 2, the second contribution** is the new data about black markets (**pubs. I, III**), hacking phenomenon (**pubs. II, V**) and overview of the security limitations of anonymity on Tor (**pub. IV**). The results highlight the fact that Bitcoin and Tor transform criminality (**pubs. I, III)**.

To satisfy **objective 3, the third contribution** is to merge this understanding in the context of illegal reasons to use Tor. Why criminals find these technical tools feasible to use? (**pubs. I-V**).

Publications supporting the answers to the research objectives

I	How Finnish illegal online drug trade concentrated on one market place in the Tor network.
II	Observing those who spy onion name directories with our onion service honeypots.
III	Following online drug trade in real-time and revealing behaviour patterns from the data.
IV	Looking the limitations of online anonymity and how criminals leak their critical information.
V	Motivations of anonymous hacktivism and what is the risk of being targeted by hacktivist.

Figure 1.1 A detailed map for the reader to quickly visualise the structure of the thesis and the areas covered: connects the topic, objectives, publications, results and contributions.

Under the topic there are research questions, followed by background and related research. This is actually our research playground where we carry out our research. After that we represent investigations and results to answer our research questions. The answers yield from our five publications. Finally, we summarise conclusions and discuss wider understanding.

Background shows us how to define the anonymous online environment. Also, we look at observations, what could be observed and how anonymous online behavior has been measured previously. After defining our real-world test laboratory, the anonymous online environment in the Tor network, describing its background and referring to the related research this research answers three novel objectives. These individual three objectives of this thesis are summarised below.

Objective I Demonstrate how to carry out experiments in the Tor anonymity network.

Objective II Represent our results and conclusions of criminal usage of onion websites.

Objective III Merge this understanding with the wider scope of illicit onion service usage.

This research makes novel contributions in three complementing levels. First we carry out a series of investigations in the Tor anonymity network to explore it. A concrete outcome of the first contribution is the novel data about black markets, hacking phenomenon and description of the security limitations of online anonymity when using onions services. Finally, we contribute to the body of knowledge on the domain of how people use onion websites.

The main three contributions of this research are as follows:

Contribution I To meet **Objective I**, we contribute to the empirical body of knowledge with our technical methods to follow Tor usage, with measurements on anonymous onion services, and with abilities to monitor anonymous online drug trade in real-time.

Contribution II To answer **Objective II**, our concrete outcome of the first contribution is the novel data about black markets, hacking phenomenon and security limitations on anonymous onions services. We analyse these results and highlight the fact that Bitcoin and Tor are transforming criminal activity, including drug trade.

Contribution III To satisfy **Objective III**, we merge this understanding in the wider context. First, why criminals as well find these technical tools feasible? Then we look at these parts as a whole and we answer what are the common reasons to use Tor and how to detect these behavioral patterns.

The key result of this dissertation is the empirical process to study criminal activities in the Tor network with the concrete research demonstrations.

In publication I we showed how Finnish illegal online drug trade concentrated to Silkkitie marketplace and discussed the rise of drug marketplaces on Tor.

In publications I and III we carried out quantitative research to estimate the drug trade on Silkkitie and explained the structure of the black marketplaces.

In publication III we calculated how seller's reputation and capacity are both associated with drug sales.

In publications II and IV we went through the limitations of privacy in the Tor network.

In publication V we investigated what motivates anonymous hacktivist groups and how they coordinate their attacks and select their targets.

1.2 Structure of this thesis

The thesis is divided into six chapters and you are now reading the first chapter. This is an introduction chapter to the scope of the dissertation. The contents of the other chapters are summarised below.

Chapter 2 represents the background of online anonymity and current state of the research. Namely, how online privacy, online anonymity, digital services, digital currencies and popular practical tools like The Onion Routing (Tor), onion websites and Bitcoin create an environment where humans interact. Theoretical and practical background is described.

Chapter 3 is our investigations of anonymous usage of onion services. We demonstrate how we carry out research in this environment and gather data on various phenomenon. These experiments are carried out in the real world laboratory and we study how real usage of onion services manifests itself.

Chapter 4 Our publications give us a wide range of results to analyse. From the data we can see human behavioral patterns and answer to our research questions. In addition, we show how our contributions merge with wider range of illegal use of online anonymity.

Chapter 5 tells us what we can learn from the research. We show the contributions and provide a synthesis of the key findings.

Finally, in chapter 6, we review the existing state and the future of online anonymity and we represent our final thoughts regarding the matter.

1.3 Research methods and restrictions

1.3.1 Tradition of methodology

After Aristotle (384-322 BC) it took two thousand years before new attempts to formulate the scientific method [27, 28]. In 1543 Nicolaus Copernicus published *On the Revolutions of the Heavenly Spheres* [29]. His work with Galileo Galilei's *Discourses and Mathematical Demonstrations Relating to Two New Sciences* (1638) transformed the views in astronomy and science [29]. Galileo used experiments as a research tool and formulated his treatise. Sir Francis Bacon described his 'Baconian method' in book *New Method* (1620) [29]. Meanwhile, in 1619, René Descartes began writing his *Rules for the Direction of the Mind* to draft guidelines for scientific research [29]. No work of science has drawn more attention than Isaac Newton's *Philosophiae Naturalis Principia Mathematica* (1687) where he formulated the laws of motion and universal gravitation [30]. The scientific Renaissance that lead to the scientific revolution in Europe began, and later transformed itself to the Enlightenment movement [29]. It was seeking the principles of scientific reasoning and knowledge based on experiments [30].

This thesis adapts these principles and knowledge based on scientific experiments. The modern formulation of the scientific method is roughly represented as six phases below. It is based on the article in the Stanford Encyclopedia of Philosophy published by Stanford University and can be regarded as the latest general methodology of science [31].

1. Show a purpose and ask a question.
2. Perform the background research.
3. Construct a hypothesis that can be tested.

4. Test the hypothesis with a repeatable controlled experiment.
5. Analyse the results and draw conclusions.
6. Report the results.

In this thesis this systematic approach is adapted to determine usage of onion services. The structure of our research method follows these six steps in the specified context of Tor and online anonymity.

1. Determine a significant phenomenon enabled by anonymous onion services.
   (a) Is it possible to study this phenomenon?
   (b) How to formulate an exceptional research question?
2. Determine the background of the phenomenon, also outside of the Tor network.
   (a) What is the origin of the phenomenon and what is the role of technology?
   (b) Which economical, technical and societal parts and functions it consists of?
3. Determine an experiment and tracking parameters inside the Tor network.
   (a) Which functions of the phenomenon can be detected in the Tor network?
   (b) How to measure the most important functions of the phenomenon?
4. Create an experiment setup in the real world laboratory.
   (a) What is an ethical way to measure this private human behavior?
   (b) How to solve the technical details of the experiment setup?
5. Analyse the results and fit them to the known background.
   (a) Are the results supporting our hypothesis?
   (b) Does the interpretation of the results fit to known background?
6. Publish the results and conclusions in recognised peer-reviewed forums.

1.3.2 Scope and restrictions

The research environment is the Tor network and this is the first restrictions for this study. We identify behavioral patterns which rely on Tor as underlining technology. Especially, we determine interesting phenomenon enabled by anonymous onion services. The Tor network is the most popular online anonymity system but there are other anonymity systems and environments out there which we are not covering.

The second restriction is that we are interested in a specific category of behavior over others. What is *interesting* human behavior? The author selected to cast light to the *criminal dark side* of the Tor network, mainly black marketplaces and hacking. However, human behavior could be scientifically categorised under endless categories. Yet, there is a certain fascination to categorise it under Jungian archetypes. An archetype villain is found in nearly every story so naturally, for the sake of being interesting, we selected criminal phenomenon to study. We did not research how the Tor network is utilised for legal online privacy and several other essential ways.

The third restriction is that we only identify real detectable phenomenon that happens between real Tor users. There are several publications covering theoretical security flaws, use cases, ideas and extensions on Tor. We research reality as it is available at the moment. For instance, theoretical security threats are out of this scope along with possible undetectable user behavior.

The fourth and final restriction of this study is that we followed ethical guidelines. The author summarised and extended the ethical research principles of the Tor Project [32].

- Appropriateness:
  - The benefits should outweigh the risks.
  - Consider whether the user meant for the data to be private.
  - Use existing public data sets whenever possible.
- Minimization:
  - Only collect data that is safe to publish.
  - Do not collect data that is not needed.
  - Limit the granularity of data, for example, aggregate or add noise.

- Safe experiment:
  - Use a test Tor network whenever possible.
  - It is safest to only attack yourself and your own traffic.
  - Take reasonable security precautions, for example, limit who has access to the data sets or experimental systems.
- Lawfulness:
  - Obey the law, for example, a privacy regulation in EU law has a global reach.

This research obeys these ethical guidelines. Furthermore, in included publications we made sure that we only collected minimal amount of data that is safe to make public, made traffic analysis with our own Tor test network, and took reasonable security precautions.

We comply with the law of Finland, and moreover with the regulation in European Union law on data protection known as the General Data Protection Regulation (GDPR). According to the author's current knowledge, following the GDPR regulation is relative straight forward in the scientific purposes. To help other researchers the author summarised the compliance process to only five questions below.

1. What are the data and does it contain personal data?
2. Where are the data stored?
3. Why are the data collected?
4. Who can access the data?
5. How are the data processed?

If a research is handling personal data of the individuals who live in the European Union then, at least, answer to these questions and check the restrictions from the regulation law text [33]. You have to know if you store personal data in your data set. Even in this case, GDPR is flexible: it is permitted to process the data for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes [33]. Still, you have to document these motives.

2 BACKGROUND OF ONLINE ANONYMITY ENVIRONMENT AND RELATED RESEARCH

Let us now represent the real-world test laboratory, the anonymous online environment in the Tor network. Also, let us look to the background and the related research in this field. An overview map below shows the research environment.

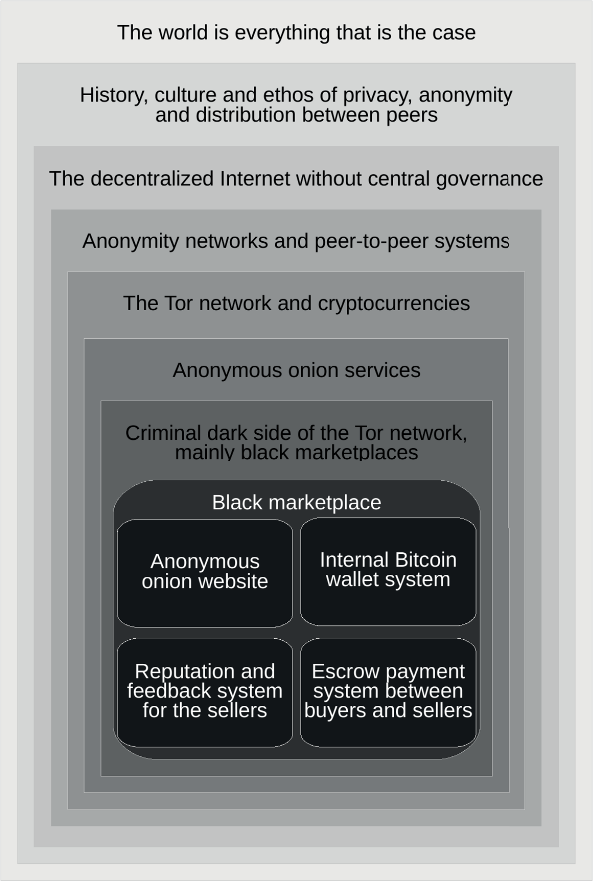

The world is everything that is the case
History, culture and ethos of privacy, anonymity and distribution between peers
The decentralized Internet without central governance
Anonymity networks and peer-to-peer systems
The Tor network and cryptocurrencies
Anonymous onion services
Criminal dark side of the Tor network, mainly black marketplaces
Black marketplace
Anonymous onion website
Internal Bitcoin wallet system
Reputation and feedback system for the sellers
Escrow payment system between buyers and sellers

Figure 2.1 An overview map for the reader to quickly visualise the research environment.

2.1 A brief history of the Internet and its revolutionary ethos

On the beginning of the 21st century we as a species reached the new age of digitalization: the Information Age [34]. This period in human history is characterised by the rapid shift from traditional industry to information technology based economy. Technology is the new host of life: a modern city would collapse to chaos without electricity within days. As the Industrial Revolution marked the onset of the Industrial Age, the onset of the Information Age is the Digital Revolution [34].

Why we have this one global distributed, open, transparent, decentralised and heterogeneous network? Let us briefly examine the history and the revolutionary ethos behind the Information Age. This is an incomplete short overview to show a background for decentralization. As you know, the Internet operates without a central governing body and there is this idea that everyone can attach their devices to the network and share information.

Initial demonstrations in 1969 enabled operational Internet called ARPANET [35]. It was a pioneering network for sharing digital resources among geographically separated computers developed by DARPA [35]. At first, two computers were communicating over Silicon Valley, California [36]. A computer in San Francisco was connected via telephone lines to the second computer 50 kilometers away in Menlo Park at Stanford Research Institute (SRI) [36].

The ARPANET was the first network to implement the protocol suite Internet Protocol (IP) address and Transmission Control Protocol (TCP) [37]. It was developed by the Defense Advanced Research Projects Agency within the U.S. Department of Defense [35]. It was the Cold War era: In the military point of view the network was interesting because it was designed to survive subordinate-network losses [38]. This means that the network can re-coordinate itself in the face of significant interruption by re-computing the routing tables. It took more than a decade before, in 1983, the operational network MILNET (Military Network) and experimental network ARPANET became separate networks [39].

The U.S. Department of Defense was not the only one interested in building a communication network which operates without a central governing body. Distribution of power and communication between peers and decentralization were wider goals. The Cold War was one of the quite many parts of the zeitgeist.

The year 1969 is associated with several other events as well: Apollo 11 made the first manned landing on the Moon, a music festival Woodstock was visited by more than 400,000 young people, UNIX operating system development started, the United States conducted lotteries to call men to military service in the Vietnam War, the hippie counterculture took off in the mainstream and many young student openly protested against the Vietnam War [40].

Later these students contributed to the development of network protocols. We can compare the similarities in the hippie communalism, libertarian politics, and DARPA's research vision to create a fault tolerant distributed communication network.

In 1972, Rand Corporation wrote about the vision that is the networks of computers involves the interconnection of heterogeneous computing resources and this trend may lead to a computer utility where computer resources will be available and marketed the same as electric power and telephone services [37].

In 1980, ARPANET newsletter stated that the ultimate goal is to provide a network that uses several different physical networks and commercial networks to provide interconnectivity between users while still maintaining network transparency [41].

In 1985, John Gilmore drafted the Dynamic Host Configuration Protocol (DHCP) [42]. A few years later he stated that *"The Net interprets censorship as damage and routes around it"* [38]. Accordingly, in 1990, Gilmore, John Perry Barlow and Mitch Kapor founded an international non-profit digital civil liberty group called the Electronic Frontier Foundation (EFF) in San Francisco, California [43]. Barlow, also a former lyricist for the band The Grateful Dead, published *A Declaration of the Independence of Cyberspace* [44, 45]. He demanded independence and sovereignty of the Internet over government control and argued that the enlightened online community is seeking a common good by itself: *"We are creating a world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth."* [45].

These and many other cultures formed the ethos and the roots of the Internet. Today, in the western world everyone can attach their devices to the Internet, publish online content, access information and use or develop peer-to-peer systems with no central coordination between the peers. The peer-to-peer systems include anonymity network Tor and virtual currency Bitcoin.

At least the pathos and the rhetoric of the key persons behind a few popular online services hints to counterculture, libertarian politics and technical search for a fault tolerant distributed communication. In 2008, anonymous creator of Bitcoin, Satoshi Nakamoto, wrote via the MetzDowd cryptography cypherpunk mailing list the following statement regarding his Bitcoin project [46].

> *"Yes, but we can win a major battle in the arms race and gain a new territory of freedom for several years.*
>
> *Governments are good at cutting off the heads of a centrally controlled networks like Napster, but pure P2P networks like Gnutella and Tor seem to be holding their own."*

A few years later many founders of anonymous black marketplaces in the Tor network wrote similar anarchistic statements. These marketplaces use Bitcoin as the payment method.

2.2 Defining online privacy

2.2.1 What is privacy?

Privacy is a useful concept, however, as with similar concepts, we do not have one definition for privacy that explains its meaning [47]. Luckily, however, we can use the word privacy without logical definition of the word [47]. Therefore, here we try ostensive method to scope privacy by pointing out examples.

Each of us have quite a few examples of what privacy is and the operational definition of privacy depends on the situation. Privacy is a part of security and motivates the security research with anonymity systems. This is the field of technical security and privacy.

Aristotle defined the division between two spheres of life: public (polis, city) and private (oikos, home) [15]. This division is fundamental for understanding privacy.

Privacy is also covered in legal frameworks. The constitution of Finland does not explicitly define privacy but implies that "The privacy of a letter, audio call and other confidential message is absolute" (author's translation). However, there are limitations to this absolute confidentially in the other sections of the law. The Fourth Amendment to the United States Constitution describes similar right to privacy stating that "The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated". Again, however, this law adds limitations to this right.

The Michigan Supreme Court Justice defined privacy as "the right to be let alone" [15]. General Data Protection Regulation in European Union law sees it very similarly and extends it by adding "the right to be forgotten" [33].

At an international level privacy is defined in the 1948 United Nations Universal Declaration of Human Rights in the article 12, and the article 17 of the International Covenant on Civil and Political Rights recognise privacy as a basic human right [15].

In scientific texts there are several descriptions of privacy depending the field of science. For instance, in the field of security research, Ian Avrum Goldberg represents in his dissertation a definition of privacy [11].

> *"Privacy refers to the ability of the individual to control the distribution of information about himself. Note that this does not necessarily mean that your personal information never gets revealed to anyone; rather, a system that respects your privacy will allow you to select what information about you is revealed, and to whom. This personal information may be any of a large number of things, including your reading habits, your shopping habits, your nationality, your email or IP address, your physical address, or of course, your identity."*

The author would like to append that privacy is a part of security and liberty. Furthermore, if something tries to trick or pressure people to give up their privacy this cannot be considered liberty and respect of privacy.

Unfortunately, currently, many popular applications will suddenly show a message like *To continue to use our online services, you must now click OK to accept the new Terms of Service*. In several cases these new terms are painfully difficult to read and extremely long. Whatever this is, it is not respect of privacy.

2.2.2 Internet privacy

In the online context privacy is an enormous scale technical problem. By default, transport and routing protocols are not providing privacy: the TCP/IP protocol suite is not designed to protect the privacy of Internet users [7, 8]. As a result, a service that the user is accessing discovers the IP address of the user and this IP address can be connected to the certain physical place, for instance, to the home address of the user. Secondly, an attacker who spies the connection will learn what the user is doing online.

Let us examine a basic situation where a user access a forum website over Hypertext Transfer Protocol (HTTP). Accordingly, the website discovers the IP address of the user and commonly this information is saved to server logs. Later with this information it is possible to find the real physical home address of the user. Furthermore, because HTTP is not encrypted the content of the communication can be read in the network between the user and the service. For example, it is possible to discover both the writer and the content of the message to this forum website by monitoring the connection.

Privacy of the content of the messages can be secured with encryption. In this case, instead of sending the plain text version of the message, the text is re-written using an encryption algorithm and it can be decrypted and read only by the receiver. For example, the forum website could offer HTTP Secure (HTTPS) which is encrypted with Transport Layer Security (TLS) between the user and the website [15]. However, still, the website knows the IP address of the user. Secondly, if an attacker is monitoring the network connection then the attacker sees that the user is accessing the site. Although, the attacker cannot read the messages between the user and the site, the attacker knows which website the user is browsing [15].

It is well known that online activities are monitored by several attackers, such as cyber criminals, surveillance agencies and Internet service providers [11, 17, 18]. Also, multiple companies offer *free* online services to gather private data and sell services based on the data to their paying customers [16].

In response to privacy threats, several online applications try to provide security measures against online surveillance, add encryption layers, and increase the confidentiality of the communication. The field of security research supports this effort [8, 9, 11, 12, 13, 14, 15, 16, 17, 18].

Repeatedly web services accidentally leak their logs which include personal data. This is common although services are demanded to gather minimal amount of personal data and make sure that the data cannot be accessed without permission. In the law of European Union personal data is globally protected under GDPR and this protection covers information that could be utilised to identify any aspect of an individual's life [33]. Commonly IP address itself is personal data. Other examples include name, online user name, email address, home address and phone number.

Next we will look at how to have Internet privacy and why one has to use energy and special solutions. In this dissertation we recognise the background fact that online activities tend to produce data logs and these data logs tend to leak [13, 48]. Unfortunately, this irreversible process yields from the second law of thermodynamics that states that the entropy of an isolated system never decreases [13, 48]. As a result, Internet privacy and information security is a constant process to use energy to minimise sprawling of data [13].

2.3 Characterization of online anonymity

Anonymity and pseudonymity are two forms of privacy of identity frequently both referred to as simply anonymity [11]. Under anonymity the user controls who can learn his identity [11]. Anonymity system offers this functionality and thus an adversary cannot discover the user's identity against his wishes [11]. Pseudonymity means that the user maintains anonymous identity, for instance, a nickname. Other users can interact with this nickname that is persistent. Without pseudonymity there is no persistent identifier and, as a result, anonymous users cannot be recognised from each other.

2.3.1 Evolution of anonymity systems

2.3.1.1 Type 0: remailing system

In 1993, in Finland, Johan Helsingius (also known as under pseudonym Julf) began to operate an email service, anon.penet.fi, providing anonymous email accounts [9]. The email relay kept a table of correspondences between real email addresses and pseudonymous addresses [9]. As a result, an email to a pseudonym would be forwarded to the real user email address and an email from a pseudonym was forwarded to the recipient without identifying information [9].

The anon.penet.fi remailer was one of the first examples of online anonymity systems.

Technically, this service provides anonymity against the users receiving or sending emails to a pseudonym [9]. They are not able to find the real email address of the pseudonym [9]. However, it is trivial for a passive attacker who can eavesdrop on Internet traffic, or the service itself to look which pseudonym email address and real email address are connected [16].

The user was assigned a pseudonym at anon.penet.fi and the remailer maintained a secret identity table matching up the real email address with the pseudonym. It provided weak anonymity by stripping identifying headers from outbound remailed messages [11]. This system is referred as type 0 anonymity service [11, 16, 17]. The service itself can reveal the identities. Eventually, this limitation was the reason why Helsingius closed his remailer service [9, 11, 16].

The anon.penet.fi remailer was the most popular remailer and the most famous one [11]. The court case of the service was noted by anonymity research community [9, 11, 16].

The judge ruled the privacy rights in the constitution of Finland provide no means not to serve as a witness [9]. If the service owner, in this case Helsingius, is called as a witness to reveal the connection between a certain anonymous address and corresponding real email address then he must reveal this information [9]. As a result, Helsingius closed his service in 1996 because he could no longer guarantee the anonymity of its users [9].

While writing this chapter the author noticed that Johan "Julf" Helsingius joined the conversation on Internet Relay Chat (IRC) on #effi channel. The author asked if Mr. Helsingius could provide his thoughts on anonymity to this thesis. Julf kindly replied.

> *"These days there are enormous amounts of information about us being automatically collected, combined and processed. I ofter hear 'But I have nothing to hide', but I am pretty sure all of us have things we wouldn't like other people – relatives, friends, neighbors, colleagues or employers - to know. What worries me even more is that a lot of that data is being stored and handled rather haphazardly. It might very easily end up in the wrong hands. As an almost overly dramatic example of this, nobody in 1935 could foresee that the excellent population register of the city of Amsterdam, that held not only home addresses, but also the religious affiliation of every inhabitant, would lead to the Jewish population ending up in concentration camps."*

2.3.1.2 Type I: the Cypherpunk remailer

The simple design of type 0 pseudonymous remailers resulted in two problems: the first is vulnerability to traffic analysis and the second is that the connection between pseudonyms and real addresses is kept on the remailing server [9, 16].

Problem 1. Type 0 remailer cannot protect anonymity against traffic analysis between incoming and outgoing messages. An adversary with the ability to monitor traffic can efficiently correlate the messages because outgoing messages have similar sizes as incoming messages and both messages travel within a short time frame [9, 16].

Problem 2. Type 0 remailer can be hacked or forced to disclose the connection between pseudonyms and real addresses for legal reasons. Whoever operates the most sensitive part of the system can uncover pseudonymity of all recipients in the system. Users need to trust the remailer operator in keeping this table secret and protecting it against hacker attacks. [9, 16]

To solve these two problems type I remailing system, the Cyperpunk remailer, was developed. The security of the design of Cyperpunk remailer is represented by Goldberg (2000) [11], Danezis (2004) [9], Loesing (2009) [16] and Bauer (2011) [17].

To address the problem 1, to prevent that the adversary trivially matches up incoming and outgoing messages, the Cyperpunk remailer accepts encrypted email and after decrypting it the remailer sends the resulting message. To address problem 2, type I remailer does not support pseudonyms.

Multiple remailers can be chained together combined with encryption to add security. The sender encrypts the message with the public keys of each remailer in the chain. The original message and destination address are encrypted with the public key of the last remailer in the chain. Therefore only the last remailer will see the destination address. The multilayer encrypted message is sent to the first remailer which decrypts the address of the next remailer and forwards the message.

For example, if we chain three remailers together the first remailer sees the original address but not the address of the final destination; the second remailer sees only the address of the first remailer and third remailer; finally the third remailer sees the address of second remailer and final destination address. Now a single server cannot learn which original email address is in contact with the destination address.

An attacker must compromise each the remailers in the chain to trace messages.

Only one non-compromised remailer is needed in the chain and the attacker cannot follow messages. Moreover, type I remailers randomly reorder outgoing messages to address the problem 1 with the traffic analysis between incoming and outgoing messages.

2.3.1.3 Type II and III: Mixmaster remailer

In the mid-nineties, next development of the remailer technology is so called type II and III Mixmaster remailer. It extends the techniques utilised in a type I remailer to provide enhanced protection against traffic analysis [9, 11, 17].

Mixmaster remailer sends messages in fixed-size packets through a Chaumian mix network [17]. It does not support reply messages [17].

The type II Mixmaster remailer extends type I remailer following four methods [11]:

1. A Type II remailer is always chaining multiple remailers and adding encryption between chains.
2. Messages have constant-lengths to prevent matching incoming and outgoing messages by size with passive traffic monitoring.
3. A type II remailer is a stateful remailer and does not resend messages to prevent replay attacks. Otherwise an adversary could intercept a message and resend it multiple times to correlate where these emails are heading to.
4. A type II remailer is reordering messages to add security against eavesdropping.

After development of type II remailers the next generation system is called type III Mixminion anonymous remailer. It offers anonymous reply addresses, TLS encrypted Simple Mail Transfer Protocol (SMTP) with unique encryption keys for each message, exit policies to enable mixminion node operators to filter abusive email messages [17]. Dummy noise traffic is added to mitigate traffic analysis vulnerabilities [17].

Email remailers cannot solve the problem that TCP/IP stack does not provide privacy features by default. However, most of these protection methods will be seen again in the next section where we show how Tor software implements full support for anonymous TCP/IP.

2.4 Tor enabling strong practical anonymity

In 1996, the design of Onion Routing was published: it is the heir of mix remailing systems and implements a general low-latency communication system [19]. It took a few years before the final technical implementation of the routing network was ready. In 2004, Syverson, Dingledine, and Mathewson published their article *Tor: The Second-Generation Onion Router* and the source code of The Onion Routing (Tor) [20]. The Tor network started to provide anonymous TCP/IP connections for everyone.

In the previous section there were the design principles of anonymous remailers. Each remailer in the chain decrypts, delays, and re-orders messages before sending them to the next remailer [20]. These high-latency networks resist strong global adversaries but have long latency and because of this lag this design is not suitable for interactive tasks like web browsing or Secure Shell (SSH) connections [8, 20].

Tor uses these protection methods but extends the system to support TCP traffic, not only sending emails [8, 49]. As such, Tor is a low-latency anonymity network [20]. As a result, it is possible to use multiple Internet applications anonymously behind Tor. Everything that communicates over TCP; User Datagram Protocol (UDP) protocol is not supported [20].

Tor hides the original IP address of the user. At the moment, Tor is the most widely used anonymity tool and can be used for both legal and illegal purposes [21]. It is utilised by ordinary citizens concerned regarding their privacy, corporations who want to hide information from their competitors, and law enforcement agencies who need to carry out their work without being noticed. Furthermore, human right activists and journalist use Tor for safe communication.

To provide online privacy and anonymity Tor combines the following main techniques [20].

1. The user is not directly connected to the destination. Instead, the user connects to the Tor network and through the network his connection reaches the destination.

2. The final destination does not see the original IP address of the user; instead, the destination sees a Tor relay's IP address.

3. There are three relays chained and this is called a Tor circuit: the first relay

knows the IP address of the user, but does not learn the destination. The second relay only sees incoming traffic from the first relay to the third relay. The third relay knows the destination but does not know the original IP address of the user.

4. The user encrypts the traffic with each of the three relays. There are separate encryption layers and every relay decrypts one layer off. It is analogous to layers of an onion and this is the reason why the protocol is called onion routing.
5. The Tor network is open and free for everyone: this attracts a lot of users and relays. As a result, it is difficult to introduce global traffic and timing analysis against the Tor network.

The three relay circuit includes end-to-end encryption (the 128-bit Advanced Encryption Standard (AES) cipher in counter mode) and checksums for integrity checking: the relay cell payload and header are encrypted and decrypted as the relay cell moves along the circuit [20].

Moreover, Tor is a lively development project and new security features have been added over the years. To prevent traffic and timing correlation attacks, a padding mechanism was introduced to the latest version of Tor [14, 50].

Still, a low-latency anonymity network is vulnerable to traffic correlation attacks by an attacker that eavesdrops on both ends. This limitation cannot be completely removed from a low-latency anonymity network [12]. However, the Tor network is huge and an adversary must perform global enormous scale monitoring [14, 20, 50].

Tor software is developed by the Tor Project, Inc. and it is a Massachusetts-based 501 research-education nonprofit organization, see `https://www.torproject.org/`. Tor is free software and the Tor network is the network of computers that communicate using Tor software [20].

In the Tor network there are different roles for these computers. Voluntary people have installed Tor software in a routing mode and permit their computers to receive traffic on the Tor network and pass it along. These computer servers are commonly called Tor routers, relays or nodes. In 2018, according to `https://metrics.torproject.org/`, there were around 7000 Tor relays all over the Earth. Normally, without modifications, Tor is a client software that connects to the Tor network and opens socks proxy. Tor users are not transferring traffic by default.

In addition, because some countries try to block connections to the Tor network by filtering IP addresses of all publicly known Tor relays there are Tor relays which are not publicly listed as part of the Tor network. These Tor relays are called Tor bridge relays. Bridges are not listed in the main Tor directory so nobody has a complete list of the IP addresses of these relays. As a result, it is impossible to completely block Tor connections by filtering connections to known Tor relay IP addresses. In 2018, according to `https://metrics.torproject.org/`, there were around two thousand Tor bridges.

Tor bridges provide several sophisticated pluggable transport protocols to obfuscate the connection [51]. In this case, the traffic between the client and the bridge is not identifiable as a Tor connection [51]. It is laborious and complex to build a censor system like this: the system needs to apply deep packet inspection (DPI) data processing to inspects in detail all the traffic to detect these connections [51]. Still it might be difficult to decide if the connection is a Tor connection.

In 2018, according to `https://metrics.torproject.org/`, every moment there were over two million users connected to the Tor network (see Figure 2.2 below).

How safe is Tor and what are the limitations? We will answer to these questions in the upcoming chapter and in our publication IV.

Tor enables TCP-based applications to obtain online anonymity without modification: it offers the standard and near-ubiquitous socks proxy interface [20]. We list here multiple technical examples how one can use Tor.

Tor enables anonymous web browsing. Although it is possible to setup a web browser to route traffic through Tor's socks proxy this is not recommended because complex modern web browser can leak identifying information so many ways. To offer a safe anonymous web browser the Tor Project develops Tor Browser, their main privacy-aware application [52]. The latest information and the browser is available on `https://www.torproject.org/projects/torbrowser.html.en`. This Tor Browser is as easy to use as a common web browser: it is a modified Mozilla Firefox Extended Support Release (ESR) with best-practice default settings and extensions, such as NoScript and HTTPS Everywhere. Most importantly Tor Browser starts Tor background process and routes all web browsing traffic through the Tor network. In addition, this browser removes all possible fingerprinting methods, including faking the information about operating system and hardware. Tor Browser removes local privacy-sensitive data, such as the browsing history, cache and cookies.

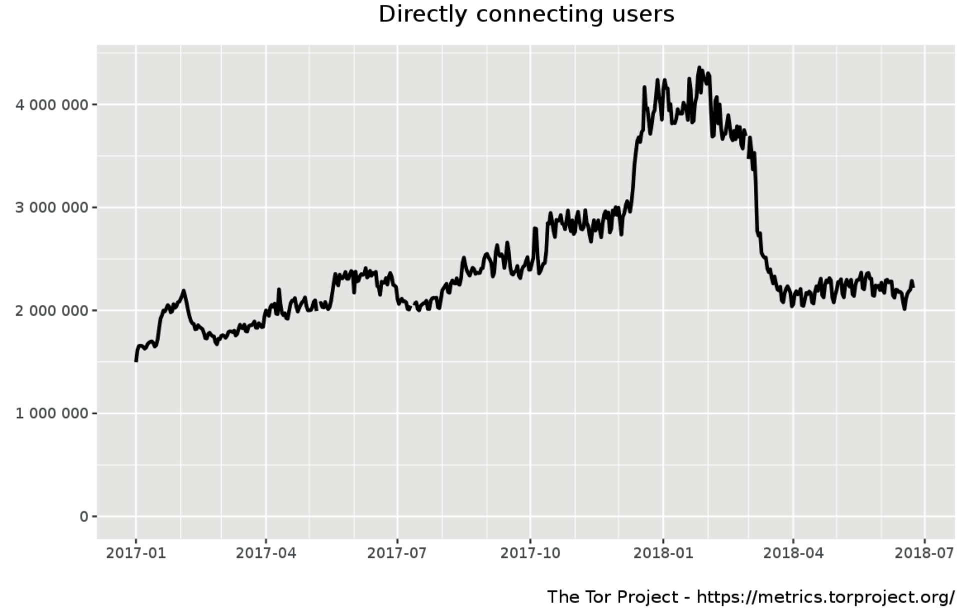


Figure 2.2 In 2018 every moment there were more than two million Tor users connected to the Tor network. The number of simultaneously connected users even peaked over four million.

Tor enables anonymous web publishing. Similar to anonymous web browsing, deploying an anonymous website is simple. Tor offers a censorship-resistant and distributed platform that provides easy-to-implement anonymity for server applications, including websites. This means that a web server is only available through the Tor network and hides its real IP address and its physical location. These web sites can be accessed using Tor browser. In this case, both the user and the website have anonymity.

Tor enables anonymous online communication. Tor offers socks proxy for a TCP application. As a result, communication applications like SSH and IRC can connect through the Tor network. In practice, the Tor Project offers torsocks wrapper application which forces DNS requests and TCP traffic of an application to use this socks proxy. It rejects other traffic than TCP of the application, such as possible UDP. For example, to use SSH or IRC anonymously one can launch the application with torsocks on commandline:

SSH connection: torsocks ssh user@example.com

IRC with irssi application: torsocks irssi

All in all, Tor is a general software to obtain privacy, it enables a TCP-based application to route traffic anonymously, and the Tor network has a huge user space. The enormous number of various users creates a significant anonymity set against Internet surveillance.

2.5 Onion websites on Tor

Tor hidden services are Internet services that are only available through the Tor network and conceal the real IP address and location of the server [16, 17, 20]. In this thesis we call these services onion services because this is the current official naming practice of the Tor Project. Onion services have onion addresses and it is possible to deploy a TCP service on the address.

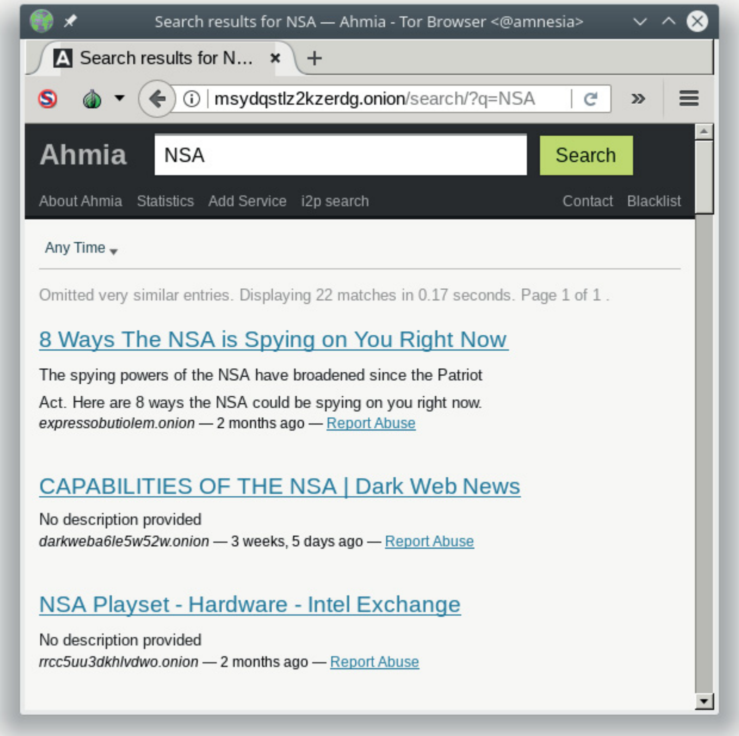
Location-hidden services use a virtual top level domain called x.onion and the first part x is a 16 character hash of the public key of the onion service. For example, `msydqstlz2kzerdg.onion` is one valid onion address. This onion service is a website (actually identical to `ahmia.fi`) and can be browsed using the Tor Browser (see Figure 2.3).

Onion services do not use the regular Domain Name System (DNS). Instead, Tor keeps a decentralised distributed hash table (DHT) between onion relays [20]. This is a key-value list of onion addresses and introduction points of the onion services [20]. Introduction points are onion relays which an onion service has selected as contact points and it keeps three hop Tor relay circuit open to these points.

When a Tor user connects to an onion service the user first queries the introduction points from the DHT [20]. Then the user connects to one of the introduction points and informs the onion service what is the rendezvous point [20]. A rendezvous point means a relay which a user has selected as a meeting point between the user and an onion service. After the onion service receives the information regarding the rendezvous point through introduction point it connects to the rendezvous point [20]. Now they both have three hop Tor relay circuit to this point and they can communicate anonymously [20].

This is how onion service protocol enables Tor users to connect to onion addresses. Let us summarise these steps [20].

1. An onion service randomly selects three relays, builds three hop circuits to them, tells them its public key, and these relays act as introduction points.



Omitted very similar entries. Displaying 22 matches in 0.17 seconds. Page 1 of 1 .

8 Ways The NSA is Spying on You Right Now

The spying powers of the NSA have broadened since the Patriot Act. Here are 8 ways the NSA could be spying on you right now.

expressobutiolem.onion — 2 months ago — Report Abuse

CAPABILITIES OF THE NSA | Dark Web News

No description provided

darkweba6le5w52w.onion — 3 weeks, 5 days ago — Report Abuse

NSA Playset - Hardware - Intel Exchange

No description provided

rrcc5uu3dkhlvdwo.onion — 2 months ago — Report Abuse

Figure 2.3 The Tor Browser is a special web browser which routes all its traffic through the Tor network. The address `msydqstlz2kzerdg.onion` is an onion address and a fraction of these addresses publish web content. Websites inside the Tor network are location-hidden. These anonymous onion websites can be viewed, crawled and indexed to provide a search engine.

.

2. The onion service creates an onion service descriptor. This descriptor contains the public key and list of the introduction points, and it is signed with the private key of the onion service. The onion service uploads this descriptor to the DHT.

3. A Tor client can request the descriptor by querying the x.onion address of the service. The client downloads the descriptor from the DHT and verifies that the descriptor is signed by the onion service x.onion. Now the client knows the list of introduction points and the public key of the onion service.

4. The client randomly selects one relay, builds three hop circuits to it and asks it to act as a rendezvous point by telling it a one-time secret.

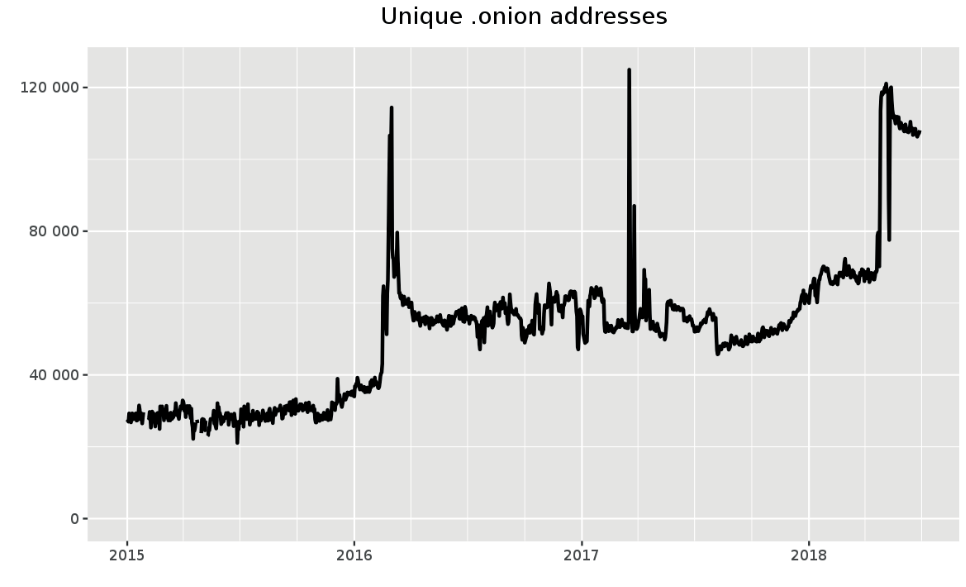
5. Next the client assembles an introduce message which is encrypted with the public key of the onion service. The message consist of the rendezvous point information and the one-time secret. The client sends the message to one of the introduction points. This introduction point delivers the introduce message to the onion service.
6. The onion service decrypts the introduce message and knows now the rendezvous point and the one-time secret. It connects through a circuit to the rendezvous point and sends rendezvous message with the one-time secret.
7. Finally, both the client and the onion service have their circuits to the rendezvous point. The user and the onion service communicate end-to-end encrypted messages between each other. There are five relays between the user and the onion service passing the messages.

Onion addresses are not using DNS and rather than revealing the real IP addresses onion services can only be accessed through the Tor network. Because of this it is impossible to censor these services. Moreover, because every connection from the onion service is behind a three hop circuit it is extremely difficult to locate the actual server. It is even possible to host onion services behind firewalls or network address translators (NAT) without modifications. If the owner of the service does not reveal oneself the operator is able to maintain censorship resistant and location-hidden Internet service, for instance, a website that is publishing news anonymously.

Onion services and users encrypt their communication from end-to-end so HTTPS is not necessary for websites. Furthermore, the onion address itself is a hash checksum from the public key of the onion service. As a result the encryption and verification between the user and the onion service prevents Man-in-the-middle (MITM) attack.

According to the Tor Project's measurements `https://metrics.torproject.org/` there are over 100 000 active onion addresses (see statistics in the Figure 2.4 below). However, only a fraction of these onion services are web services. Search engine `ahmia.fi` estimates that there are around 15 000 active onion websites in a given moment and multiple of these websites have only a lifespan of weeks.

Onion services enable anonymous web publishing. As a result, the published content is diverse. Undoubtedly, some onion websites share pictures of child abuse or operate as illegal marketplaces for drugs. These few services are controversial



The Tor Project - https://metrics.torproject.org/

Figure 2.4 The Tor Project is measuring the number of onion addresses. These addresses can provide Internet services over TCP, including web sites, BitTorrent trackers, IRC servers and other chat protocols.

and frequently pointed out by critics of Tor and anonymity. Still, onion websites are sharing legal content and multiple of them are even devoted to human rights, journalism and publishing content that is censored by oppressive governments.

The Tor Browser is easy to use and, similarly, deploying an anonymous onion website is simple. Tor offers a censorship-resistant and distributed platform that provides easy-to-implement anonymity to web users, onion websites, and other Internet services.

2.6 Decentralised digital currency revolution

The illegal marketplaces in the Tor network need more than leverage of onion services to operate. A payment method is needed and normal bank transfers or credit card payments are not suitable for illegal trade. Pseudonymous unregulated electronic money arrived in 2009 and two years after that the first marketplaces were created. These marketplaces combined the anonymous website publishing capabilities of the Tor network and the unregulated new currency as their payment system.

Bitcoin is the first decentralised digital currency. Anonymous entity known as Satoshi Nakamoto created the peer-to-peer software in 2009 [53]. It forms a blockchain based currency that works without central organizations [53]. Bitcoin is a digital asset created without a central bank [53].

Bitcoins can be transferred on the peer-to-peer Bitcoin network directly between the users of the system [53]. These transactions are written to the shared Bitcoin blockchain [53]. With hash checksums and proof-of-work cryptography the network verifies these transactions and it is not feasible to spoof transactions [53]. The proof-of-work solves the problem of determining representation in majority decision making: as long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network an attacker cannot counterfeit transactions [53].

The price of bitcoins is determined by traders who exchange bitcoins for other currencies [54]. The Bitcoin network creates new bitcoins to nodes verifying transactions and eventually the network reaches the hard limit of 21 million bitcoins [53]. As a result, the price is based on the continuous demand to buy and sell this limited number of bitcoins.

Bitcoin is pseudonymous and the identifier is the user's bitcoin address. Although owners of the bitcoin addresses are not explicitly combined together every transaction between users is permanently saved to the public blockchain [54]. As a result, the anonymity is very limited: when the user exchanges his bitcoins to traditional currencies the trading market will commonly require personal details and a bank account. In 2018, one study estimated that 25 percent of users (approximately 24 million) and 44 percent of transactions are associated with illegal drug trade [55].

Even though the Bitcoin system does not offer technical anonymity, however, bitcoins are widely utilised in illegal trade in the Tor network. Furthermore, the Bitcoin system is the first of so called cryptocurrencies but not the only one: in 2018 popular trading places recognised several popular electronic peer-to-peer currencies. You can check out the current market prices of the popular currencies here `https://coinmarketcap.com/coins/`. Obviously there is an enormous support for legal use of blockchain based electronic payment systems, however, criminals have adapted Bitcoin and these other cryptocurrencies as well.

2.7 Earlier research on Tor usage

Let us take a look at the related research on Tor usage. Especially, how this research sheds lights on how people use Tor and onion services. Moreover, we created a categorization of the research methods: we can see eight technical data collection methodologies and fit the previous research to our categorization.

1. Tor exit relay traffic monitoring [56, 57, 58]. It is important to note that the exit relay can observe the decrypted payload before it sends it to the final destination. Moreover, obviously, the exit relay learns what is the final destination of the Tor user. If the user does not use encryption, for instance, HTTPS, then the exit relay could (illegally and unethically) wiretap the actual content of the user's communication.

2. Exploiting unfixed Tor protocol weaknesses and thus revealing usage information [50, 59].

3. Passive traffic and timing correlation attacks where an attacker correlates connections between endpoints and reveals Tor user's identity [60, 61, 62, 63, 64, 65, 66].

4. Active traffic and timing correlation attacks where an attacker injects traffic or delays traffic passing through the Tor network and this way shapes the traffic pattern [67, 68, 69, 70].

5. Port scanning onion services to probe for open ports and services [59, 71, 72]. This method reveals what Internet services onion addresses are providing.

6. Official Tor metrics provided by the Tor Project. The Tor Project gathers non-sensitive data for metrics [73]. The goal is to provide data for monitoring, research, and developing the Tor network. Also, the data can be utilised to detect censorship events or attacks against the Tor network. Technically, Tor relays estimate the number of users per country, number of different types of relays, number of Tor applications downloads, censorship events, amount of Tor traffic, performance of the network, and number of active onion addresses.

7. Research honeypots in the Tor network reveal, for instance, how attackers are spying information: honeypots can send traffic through exit nodes containing decoy credentials [74], or onion service honeypots can be installed to detect misuse of the Tor network as our publication II demonstrates.

8. Gathering data by crawling onion websites as our publications I and III show along with earlier research [22, 23, 26, 71, 75, 76, 77, 78, 79].

In addition, non-technical methods such as surveys could be utilised as well, however, in this thesis we only cover technical methodology. Also, although each of these technical methods could be utilised to determine how people use the Tor network many of them are illegal and unethical. For instance, in the EU, Tor exit node operators are generally protected from liability for the content passing through the exit node, and, furthermore, it is illegal to wiretap and spy on Tor network traffic [21]. Also, in the EU, national laws protect Tor exit relay traffic [21]. Table 1 shows what these eight technical methods could reveal regarding the usage of the Tor network.

In the EU, anonymity is an integral part of established human rights, and therefore indiscriminate monitoring of is not allowed: this means that the Tor network is subject to the same requirements as the interception of other Internet traffic and using Tor is not labeled as a sign of criminality [21].

Technical methodology	**What the resulting data could reveal regarding the usage of the Tor network?**
1. Monitor exit relay	Analysis of Tor exit relay traffic can show which Internet services Tor users access, for example, which websites are browsed anonymously.
2. Exploit protocol weakness	Unfixed protocol weaknesses have leaked information of how people use Tor, for instance, which onion services are the most popular.
3. Passive traffic correlation	An attacker correlates connections between endpoints and reveals, for instance, what the Tor users from a certain country are accessing.
4. Active traffic correlation	An attacker shapes the traffic pattern and, for example, shows who is accessing certain onion services.
5. Port scan onion services	Results from port scanning onion services reveal wide range of Internet services in the Tor network and we know that only a minor fraction of the onion services are websites.
6. Official Tor metrics	From the public statistics we can learn number of users per country, number of relays, how much traffic the Tor network handles, how quick and reliable it is, number of onion services, and how many times Tor browser is downloaded and updated.
7. Research honeypots	Honeypots detect how attackers try to exploit services and spy Tor users.
8. Crawl onion websites	Finally, using leverages of web crawling and scraping techniques, research shows what kind of web content there is available on onion websites.

Table 1. Earlier research sheds lights on how people use Tor and onion services.

We created a categorization of the research methods: we can categorise the research according to technical data collection methodologies.

In this thesis we investigate empirically how the Tor network is used with research honeypots (method 7) and by web crawling and scraping onion websites (method 8) and fit our results to earlier research.

Before explaining our investigation let us take a look at inspiration that earlier research on illegal marketplaces on Tor gave us. Paradoxically illegal activities are more transparent in the Tor network because anyone can access these onion sites and look at the product listings. As a result, the illegal trade can be studied by the means of web-crawling and scraping (method 8).

2.8 Earlier research on marketplaces operated on Tor

In 2011, the first black marketplace, Silk Road, was founded: it was an onion website providing a platform for selling and buying illegal products, mostly drugs [22]. Silk Road adapted Bitcoin as its payment system and hid the location of the web servers behind the Tor network. This was the first combination of the technologies and enabled an enormous scale of online markets for illegal products.

Silk Road captured worldwide media attention and alarm from law enforcement agencies already not making progress with the drug prohibition policies [24]. According to the U.S. Congress Silk Road is *the most brazen attempt to peddle drugs online we have ever seen* [24].

In Silk Road, vendors and buyers have their Bitcoin wallets for payments [23]. It is mandated by the site to use the *escrow system* [23]. This system takes a commission fee to the site and locks the payments between a buyer and a vendor [23]. When the buyer receives the product the buyer frees the payment to the vendor. In most cases the buyer receives desirable product and the Bitcoin payment is transferred to the seller. This is because the reputation of the seller is very crucial and the research shows that the sellers fulfill their promises [79]. If the buyer is not happy with the product the buyer can make a dispute claim and the site creates a resolution between the seller and the buyer [23]. The buyer gives a public feedback to the seller, and other buyers see the reputation of the seller.

Silk Road was a very ideal place to study how online communication technologies transform crime [22]. Silk Road was the first marketplace that made illegal trade transparent and this trade can be studied with crawling and scraping. As a result, several academic scholars published research regarding different aspects of Silk Road marketplace [22, 23, 24, 25, 26].

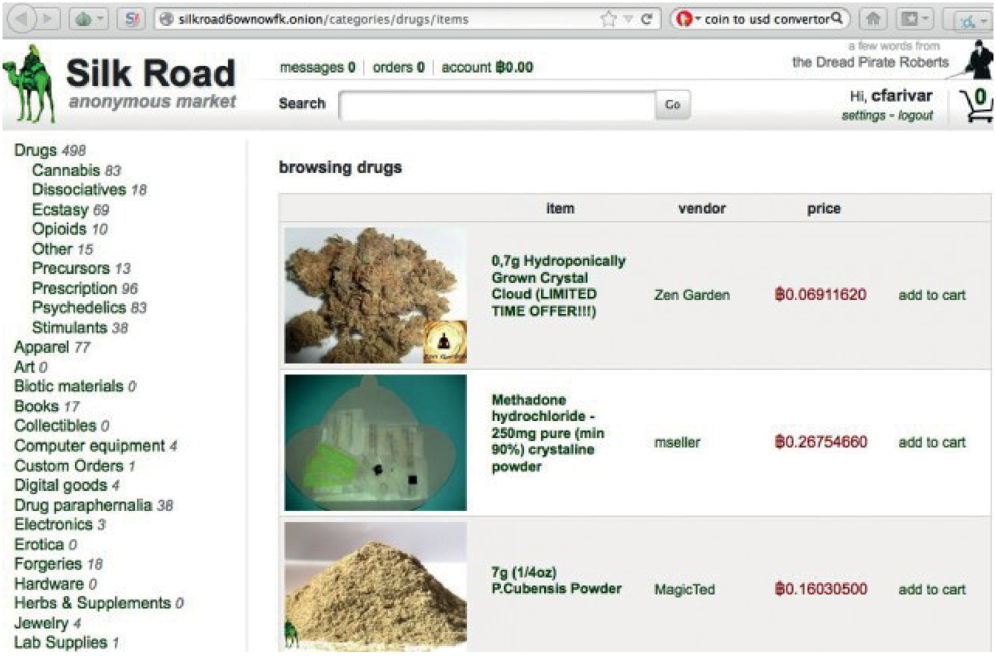
For example, one study scraped data on cannabis listings from November 2013 to October 2014 and showed that the reputation acts as a sufficient self-enforcement mechanism for the marketplace [79]. Another study showed that the Silk Road was a profitable marketplace with a growing and loyal consumer base and the trade volume of illegal drugs was millions of USD per month [75].

Researchers noted revolutionary libertarian ethos of this new online culture [76]. Reputation system replaced the rule of law and customers were satisfied with high quality products [79]. In the economy point of view these marketplaces have no troubles [79]. Also, notably the law enforcement has significant troubles to shut down these marketplaces [76]. Still, it was noted that the greatest threats to Silk Road and other marketplaces are difficult technical security details to operate these online services [76].

Eventually Silk Road was shut down in October 2013 by the Federal Bureau of Investigation (FBI) [80]. The FBI arrested Ross William Ulbricht, a 29-year-old software engineer and outspoken libertarian, who created the Silk Road marketplace [80]. Ulbricht did several technical and operation security mistakes which lead the FBI to finally locate him and where his onion website was hosted [80]. In the upcoming chapter 3 and in our publication IV we will carefully explain the details and what are the typical technical and operation security mistakes.

Immediately a month after the original Silk Road was shut down Silk Road 2.0 was launched to fill the void [75]. Again this marketplace was studied using crawling and scraping and the results showed that the trade was shifting to Silk Road 2.0 and growing [75].

There is a customer base and economical drive for these marketplaces. As a result, currently, there are around 20 different very popular black marketplaces in the Tor network and there is competition between the marketplaces. In short, these marketplaces operate like the Ebay online service but for illegal products. Each of these popular marketplaces implement internal Bitcoin wallet systems, escrow systems, and reputation systems. Take a look at the screen capture of the Silk Road 2.0 website when it was launched in 2013 (Figure 2.5).



Drugs 498
Cannabis 83
Dissociatives 18
Ecstasy 69
Opioids 10
Other 15
Precursors 13
Prescription 96
Psychedelics 83
Stimulants 38
Apparel 77
Art 0
Biotic materials 0
Books 17
Collectibles 0
Computer equipment 4
Custom Orders 1
Digital goods 4
Drug paraphernalia 38
Electronics 3
Erotica 0
Forgeries 18
Hardware 0
Herbs & Supplements 0
Jewelry 4
Lab Supplies 1

browsing drugs

	item	vendor	price	
	0,7g Hydroponically Grown Crystal Cloud (LIMITED TIME OFFER!!!)	Zen Garden	฿0.06911620	add to cart
	Methadone hydrochloride - 250mg pure (min 90%) crystaline powder	mseller	฿0.26754660	add to cart
	7g (1/4oz) P.Cubensis Powder	MagicTed	฿0.16030500	add to cart

Figure 2.5 New marketplaces were founded after the FBI shut down the original Silk Road. *"Just a month after shutdown, Silk Road 2.0 emerges - another anonymous person online has taken up the persona of Dread Pirate Roberts and created The Silk Road 2.0"*, source `https://arstechnica.com/information-technology/2013/11/just-a-month-after-shutdown-silk-road-2-0-emerges/`

Some academic researches pointed out that these black marketplaces may have positive impacts to the society. In the consumer point of view the price and quality of illegal drugs are excellent because the reputation system works, the law enforcement cannot efficiently trace down buyers or sellers, and actually these marketplaces provide safer alternative for regular drug deals. In the publication *Responsible vendors, intelligent consumers: Silk Road, the online revolution in drug trading (2014)* [76] Van Hout and Bingham wrote:

> *"Given Silk Road's potential built in quality mechanisms for safer drug use, previous commentaries have suggested the need for a shift in drug policy focus toward reducing consumer demand for web retailed drugs and optimising on the site's capacity for encouraging safer forms of drug taking amongst a very hard to reach and informed drug using population."*

The digitalization of the drug trade is difficult to tackle and it is a growing trend.

The law enforcement is unable to prevent this shift. More research is needed and we can use novel technical means to carefully study how people use these marketplaces.

3 INVESTIGATING USAGE OF ONION SERVICES

Now we carry out our experiments in the Tor network and we study how real usage of onion services manifests itself. An overview map below shows the anonymous online environment where we perform our investigations.

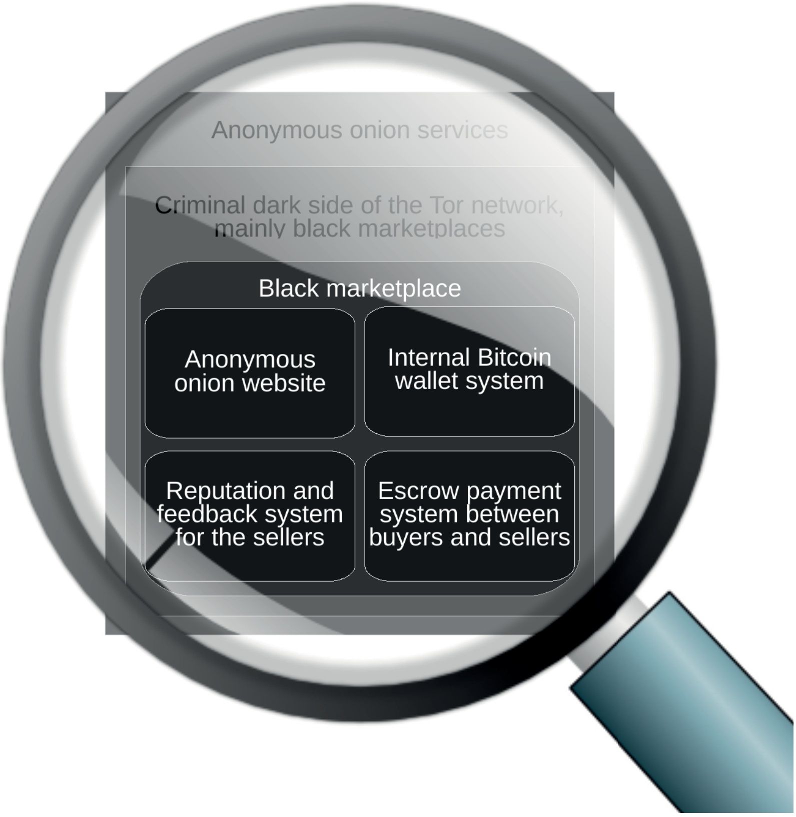

Anonymous onion services
Criminal dark side of the Tor network, mainly black marketplaces
Black marketplace
Anonymous onion website
Internal Bitcoin wallet system
Reputation and feedback system for the sellers
Escrow payment system between buyers and sellers

Figure 3.1 The aim of this thesis is to determine the usage of onion services in the Tor anonymity network. To be more precise the aim is to discover and measure human activities on Tor and on anonymous onion websites.

Objective I We demonstrate how to carry out experiments in the Tor anonymity network. **Objective II** We represent our data of criminal usage of onion websites. To understand the context, we categorise the security limitations of the Tor network.

3.1 Limits of anonymity

How safe is Tor and what are the limitations? We cover this topic in our publication IV, where we cover what are the real Tor de-anonymization techniques which have been applied in the real-world. If you like, instead of reading the publication, you can also watch the author's one hour lecture regarding the topic `https://www.youtube.com/watch?v=zjp3wvX3Rqc`.

Earlier we mentioned surveillance agencies and global adversaries which are spying Internet traffic. Fortunately, we know a little bit regarding their technical capabilities and their attacks against Tor.

In 2013, ex-NSA contractor Edward Snowden revealed that the United States National Security Agency (NSA) has been spying both U.S. citizens and foreigners all over the world [81]. During the same year, a federal judge in Washington, Judge Richard Leon, ruled that these *Orwellian programs probably violate the fourth amendment in the U.S. constitution* [81]. We gain interesting insight to the security of Tor by reading the actual top-secret NSA documents. The NSA has tried without achievements to attack Tor users according to their own presentations on Tor [82, 83].

Remember that Tor is designed to support free journalism. An interesting detail is that Snowden actually used Tor to obtain anonymity while contacting journalists. Famously, the first news photo of Snowden with his laptop shows two stickers, a Tor sticker and an EFF sticker (see Figure 3.2).

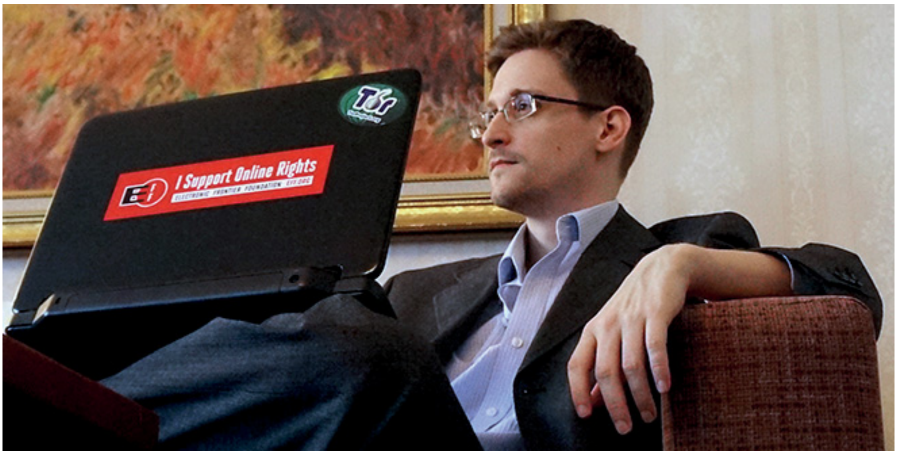
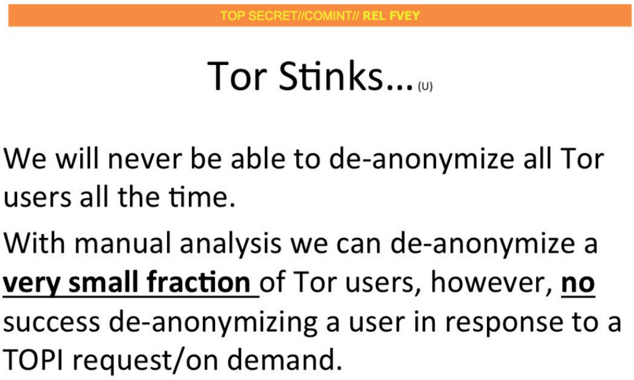


Figure 3.2 Ex-NSA contractor Edward Snowden revealed that the United States National Security Agency (NSA) has been spying both U.S. citizens and foreigners all over the world, and tried to track Tor users. He used Tor himself to anonymously contact journalists. His laptop has a Tor sticker and an EFF sticker. Photo: Barton Gellman for The Washington Post.

The NSA documentation *Tor Stinks* (see Figure 3.3) states that "We will never be able to de-anonymise all Tor users all the time" but "with manual analysis we can de-anonymise a very small fraction of Tor users" [83].



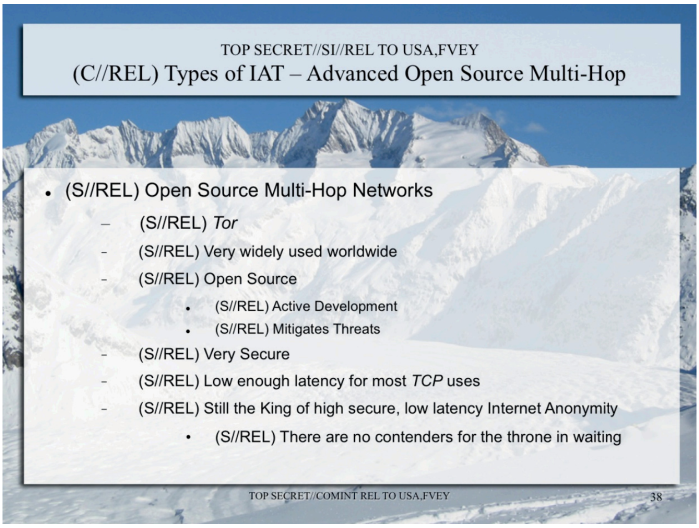
TOP SECRET//COMINT// REL FVEY

Tor Stinks... (U)

- We will never be able to de-anonymize all Tor users all the time.
- With manual analysis we can de-anonymize a **very small fraction** of Tor users, however, **no** success de-anonymizing a user in response to a TOPI request/on demand.

Figure 3.3 The NSA's top-secret document states that "We will never be able to de-anonymise all Tor users all the time" but "with manual analysis we can de-anonymise a very small fraction of Tor users".

Another document (see Figure 3.4) writes that Tor is *the king of high-secure, low-latency anonymity* and *there are no contenders to the throne in waiting* [82].



TOP SECRET//SI//REL TO USA,FVEY

(C//REL) Types of IAT – Advanced Open Source Multi-Hop

- (S//REL) Open Source Multi-Hop Networks
  - (S//REL) *Tor*
  - (S//REL) Very widely used worldwide
  - (S//REL) Open Source
    - (S//REL) Active Development
    - (S//REL) Mitigates Threats
  - (S//REL) Very Secure
  - (S//REL) Low enough latency for most *TCP* uses
  - (S//REL) Still the King of high secure, low latency Internet Anonymity
    - (S//REL) There are no contenders for the throne in waiting

TOP SECRET//COMINT REL TO USA,FVEY 38

Figure 3.4 Tor is *the king of high-secure, low-latency anonymity* according to the NSA.

This reveals that Tor provides high-level online anonymity and is safe even against global adversaries like the NSA. However, throughout the years a few users have lost their anonymity. In publication IV we researched challenges that Tor users will encounter and what are the real world cases where someone has lost their anonymity. We noticed that these cases fall under four categories.

1. Operation security (OPSEC) is difficult.
2. End user devices may still be vulnerable to conventional cyberattacks.
3. Onion services may be attacked by exploiting unintentional features of server software that is not designed to be installed on Tor.
4. Under special conditions it is possible to carry out traffic and timing correlation attacks.

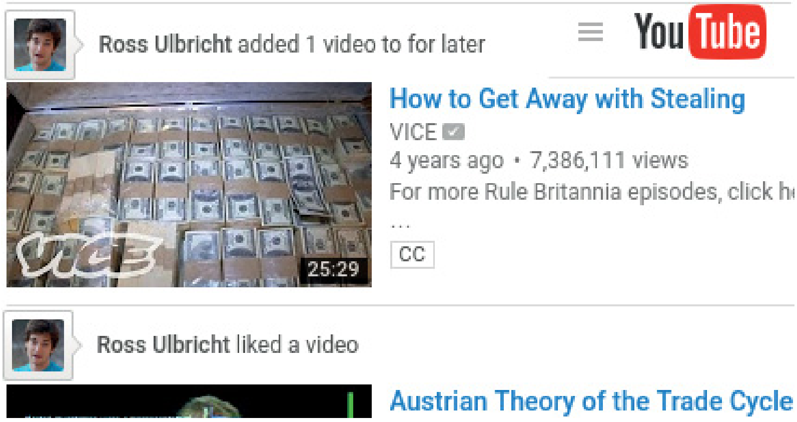
3.1.1 Operation security is difficult

First, OPSEC seems to be very difficult even for technically clever people who understand these limitations, for instance, the founder of the first Silk Road, Ross

William Ulbricht, also known under pseudonyms *Dread Pirate Roberts*, *frosty* and *altoid*, leaked pieces of information over time.

In 2011, a user called *altoid* posted publicly on the Bitcoin Talk forum a message titled “a venture backed Bitcoin startup company” with his recognisable email address *rossulbricht@gmail.com* [84]. After that *altoid* advertised Silk Road. Furthermore, simultaneously, *altoid* advertised Silk Road on a magic mushroom discussion board on `shroomery.org`.

Ulbricht’s Google Plus page and his YouTube profile both make multiple references to the Austrian economic theory. At the same time on the Silk Road forums, *Dread Pirate Roberts* shared these same links and cited same theories. Figure 3.5 demonstrates an example of Ulbricht’s behaviour pattern on YouTube.



Ross Ulbricht added 1 video to for later

How to Get Away with Stealing

VICE

4 years ago • 7,386,111 views

For more Rule Britannia episodes, click h

...

CC

Ross Ulbricht liked a video

Austrian Theory of the Trade Cycle

Figure 3.5 Ulbricht and *Dread Pirate Roberts* shared distinguishable interest to similar Austrian economics videos.

Furthermore, later *Dread Pirate Roberts* mentioned that he is in the Pacific time zone.

According to the FBI’s criminal complaint Ulbricht posted the question *“How can I connect to a Tor hidden service using curl in php?”* on stackoverflow.com[1] under his own real name and later changed his username to frosty [85].

On top of everything else, Ulbricht purchased nine counterfeit identification documents with his face but with different names from Canada. U.S. border customs intercepted these counterfeit identification documents, which were addressed to Ulbricht’s San Francisco apartment. The law enforcement suspected that he is using

[1] https://stackoverflow.com/questions/15445285/how-can-i-connect-to-a-tor-hidden-service-using-curl-in-php

these counterfeit documents for illegal activities.

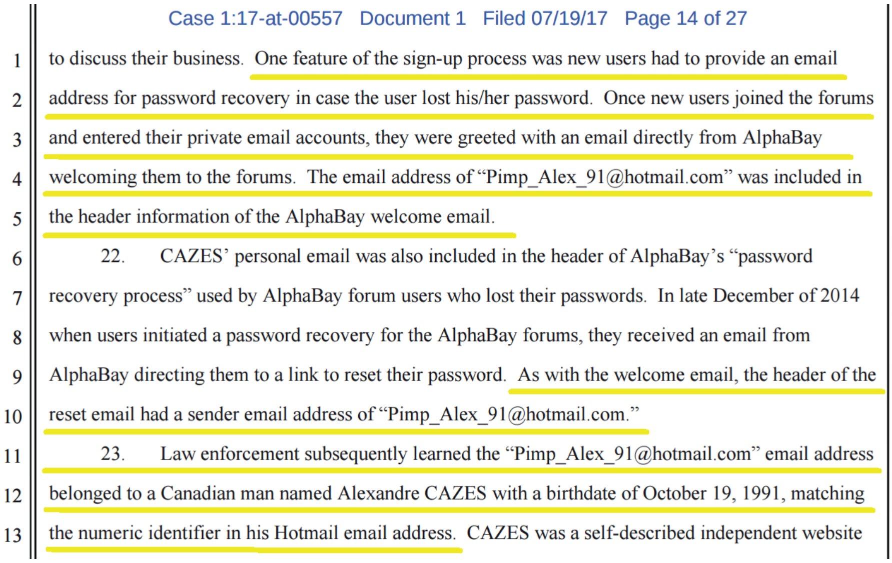
We combined this cumulative data trail in publication IV and created timeline in Table 2.

Date	OPSEC: Leak of critical information
01/2011	Silk Road onion service http://tydgccykixpbu6uz.onion is created.
01/2011	Silk Road portal silkroad420.wordpress.com is created.
01/2011	Silk Road portal starts to advertise Silk Road onion service.
01/2011	*altoid* advertises Silk Road on shroomery.org forum.
01/2011	*altoid* advertises Silk Road on Bitcointalk forum.
10/2011	*altoid* posts a job offer on Bitcointalk, rossulbricht email.
03/2013	Question about Tor and PHP is posted on stackoverflow.com.
03/2013	Ulbricht changes his real name on Stack Overflow to *frosty*.
07/2013	A routine border search intercepts a package of fake IDs.

Table 2. The main OPSEC-related events that linked Ulbricht to Silk Road through the pseudonyms *altoid*, *frosty*, and *Dread Pirate Roberts*.

Eventually, the FBI closed in on the suspect Ross Ulbricht with warrants, technical investigation and by following him. In 2013 he was arrested. The FBI seised his open laptop in a public library and this laptop provided the final evidence to convict Ulbricht as the founder of Silk Road marketplace.

Another similar case happened in 2017. Alexandre Cazes, the 25-year-old Canadian and founder of the AlphaBay marketplace, used same email address in his Linkedin profile and on Alphabay website (Figure 3.6) [86]! The email address was visible to every new user in a welcome email when they signed up to the world's largest online drug marketplace. The identical email address was visible in the Linkedin profile of Alexandre Cazes.



Case 1:17-at-00557 Document 1 Filed 07/19/17 Page 14 of 27

to discuss their business. One feature of the sign-up process was new users had to provide an email address for password recovery in case the user lost his/her password. Once new users joined the forums and entered their private email accounts, they were greeted with an email directly from AlphaBay welcoming them to the forums. The email address of "Pimp_Alex_91@hotmail.com" was included in the header information of the AlphaBay welcome email.

22. CAZES' personal email was also included in the header of AlphaBay's "password recovery process" used by AlphaBay forum users who lost their passwords. In late December of 2014 when users initiated a password recovery for the AlphaBay forums, they received an email from AlphaBay directing them to a link to reset their password. As with the welcome email, the header of the reset email had a sender email address of "Pimp_Alex_91@hotmail.com."

23. Law enforcement subsequently learned the "Pimp_Alex_91@hotmail.com" email address belonged to a Canadian man named Alexandre CAZES with a birthdate of October 19, 1991, matching the numeric identifier in his Hotmail email address. CAZES was a self-described independent website

Figure 3.6 AlphaBay Market was the largest online black marketplace when it was shut down in 2017. According to the complaint affidavit the FBI investigated an email address found from the site: the founder Alexandre Cazes used the same email address in his Linkedin profile and on Alphabay website [86]. Underlining added by the author.

It is easy to conclude that anonymous Tor users should follow a strict OPSEC process, however, as several cases like these indicate, it is extremely difficult even for security focused software developers.

> *"The art of war teaches us to rely not on the likelihood of the enemy's not coming, but on our own readiness to receive him; not on the chance of his not attacking, but rather on the fact that we have made our position unassailable."* [2]

It is easy to give away the advantage of the perfect technical anonymity by a cer-

[2] The Annotated Art of War, Parts 8.3-11: Advantages, written originally by Sun Tzu around 500 BC

tain behavioral pattern or a combination of identifying information. An adversary can follow this cumulative data trail over time and learn the real identity of the Tor user.

3.1.2 Weak security of the end user device

Secondly, the end user devices may still be vulnerable to conventional cyberattacks. Applications which are installed behind the Tor network have been exploited even if the Tor network itself is safe. In publication IV we studied these cases.

In 2013, the FBI seised *Freedom Hosting* service provider which was hosting several onion services. Instead of closing these services the FBI spread a special malware software through quite a few of them. The malware was a memory-management exploit against the (Firefox) Tor Browser [87]. *Freedom Hosting* was facilitating child abuse onion websites and the FBI injected their malicious JavaScript exploit code to these sites. Before Firefox developer Mozilla patched the vulnerability, the JavaScript exploit was able to identify the MAC address and hostname of the user and sent it as HTTP requests to the FBI. The HTTP request exfiltrated the user's real IP address and exposed this to the FBI.

In 2015 the FBI carried out a similar operation and took over a child abuse onion website called *Playpen* and operated it two weeks before closing it down; the FBI continued to distribute child abuse media material with a special piece of malware included [88]. This malware revealed the real IP addresses of the visitors [88].

Then as well, in 2016, an anonymous warning was sent to the popular tor-talk mailing list. This email published a previously unknown exploit that was injected to an onion website that was sharing child abuse material [89]. Again, a piece of malicious JavaScript code was exploiting the (Firefox) Tor Browser and exposed the user's real IP address. Identically to the one found in 2013, this memory-management exploit was able to call kernel32.dll on Windows operating systems and executed the attacker's commands. Again, Mozilla quickly fixed this Firefox vulnerability.

Most software applications are not designed for anonymity. Repeatedly they leak information regarding the end user device. Modern web browsers are not able to open all types of files and malicious files can be shared to the user. Although the Tor Browser clearly warns that opening these files outside of the browser is not safe, because applications could connect to the Internet without Tor, still it is very easy to make the mistake.

The only known secure way to mitigate these risks is to create a separate environment which only permits traffic through the Tor network and prevents other connections to the Internet. This way every application operates behind the Tor network. There are two Linux distributions that provide this anonymous environment for every application, Tails and Whonix [90, 91].

Actually, the adversary does not need to find novel exploits against software; instead, it is feasible to exploit existing features of applications that are not designed to protect anonymity. For example, popular document viewers fetch online resources, such as images and style sheets. This will cause the applications to leak the real IP address of the user to the attacker when these kind of files are opened.

It is difficult to perceive which file types are safe to open without side effects. Modern applications are so complex that it is impossible to determine what kind of side effects for privacy they are causing. We studied the most simple examples to demonstrate this danger for anonymity.

For example, an mp3 music file format and mp3 player applications are rather simple. Still, a simple mp3 file with a popular m3u file is a dangerous combination for privacy. Popular media players look up the album image cover, which is in the m3u file, from the online source. The attacker can share this kind of mp3 and m3u files together and detect the incoming connections from the user's real IP address.

It is important to conclude that in the point of view of the media player developers this is not even an important security feature because they are not intending their player to provide anonymity. When the author tested very popular free media players many of these fetched the album cover image without asking permission to do this. Even worse, the most popular media player on Linux, VLC player, had a user interface bug: the player actually asks permissions but when you select the option not to download from the online sources the player downloaded the image cover anyway [92]!

3.1.3 Unintentional features of server software behind Tor

Now, thirdly, we examine a few common mistakes and flaws that may reveal critical information regarding the location of an onion service. Again, it is important to notice that although a TCP based server software can be installed as an onion service these applications are not designed to provide anonymity by default. Instead, there are several unintentional ways server applications reveal critical information when installed behind the Tor network.

Let us examine a typical SSH service that is installed to operate through an onion address. The SSH shows a unique fingerprint of the service before login. The feature is intended for validating the SSH server's identity against man-in-the-middle attacks. However, repeatedly people install the same SSH service on a public IP address and through an onion address. Unfortunately, this reveals the IP address of the onion service. The following demonstration tests SSH connections to the onion address and to the public IP address.

```
# torsocks ssh root@msydqstlz2kzerdg.onion
RSA key fingerprint a7:93:84:a6:97:fa:25:65:77:c9:58:bb:fe:8e:e2:2f
# ssh root@ahmia.fi
RSA key fingerprint a7:93:84:a6:97:fa:25:65:77:c9:58:bb:fe:8e:e2:2f
```

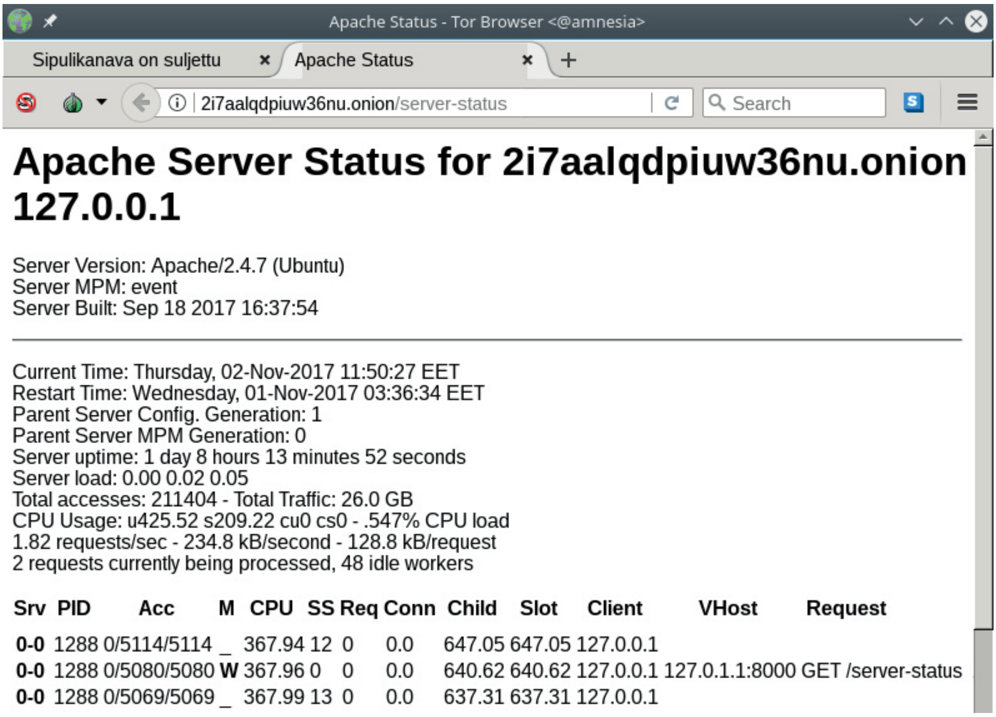
As a result, we see that `ahmia.fi` and `msydqstlz2kzerdg.onion` SSH servers have the same fingerprint. We revealed the real address of the onion service.

Tor software cannot warn against this unintentional configuration. It is easy to share the same Internet service simultaneously through a public IP address and an onion address and reveal the location of the service.

Another Achilles' heel is an aftereffect from the fact that server applications use Tor through SOCKS. In the point of view of the server software the connections are coming from localhost. This causes a new danger for anonymity: major web frameworks treat localhost as a safe zone and offer special features to the users coming from the localhost. In this case, every onion website visitor can access these features.

A very popular Apache HTTP Server has the Apache Server Status module which provides a statistical information view to localhost connections `http://127.0.0.1/server-status/`. Normally users coming from localhost have already login access to the server computer itself and it is natural to consider a localhost as a safe zone.

However, when Apache is connected to Tor through localhost SOCKS connection the onion website displays the page `http://address.onion/server-status/` publicly. The information page shows a lot of information regarding the server itself, including uptime, number of connections and even rare cases the real IP address of the server. National Bureau of Investigation (Finland) seised one of the Finnish discussion boards and pointed its traffic to their server which was leaking server status information (Figure 3.7). For anonymity, even one of these pieces of information could be critical and lead to de-anonymisation of the onion service.


Apache Status - Tor Browser <@amnesia>
Sipulikanava on suljettu
Apache Status
2i7aalqdpiuw36nu.onion/server-status
Search

Apache Server Status for 2i7aalqdpiuw36nu.onion 127.0.0.1

Server Version: Apache/2.4.7 (Ubuntu)
Server MPM: event
Server Built: Sep 18 2017 16:37:54

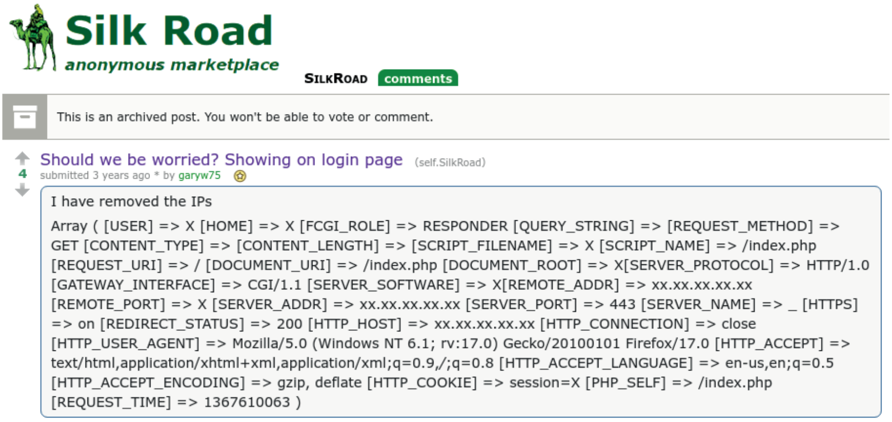
Current Time: Thursday, 02-Nov-2017 11:50:27 EET
Restart Time: Wednesday, 01-Nov-2017 03:36:34 EET
Parent Server Config. Generation: 1
Parent Server MPM Generation: 0
Server uptime: 1 day 8 hours 13 minutes 52 seconds
Server load: 0.00 0.02 0.05
Total accesses: 211404 - Total Traffic: 26.0 GB
CPU Usage: u425.52 s209.22 cu0 cs0 - .547% CPU load
1.82 requests/sec - 234.8 kB/second - 128.8 kB/request
2 requests currently being processed, 48 idle workers

Srv	PID	Acc	M	CPU	SS	Req	Conn	Child	Slot	Client	VHost	Request
0-0	1288	0/5114/5114	_	367.94	12	0	0.0	647.05	647.05	127.0.0.1		
0-0	1288	0/5080/5080	**W**	367.96	0	0	0.0	640.62	640.62	127.0.0.1	127.0.1.1:8000	GET /server-status
0-0	1288	0/5069/5069	_	367.99	13	0	0.0	637.31	637.31	127.0.0.1		

Figure 3.7 National Bureau of Investigation (Finland) seised one of the Finnish discussion boards and when they directed its onion address to their Apache server they were leaking their server's information.

In addition, because popular web frameworks that are not designed to be installed on Tor and are not designed to provide anonymity for the server, popular web frameworks frequently leak error messages that include critical server information. This happened to the notorious Silk Road marketplace: it showed for some time an error message which contained the IP address of the server, see screen capture image 3.8 from Reddit.com discussion website.

Unfortunately, Tor adds an extra layer of complexity to already complex web framework security. In addition, it only requires to leak once a tiny piece of information to completely compromise anonymity.



This is an archived post. You won't be able to vote or comment.

Should we be worried? Showing on login page (self.SilkRoad)

4

submitted 3 years ago * by garyw75

I have removed the IPs

```
Array ( [USER] => X [HOME] => X [FCGI_ROLE] => RESPONDER [QUERY_STRING] => [REQUEST_METHOD] =>
GET [CONTENT_TYPE] => [CONTENT_LENGTH] => [SCRIPT_FILENAME] => X [SCRIPT_NAME] => /index.php
[REQUEST_URI] => / [DOCUMENT_URI] => /index.php [DOCUMENT_ROOT] => X[SERVER_PROTOCOL] => HTTP/1.0
[GATEWAY_INTERFACE] => CGI/1.1 [SERVER_SOFTWARE] => X[REMOTE_ADDR] => xx.xx.xx.xx.xx
[REMOTE_PORT] => X [SERVER_ADDR] => xx.xx.xx.xx.xx [SERVER_PORT] => 443 [SERVER_NAME] => _ [HTTPS]
=> on [REDIRECT_STATUS] => 200 [HTTP_HOST] => xx.xx.xx.xx.xx [HTTP_CONNECTION] => close
[HTTP_USER_AGENT] => Mozilla/5.0 (Windows NT 6.1; rv:17.0) Gecko/20100101 Firefox/17.0 [HTTP_ACCEPT] =>
text/html,application/xhtml+xml,application/xml;q=0.9,/;q=0.8 [HTTP_ACCEPT_LANGUAGE] => en-us,en;q=0.5
[HTTP_ACCEPT_ENCODING] => gzip, deflate [HTTP_COOKIE] => session=X [PHP_SELF] => /index.php
[REQUEST_TIME] => 1367610063 )
```

Figure 3.8 According to Reddit.com conversation the original Silk Road marketplace website showed this error message which contained the public IP address of the server.

3.1.4 Special conditions enable traffic and timing correlation attack

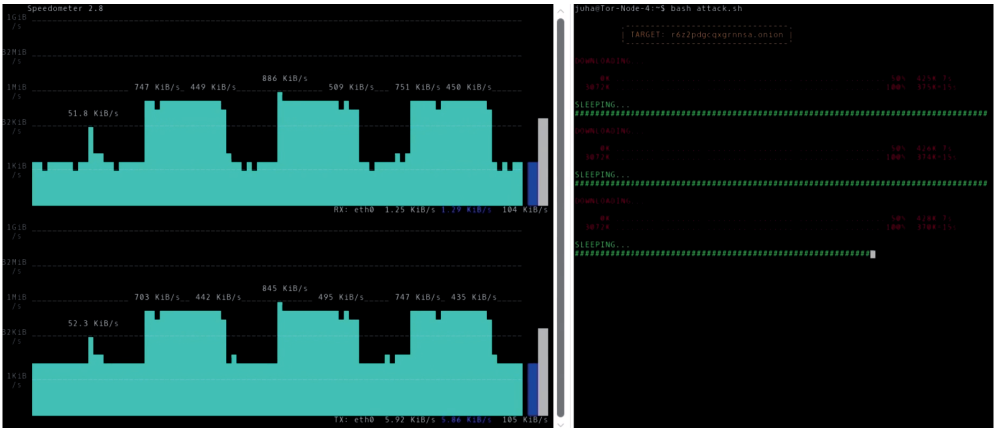
As the background shows Tor traffic could be de-anonymised using end-to-end traffic and timing correlation attack. When the attacker is observing traffic of the first relay (entry guard) and traffic of the destination (onion service, exit relay) statistical analysis to discover the real IP address of the user can be applied. Through correlation attacks it is possible to attack against anonymity by passive traffic monitoring without actually having the access to the relays themselves. It is enough that the attacker observes the traffic.

According to the leaked NSA documents, intelligence agencies and other large attackers can catch only a fraction of the Tor traffic so, in general, it is not feasible to carry out this attack [83, 93]. However, having said that, occasionally the correlation attack does not require enormous traffic monitoring and sophisticated statistical analysis. There are known cases where the adversary effectively de-anonymises a Tor user with traffic and timing correlation. In publication IV we studied these cases.

For instance, in 2013, a 20-year-old Harvard University student sent hoax bomb threats using Tor to get out of a final exam [94]. According to the FBI affidavit, the FBI noted that these emails came from a free email service that creates temporary email addresses [94]. The service embedded the IP address in outgoing email which pointed to Tor exit node. The investigators suspected that a student used Tor to send these emails. They found out that there has been only one student connected to the Tor network from the university wireless network while these emails were sent. This correlation led the investigators to interrogate the student. The student confessed and was arrested.

When the number of clients is tiny in the anonymity system, the degree of anonymity is not sufficient [95, 96]. If the context or the content of the anonymous communication contains information which narrows the anonymity set then it is possible to investigate every suspect. In previous case there was a reason to assume that the Tor use came from the university network. In this case there was only one Tor user so it was easy to interrogate him. Still, note that he had plausible deniability until he confessed.

Let us test more sophisticated traffic and timing correlation attacks against onion services and Tor users. We demonstrate an example of a traffic correlation attack against a fleet of our onion websites which are serving HTTP content. We selected our own relays as the entry guards for these onion websites. We sent a distinguishable HTTP traffic pattern to our onion websites and simultaneously observed the traffic through their entry guard relays. As a result, from this correlation, we are able to reveal the real IP address of the onion websites. Figure 3.9 shows this test and the traffic pattern between the onion websites and the entry guard. In theory, a similar passive traffic analysis could be carried out against the Tor network in wider range.



```
juha@Tor-Node-4:~$ bash attack.sh

         | TARGET: r6z2pdgcqxgrnnsa.onion |

DOWNLOADING...

     0K ..........................................  50%  425K 7s
  3072K .......................................... 100%  375K=15s

SLEEPING...
################################################################################
DOWNLOADING...

     0K ..........................................  50%  426K 7s
  3072K .......................................... 100%  374K=15s

SLEEPING...
################################################################################
DOWNLOADING...

     0K ..........................................  50%  428K 7s
  3072K .......................................... 100%  370K=15s

SLEEPING...
##########################################################
```

Figure 3.9 A distinguishable traffic pattern all the way to the guard relay of the onion service. As a result, it is possible to locate the onion service. We tested this with our onion services and relays: on right there is a script that connects to the onion websites and sends a shaped traffic pattern, and on left there is a monitoring view of the guard relay traffic.

Another possible way to execute a correlation attack and reveal Tor user's activities is to learn what kind of traffic fingerprints services have. Again, this kind of correlation is a known problem for an anonymity network that provides TCP connectivity. This type of attack can be weaponised against web browsing without decrypting the actual traffic; only the traffic pattern of a web site is relevant.

Similarly, the traffic pattern can be detected when a same website is visited several times (see Figure 3.10). Even between completely different Tor circuits the fingerprint remains the same for the same website; and different websites have recognizable fingerprints.

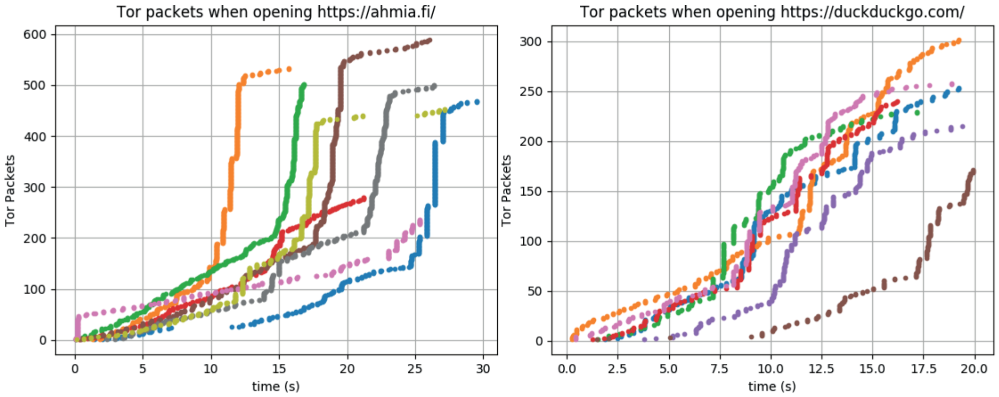


Figure 3.10 We see that a distinguishable traffic pattern is repeated when opening a website seven times. Every time there was a different Tor Browser session with completely new circuit. It is possible to gather these website fingerprints and reveal significant information regarding the possible website a user is opening without decrypting the traffic. These two different websites leave recognizable traffic fingerprints.

The limitation of this attack is that there are hundreds of millions of websites in the world and similar websites leave identical traffic fingerprints. It is not feasible to gather fingerprints of every website, and certain traffic fingerprints are not unique and shared between multiple websites.

It is known that the NSA and the British Government Communications Headquarters (GCHQ) already audit connections to the Tor network [93]. According to classified documents leaked by Snowden, the GCHQ gathers data from major fibre-optic cables, saves it and later searches patterns from there. However, intelligence agencies can only watch a tiny fraction of the Tor traffic but they could occasionally de-anonymise an onion service.

Tor is safe because there are over two million Tor users every moment, around ten thousand onion services, and seven thousand relays all over the world. The enormous anonymity set makes it extremely difficult to target a certain Tor client or onion service.

3.2 Watching incoming traffic to onion honeypots

In publication II we watched incoming traffic to our honeypot onion services in order to study if unknown attackers spy the DHT onion address directory in the Tor network. In general, a honeypot is a decoy service which attracts the attacker. Connections to the honeypot can be analysed in order to profile the attacker. We saw incoming traffic so the answer is yes: unknown attackers were spying onion addresses from the DHT. When publishing the article it was possible to discover this way which onion addresses exists. Later, the Tor Project has developed a recent version of the protocol to mitigate the problem.

The Tor network has a global hidden service directory (HSDir). This is a distributed hash table of onion addresses and their descriptors. Every HSDir relay has a subset of the descriptors created by onion services. An onion service's description contains a list of its introduction points. When connecting to an onion address the client sends an introduction message to one of the introduction points of the onion address.

Onion addresses publish their descriptors to a set of six responsible HSDirs once per hour. The HSDirs are actually regular Tor relays which have been online longer than 92 hours and received the HSDir flag from the Tor directory authorities. Every HSDir has certain onion addresses according to its position in the DHT. Relay's fingerprint orders their position in the DHT. Based on their position they have description-onion list of certain onion address.

HSDirs could spy their piece of the DHT and reveal unknown onion addresses accordingly. Due to the possibility of this spying method, we decided to study the following questions:

1. Are unknown attackers spying the DHT onion address directory?
2. How huge scale is the monitoring?
3. What is the level of automation?
4. Which Internet services are targeted?

To answer these questions we installed one hundred onion honeypots for the experiment. The addresses were randomly generated so they distributed all over the DHT. We did not share the onion addresses publicly. Each honeypot listened three popular service ports: port 21 (FTP), port 22 (SSH) and port 80 (HTTP). After installation we ran them for 42 days and logged every connection.

When setting up an onion address there is a certain expectation that the HSDir is not unethically harvesting addresses and after that scanning onion services. Even so, we proved that this is the case: the onion honeypots experiment reveals us that there are unknown separated entities especially examining onion websites. Of course, attribution of the anonymous attackers is impossible.

We believe that the fact that these onion addresses exist was only known by HSDirs. Likelihood to expose these addresses by guessing is extremely low, and it would take serious effort to penetrate our Linux servers to reveal the addresses.

Out of our 100 honeypot onions 30 received traffic. As a result, our conclusion is that a tiny fraction of the HSDirs were spying onion addresses. Surprisingly only port 80 (HTTP) was connected and 82 percent of the HTTP requests came from Mozilla browser (the Tor Browser) user agents. Although crawlers as well could have faked their user agents.

The experiment gave us the answer that although a minor fraction of the HSDirs are spying onion addresses they are not probably controlled by the same attacker. The attackers are only interested in websites.

3.3 Following illegal trade on black markets

In publications I and III we followed illegal drug trade on anonymous black marketplace Silkkitie, also known as Valhalla.

In 2013 the Silkkitie was founded by *Kapteeni* (captain). It started as a local marketplace, the sellers were from Finland and the only available language was Finnish. Later it opened its international English version, Valhalla.

3.3.1 The marketplace

Silkkitie followed the model developed by the original Silk Road: it uses Tor to obtain anonymity, Bitcoin as the payment method, escrow system to prevent abuse, and vendor feedback system to show transparent reputation.

After Hansa Market, Dream Market and AlphaBay Market were closed in 2017 the Silkkitie marketplace is the oldest running marketplace. It is suspected to be one of the largest marketplaces in the Tor network so far, however, it is difficult to confirm that from a reliable source.

Silkkitie offered immediately an interesting view to Finnish online drug trade: indeed, the marketplace was in Finnish, sellers and buyers were Finns, and the shipments proceeded through domestic mail in Finland. We started to study the anatomy of the function of the marketplace already in 2014. We found that an anonymous marketplace needs four components to operate.

1. Anonymous and censor-resistant way to operate: for example, onion website.
2. Online (semi-)anonymous monetary system: for example, Bitcoin.
3. Escrow payment system: internal escrow accounting.
4. Reputation and feedback: transparent reputation metric.

When these four technologies are combined the market can operate. Without one of these the market will cease its operation. Also, there is a Bitcoin mixer system inside the Silkkitie marketplace: there are multiple mixed wallets and without a distinct link in the Bitcoin blockchain between the input transaction to Silkkitie's wallet and payment from this wallet to the seller.

In addition, using web-scraping tools, the author revealed the amount of different types of illegal substances traded on Silkkitie.

Using daily web-crawling and web-scraping we extracted information from the Silkkitie onion website and gathered a database of transactions between 5 November 2014 and 23 September 2015. Our method is similar to earlier research where Silk Road 1 and 2 was studied [22, 25, 75]. Our scraper software extracted the product information from every product page on Silkkitie and collected the fields title, price, number of items in stock, nickname of the seller, positive and negatived feedbacks, country of the seller, which countries the seller is willing to send shipments, and the unique URL address of the product page.

We executed the crawler software once a day over 11 months. This gathers enough data to study day-to-day drug trade. 36 days are missing from the data because the Silkkitie website was not available. Half of these days took place in June and July

2015 when the website was offline but came back after this break. There is no information available why it was offline.

The key result is the drug trade data itself; we have now a day-to-day database of drug trade on Silkkitie. Nearly a full year can be studied from the data. The quantitative research showed the current estimations of the drug trade (see Figure 3.11).

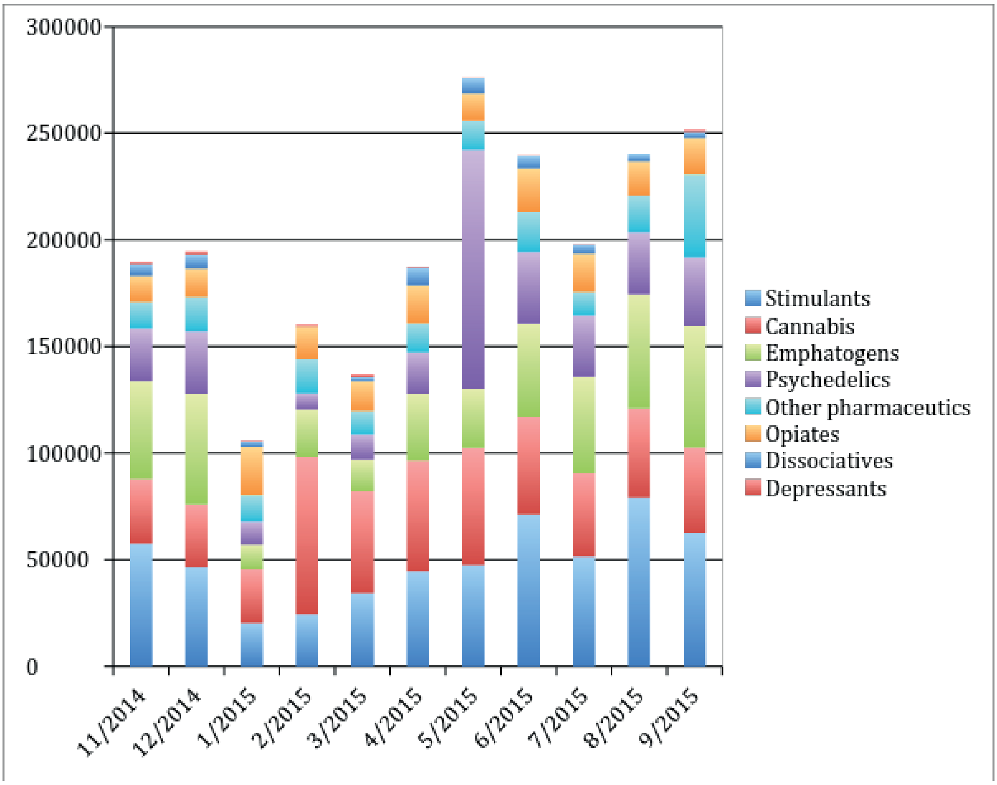


Figure 3.11 Sales of different illegal substances on Silkkitie between 5 November 2014 and 26 September 2015 between Finnish buyers and sellers. From III.

We found that sales totaled over two million euros during the study. Measured in euros, stimulants were most widely purchased, followed by cannabis, MDMA, and psychedelics. Note that when compared to the fact that cannabis is cheaper per dosage than other drug it is overwhelmingly most popular drug that was bought.

For this particular study we only calculated sales that took place within Finland by removing products where the location of the seller was abroad or the seller was willing to sell abroad.

This database includes 93 878 observations where 260 sellers offered 3823 products. Drug sales were calculated using the change on available stock. When a seller sells a product the capacity lowers so we can estimate the number of sales during the specific day.

Furthermore, we can model the marketplace dynamics and trade. In publication III fixed-effects regression models were applied to estimate the effects of reputation and capacity. Our finding was that the reputation and the capacity of a seller directly impact the sales positively.

The reputation of the seller is a combination of positive and negative feedbacks. The total reputation ranged from -19 to 3418. Surprisingly, only few sellers have total negative reputation. This means that the marketplace dynamics under reputation system forces the seller to avoid negative feedback.

Due to the anonymity and nature of illegal sales, reputation systems and capacity information have replaced the rule of law: a buyer must trust the seller's reputation because the law is not guaranteeing the delivery. The only available information is the seller's reputation and capacity and what the seller writes.

Within our data, we noticed that occasionally products were not sold at all and that several sellers remained active for only a short period of time and risk-averse online consumers avoid such sellers. At least in our data, consumers selected those sellers who had higher reputations and capacity.

3.3.2 The insider with many names

In 2018 the author contacted a prisoner inside Helsinki prison to resolve unanswered questions.

There are three popular views of this prisoner. Firstly, multiple online discussion boards, especially on Tor, warmly describe him as an honest drug lord who delivered high quality laboratory tested drugs with clear dosage and usage instructions through domestic mail in Finland. For those, he is like *Prometheus*.

Prometheus is a Titan in Greek mythology who defies the gods by stealing fire and giving it to humanity enabling the progress of the civilization. Zeus sentenced immortal Prometheus to be chained to a mountain where each day an eagle eats his liver and the liver grows back overnight to be eaten again. The hero Hercules frees Prometheus from the chains.[97]

Secondly, of course, the creator of the largest drug syndicate in Finland is comparable to *Faust*.

The story of Faust is a classic German legend, where a man, Faust, is bored and dissatisfied with his highly successful life. After his suicide attempt he calls on the Devil for all the pleasure and knowledge of the world. The Devil's representative, Mephistopheles, appears and makes a deal with Faust. The deal is that Mephistopheles will serve Faust with his magic powers and in the end the Devil will claim Faust's soul. Faust takes the deal.[98]

He coordinated multiple sellers under one brand name to sell tens of kilograms of various substances through Silkkitie marketplace - using the full leverage of Tor's online anonymity, reputation system and Bitcoin payments.

Thirdly, for a few people he is closer to *Mephistopheles* than *Faust*. He visibly advertised illegal drugs and enabled wide scale delivery. In 2014, under his pseudonym *Douppikauppa*, he sent free LSD samples around Finland.

Douppikauppa is the leader of the syndicate of multiple sellers who created a common brand in Silkkitie marketplace. In the open-source software world he is known under pseudonym *Tronic* and was the former lead developer of a container format specification known as Multimedia Container Format, MCF. Also, he kept his daily work as a software developer and simultaneously led the criminal organization. Furthermore, in 2015, he was a candidate in Finnish parliamentary elections.

A year after this 32-year-old chief technology officer and software developer Lasse Kärkkäinen was arrested. Before his arrest he had planned to launch his own marketplace on Tor. The marketplace would have been named after *Douppikauppa* brand name.

Lasse Kärkkäinen is chained for 10.5 years[3] and suffering the consequences of the deal.

Mr. Kärkkäinen is ready to clarify the author's questions.

1. What kind of motives you had in mind when you began to sell drugs?

Mr. Kärkkäinen told that he was first very interested in the Tor network as a technological phenomenon. In addition, he followed how these marketplaces operate inside the Tor network. Apparently, illegal drugs were widely available in the Netherlands. Dutch dealers had a wider range of products and their prices were only

[3] This is practically close to maximum sentences in Finnish justice system, typically given to murderers. The court noted that it is especially aggravating to sell drugs without knowledge of the buyer's age. Kärkkäinen will be probably released to parole after serving half of the sentence.

a fraction of the Finnish prices. He began to wonder how to import these drugs to Finland.

Previously he had planned his road trip through Europe for other reason and saw this as a perfect opportunity to try his skills importing drugs to Finland. He wanted to show himself that he could be a perfect drug dealer.

While selling these drugs in Finland it became apparent to him that there was a market for high quality drugs and for a high quality customer service. The first batch sold out rapidly despite the fact that his prices were above average.

2. Why were you planning to create your own market instead of using existing ones?

Mr. Kärkkäinen told that after Silkkitie marketplace was offline several weeks in the end of 2014 and in the begin of 2015 he and the other members of the sales team decided to search for better marketplaces. Eventually, due to the Silkkitie's issues, they migrated to the Evolution marketplace.

However, none of the existing marketplaces proved to be very usable. Furthermore, he pointed out that there were serious security issues in the marketplace software. Because of this the plan to create their own marketplace was on the table. It would have been closer to a tradition web shop with emphasis of usability than open bazaar. Under well-known brand name of his *Douppikauppa* syndicate. A good brand is important.

According to Mr. Kärkkäinen current marketplaces cause a lot of manual work for any seller who sells several shipments per day. From the logistics point of view they are not offering a good shipment list with addresses and quantities. Instead, there might be a large list of separate messages. Even in the case when one buyer buys multiple similar products the system showed these as separated orders.

3. How did you get caught before launching your own marketplace?

In 2014, the FBI located and seised Silk Road 2 marketplace. Now the FBI had full access to the server of the marketplace and they researched internal messages between buyers and sellers. Among many buyers there was *Redword* who seemed to be buying large amounts of drugs and transferring them to Finland.

The FBI shared the information with the Finnish authorities. The Finnish authorities concluded that there is a Finnish buyer who has bought significant amounts of drugs in the Netherlands. Even the area in the Netherlands was mentioned in the messages. The Finnish authorities started to suspect that indeed these are the shipments that *Douppikauppa* is selling in Silkkitie.

With the help of the Dutch authorities Finnish investigators obtained a list of Finns who had stayed in the area where the exchange took place in the Netherlands. Only few Finns had stayed there at the time. After this they narrowed down their investigations to Mr. Kärkkäinen.

In April 2016, he was arrested in his home by eight police officers. Before that criminal investigators wiretapped his phone and followed his car. In addition, a few members of the syndicate were arrested.

However, Mr. Kärkkäinen had already stopped selling drugs himself. Nevertheless, criminal investigators were able to open his laptop. There they discovered his old bookkeeping and seised a Bitcoin wallet containing a total of 1670 bitcoins.

3. Do you think that drug dealers should actively perform harm reduction?

Mr. Kärkkäinen thinks users are solely responsible of drug usage. A drug dealer can provide facts about the products and this helps to build an honest brand name in the marketplace.

Mr. Kärkkäinen stated that he performed reagent tests for the substances he was selling. Also, he laboratory tested substances and told what is the real purity of his products. In addition, he wrote clear descriptions of how certain substances affect and added dosage information. He even pointed out possible harms and health issues caused by highly addictive products, such as amphetamines.

3.1 If there would be an anonymous method to verify buyers age do you think that you had demanded it?

Mr. Kärkkäinen would had demanded age verification if there would had been an anonymous age verification method available.

3.3.3 Administration's point of view to anonymous conversation

To gain deeper understanding of onion service usage the author interviewed another senior technology professional who operated in the Tor network.

Kim Holviala is better known under his pseudonym which is simply *Ylläpitäjä* (Administration). Thus, he is a system administrator and the creator of a large anonymous discussion board known as *Sipulikanava* (Onion channel). He claimed that his anonymous discussion forum was one of the most popular websites in Finland before Finnish authorities seised it.

In 2014, Mr. Holviala created *Sipulikanava* after a few previous Finnish discussion boards closed down. This created a vacuum waiting to be filled. In his personal life there was vacuum too: he wanted a new technical hobby and as a senior software system consultant he knew exactly how to create and manage large web services. He installed servers in his home and published his onion service.

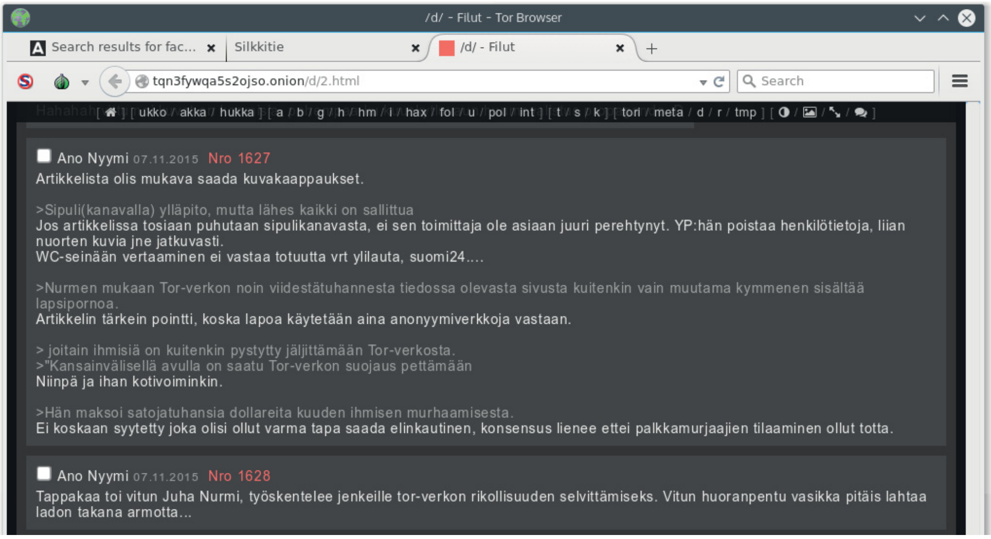
The site was a discussion forum without a need for registration and without premoderation of messages. As a result, the site soon contained plenty of legitimate but controversial content. Moreover, however, there was a section to buy and sell products. After a year this section of the site established itself as a popular place to share contact information to trade illegal drugs in Finland.

Mr. Holviala told that he was arrested because he started to reveal information about himself on purpose. He even made phone calls to prison and had several conversations with Mr. Kärkkäinen. These phone calls were wiretapped by authorities. During these calls he gave out clear clues that he is behind one of the largest discussion boards in the Tor network. Accordingly, he was soon arrested and the site was seised by the Finnish authorities.

The site itself did not implement a marketplace. Instead, a buyer contacted a seller directly with external technical communication methods, such as via anonymous instant messengers and emails. Mr. Holviala's site acted as a service which connected buyers and sellers, thus he is charged with wittingly facilitating drug sales and drug trafficking. Prosecutors demand at least seven years behind bars to him.

In Febuary 2019, Finnish district court found that while the defendant had maintained the website, he was not directly involved in selling drugs, instead, he was found guilty of facilitating the sale of illegal drugs [99]. As a result, he received a three-year and four month jail sentence [99]. But the court noted that this is the first time a website administrator was found guilty facilitating the sale of illegal drugs and did not send him to prison: the higher court level will take the case and it is still unclear what is the final sentence [99].

According to Mr. Holviala there were over two thousand messages per day, unique session cookies tracked over ten thousand users per day, and there were almost one million page loads per week. All messages were in Finnish (Figure 3.12 demonstrates an example conversation).


/d/ - Filut - Tor Browser
Search results for fac...
Silkkitie
/d/ - Filut
tqn3fywqa5s2ojso.onion/d/2.html
Search
[ ] [ukko / akka / hukka] [a / b / g / h / hm / i / hax / fol / u / pol / int] [t / s / k] [tori / meta / d / r / tmp] [/ / /]

☐ Ano Nyymi 07.11.2015 Nro 1627
Artikkelista olis mukava saada kuvakaappaukset.

>Sipuli(kanavalla) ylläpito, mutta lähes kaikki on sallittua
Jos artikkelissa tosiaan puhutaan sipulikanavasta, ei sen toimittaja ole asiaan juuri perehtynyt. YP:hän poistaa henkilötietoja, liian nuorten kuvia jne jatkuvasti.
WC-seinään vertaaminen ei vastaa totuutta vrt ylilauta, suomi24....

>Nurmen mukaan Tor-verkon noin viidestätuhannesta tiedossa olevasta sivusta kuitenkin vain muutama kymmenen sisältää lapsipornoa.
Artikkelin tärkein pointti, koska lapoa käytetään aina anonyymiverkkoja vastaan.

> joitain ihmisiä on kuitenkin pystytty jäljittämään Tor-verkosta.
>"Kansainvälisellä avulla on saatu Tor-verkon suojaus pettämään
Niinpä ja ihan kotivoiminkin.

>Hän maksoi satojatuhansia dollareita kuuden ihmisen murhaamisesta.
Ei koskaan syytetty joka olisi ollut varma tapa saada elinkautinen, konsensus lienee ettei palkkamurjaajien tilaaminen ollut totta.

☐ Ano Nyymi 07.11.2015 Nro 1628
Tappakaa toi vitun Juha Nurmi, työskentelee jenkeille tor-verkon rikollisuuden selvittämiseks. Vitun huoranpentu vasikka pitäis lahtaa ladon takana armotta...

Figure 3.12 The largest Finnish discussion board, *Sipulikanava*, in the Tor network. The site was created in 2014 by Kim Holviala and seised by Finnish authorities in 2017. The site contained heterogeneous conversations and a wide range of opinions ranging from well-written technical information to death threats.

Mr. Holviala moderated messages after they were published. He tried to keep the quality of the discussion as high as possible while allowing people to express painfully controversial views, such as their self-destructive behavior.

He described that surprisingly drug users were sharing accurate health information to each others and often honestly expressed harms too. Many of the conversations were about health issues and safety.

Various topic ranging from technical details to suicide were covered in his discussion board. Anonymity was a key feature to produce this straight forward discussion. Mr. Holviala told that he would like to see transparent conversation in our society and sees online anonymity as a tool to enable it.

4 UNDERSTANDING THE RESULTS

Let us take a final look at the results and reach conclusions. As shown, the results imply that there are significant criminal activities in the Tor network. We contribute to the body of knowledge on the domain of how people use onion websites and demonstrated empirical experiments to study criminal activities in the Tor network, including abilities to monitor anonymous online drug trade in real-time. Figure 4.1 organises the scope of this chapter.

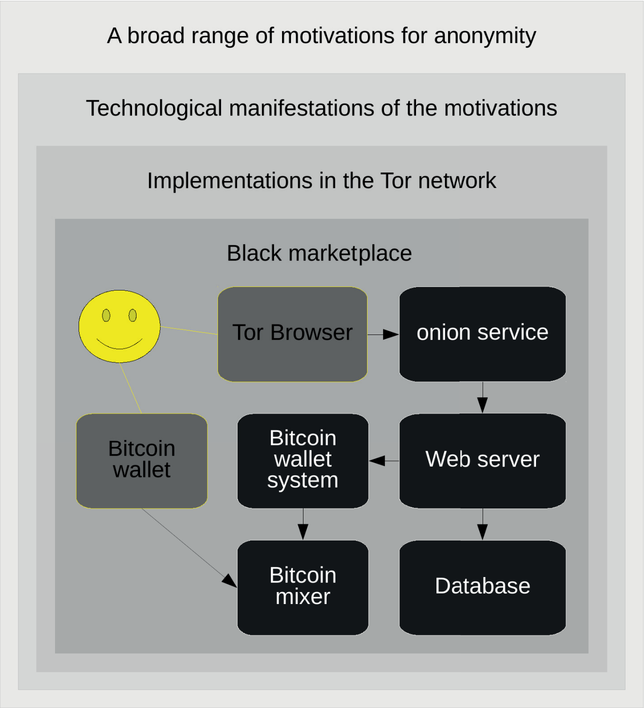

A broad range of motivations for anonymity
Technological manifestations of the motivations
Implementations in the Tor network
Black marketplace
Tor Browser
onion service
Bitcoin wallet
Bitcoin wallet system
Web server
Bitcoin mixer
Database

Figure 4.1 Motivations of anonymity manifest themselves in the Tor network: criminal activity is concentrated to black marketplaces. Here a seller or a buyer is connected to a black marketplace. Marketplaces have a certain general structure and common functional logic. We describe an empirical and technical process to research these marketplaces.

The main three contributions of this chapter to satisfy **Objective III** are as follows.

Firstly, we will merge the background knowledge and our results to examine the online anonymity as a whole. How people use Tor and what we have discovered from the usage? Why criminals as well find these technical tools feasible? Moreover, what are the real world security limitations of the Tor network?

Secondly, what is the technical structure of a marketplace, and what is the criminal logic behind it? How are Bitcoin and Tor transforming illegal drug trade?

Thirdly, the final section will be an established empirical process to follow illegal trade in the Tor network. What is the technical and empirical process of research in this field? The key result of this dissertation is the empirical process to study criminal activities in the Tor network.

4.1 Motivations and limitations of anonymity

We combined the background knowledge and our results. There are, of course, a broad range of motivations for online anonymity. The pressure for online anonymity is an implication of the Digital Revolution: services are shifting to online world.

For example, there is pressure towards online electronic voting, but privacy and auditability are open problems. A vote must be anonymous and not connected to a voter.

Electronic payments are problematic due to an asymmetric demand for anonymity: on the other hand governments demand tax accounting and audit for companies and anonymous money transactions create a vacuum; still consumers want the ability to buy products anonymous without identification, for instance, online adult entertainment videos. As a result, software projects, such as GNU Taler `https://www.taler.net/en/`, seek solutions to the asymmetric demand for anonymity: a seller's revenue is transparent for tax collection authorities while a payer is anonymous and obtains a legally valid proof of a payment.

Online anonymity is related to journalism, protection of sources and freedom of speech. For instance, global surveillance disclosures, in 2013, by Edward Snowden have these dimensions: he contacted journalists under the shield of online anonymity and revealed top-secret online surveillance programs; then journalists utilised extensive measures to secure their ability to publish the information.

Criminals have always tried to hide their identities. Currently, they adapt to digitalization and utilise online anonymity networks. In the digital form, illegal anonymous activities range from hacking to online drug trade.

We contributed to the current understanding of motivations by researching anonymous hacktivism attack motivations. In publication V we studied thirty-three hacktivism attack campaigns in manifesto level and then analysed the attacks and targets. Accordingly, we categorised these motivations under political, economic, socio-cultural, technological, environmental and legal reasons. The framework is known as PESTEL.

The PESTEL analysis framework is utilised to analyse and monitor the macro-environmental factors to organisations. In addition, PESTEL has been applied to investigate macro-economic and social trends from online data sources in order to identify and monitor early indicators of security threats [100] and *The United Nations Office on Drugs and Crime The SOCTA Handbook - Guidance on the preparation and use of serious and organised crime threat assessments* recommends PESTEL analysis for criminal activities [101].

The motivations behind hacktivism operations are categorised under PESTEL (in Figure 4.2). The categories are political, economic, socio-cultural, technological, environmental and legal.

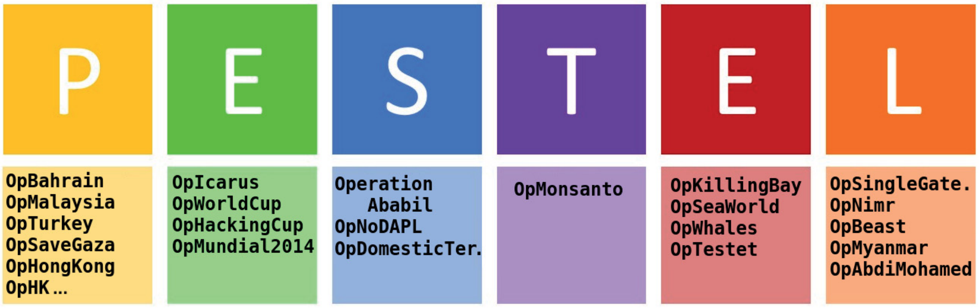


Figure 4.2 Motivations of anonymous hacktivism. Hacktivist groups publish manifestos which declare the intentions and motives of the their attacks. They call these attacks operations. We validated the operations and the targets that were selected according to these motivations. Attacks followed the stated motives: online services of industries, organization and governments were attacked accordingly. From publication V.

Most of the attacks were executed by hacktivist group *Anonymous*. It is a decentralised international group and the members of the group are participating in DDOS cyber attacks. This demonstrates how our digital world enables people with similar motivations to organise and coordinate their operations while still maintaining their anonymity. Similarly, anonymous marketplaces enable people to coordinate their motivations without the need to know each other.

The Tor Project provides means of online anonymity for everyone. The Tor network enables a free open network to various motivations of anonymity. For many, there is a critical demand for secure methods to communicate and access information. The Tor network has limitations. Numerous cases demonstrate how Tor users have lost their online anonymity and we analysed these realistic de-anonymisation techniques.

The hands-on demonstrations of anonymity exposures can be classified into four groups. The classes are 1. traffic correlation attacks, 2. electronic fingerprinting, 3. operational security failures, and 4. remote code executions. Both users and the Tor Project have separate responsibilities to mitigate these dangers.

I It is the responsibility of the Tor Project to deploy secure software.

II It is the responsibility of the Tor Project to publish clear tutorial material.

III It is the responsibility of the user to read and follow these guidelines.

IV It is the responsibility of the user to understand basic OPSEC principles.

We clarified the responsibilities of users and the Tor Project and categorised security aspects between technical and non-technical (in Figure 4.3).

Providing free and user-friendly anonymity solutions on the Internet today is an ongoing challenge for the Tor Project. Fortunately for Tor's anonymity, the network has a huge range of various users and according to current knowledge even the largest intelligence services are unable to spy Tor users.

The Tor Project offers extensive tutorials how to deploy safe installations but in every case it is hard to say what is safe and what is not safe. It is feasible to use most of the Internet applications behind the Tor network, however, unintentional configurations of the applications which are not designed to provide anonymity have revealed IP addresses.

The weakest link of anonymity is the actor leaking unintended information. In many cases, even technically talented users have failed to follow OPSEC principles and directly leaked their real identities.

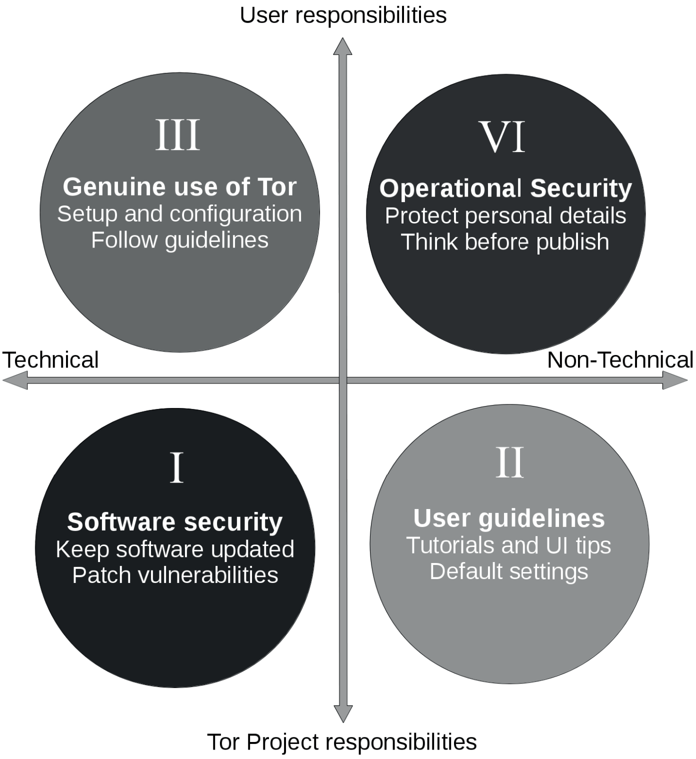


Figure 4.3 In publication IV we investigated real world dangers to Tor user's anonymity and analysed responsibilities. The Tor Project must deploy secure software and publish clear tutorials; the user must follow the guidelines and have a basic OPSEC understanding.

4.2 Anatomy of the marketplace

To shed light to drug trade and the structure of the marketplace we interviewed a former drug lord, current prisoner, Lasse Kärkkäinen, who led a drug syndicate and was planning his own black marketplace. We gained insight of the operations of drug dealers who sell their products through anonymous online markets.

First, this knowledge, combined with the background information, implies that the general components of a marketplace in the Tor network are the following: one or more Tor client software providing onion service, possible proxy and firewall solutions to sandbox the system, a web front-end server, a database server, a backup service, a Bitcoin wallet system, a Bitcoin transaction mixer, an escrow system, a user management system, a reputation and feedback system, a system that notifies the seller when a product is bought, and internal message system.

Not every marketplace implements every component; nay, for example, there are even very minimal proof-of-concept open-source marketplace implementations available, such as Python and Flask based *Laffka* `https://github.com/eruina/laffka`. *Laffka* provides a minimal web shop functionality to sellers who want to operate their own web shops.

The second part of illegal trade are the activities of sellers. We concluded that although the seller's operations are close to normal drug trade without digital aspects there are new demands for sellers: an established brand name is needed, customers demand a good reputation and expect a high quality customer service.

Thirdly, a buyer's role is simple. A buyer buys bitcoins, creates an account to the marketplace, transfers bitcoins to the wallet of the account, selects a product and provides shipment information to the seller. The escrow system locks the sum of the payment from the buyer's wallet. Then the seller sends the product and the buyer gives public feedback to the seller. After this step the escrow system transfers the payment to the seller.

4.3 Technical research methods to understand marketplaces

In Chapter 2 we created a categorization of the current technical research methods. In Table 3 we analysed how these methods could provide data regarding marketplaces in the Tor network.

Technical methodology	**Data on marketplace in the Tor network**
1. Monitor exit relay	Exit relay traffic does not reveal market data. Only aggregated data can be legally collected.
2. Exploit protocol weakness	Unfixed protocol weaknesses have revealed access rate of onion services, including marketplaces.
3. Passive traffic correlation	Correlation of traffic patterns between endpoints may reveal limited data.
4. Active traffic correlation	A shaped traffic pattern and traffic monitoring may reveal limited user information.
5. Port scan onion services	Port scanning could reveal if marketplaces have administration services through onion address.
6. Official Tor metrics	Official statistics provide comparison data. For instance, total number of Tor users per country.
7. Research honeypots	Honeypot services could collect data regarding marketplace users. For instance, an external link in user profile gathers activity information.
8. Crawl onion websites	Web crawling and scraping techniques are powerful tools to capture snapshots of the transparent trade on the marketplaces.

Table 3. We created a categorization of the earlier research methods and show how these methods provide data regarding marketplaces in the Tor network.

The market website itself provides a transparent view to illegal trade. In addition, most of the crypto currencies have a public blockchain that can be studied. In Figure 4.4 there is an overview of direct data sources which can be utilised to understand illegal online trade.

Most of the known Tor usage investigation methods (in Table 3) are not very effective to gather data on illegal markets in the Tor network. At the moment there are a few straight forward technical methodologies to gather quantitative research data. The main technical marketplace research methods are categorised here to five

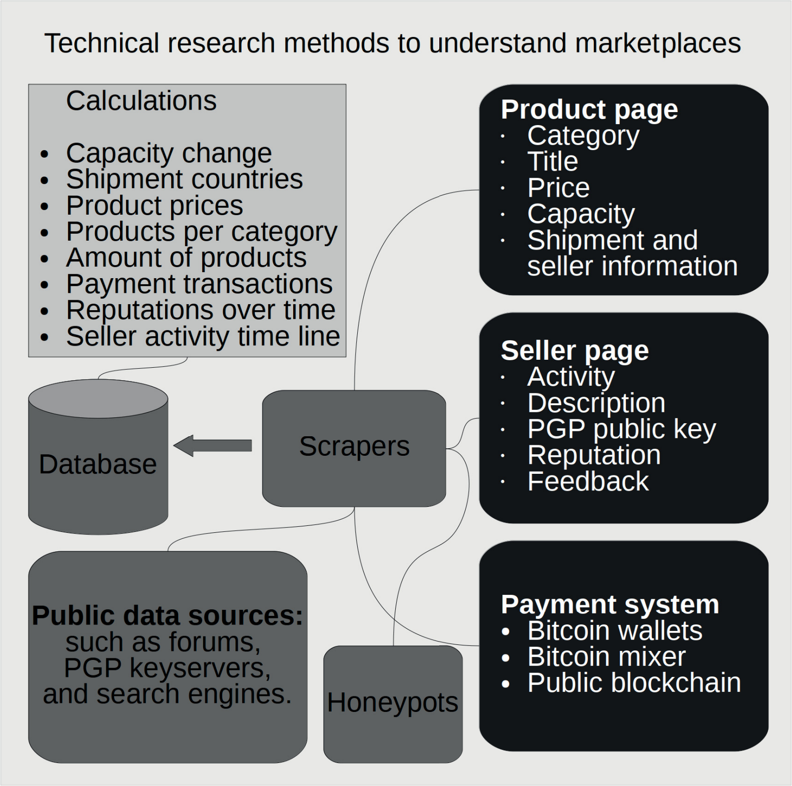


Figure 4.4 Available trading information and data collection process for anonymous marketplace investigations.

groups.

I **Follow the money**: scrape transactions from the public Bitcoin blockchain.

II **Follow the reputation of the sellers**: scrape feedbacks and calculate statistics.

III **Follow products available and capacities**: reveal metrics of the trade.

IV **Connect the dots**: combine public information to follow the seller's operations.

V **Setup honeypots**: for instance, gather data by creating a fake seller account.

In leverage of these five methods it is practical to gather quantitative data. Bitcoin provides unregulated transactions between buyers and sellers which are difficult to trace, however, in many cases, the amount of trade can be estimated [55]. The data enables empirical research of criminal activities of the various marketplaces in the Tor network.

5 CONCLUSIONS

We demonstrated how to investigate anonymous Tor usage. With the concrete case studies we showed how to gather data and research criminal activities.

In publication I we showed how Finnish illegal online drug trade concentrated to Silkkitie and discussed the rise of drug marketplaces on Tor. We predicted then that online drug markets are growing and the law enforcement is unable to prevent this shift. Now, four years later, Silkkitie is still operating and it is one of the several large online marketplaces on Tor. In fact, Bitcoin and Tor are transforming criminal activity, including drug trade. In the future societies need to resolve how to reduce the harms of drug use using education, illness and injury prevention, effective treatment, and by technical measures.

In addition, in publications I and III we carried out quantitative research to estimate the drug trade on Silkkitie. In 2015, the sales totalled over two million euros between Finnish buyers and sellers. In spite of anonymity, these methods reveal very precise data on Tor usage and a new way to measure drug trade.

Publication III shows how seller's reputation and capacity are both associated with drug sales. This means that the feedback system of the marketplace forces seller to maintain excellent reputation. Brand name is the most valuable asset of the seller. A reputation system is an important part of the illegal anonymous trade.

Publications II and IV show the limitations of privacy in the Tor network. In publication II, we used onion service honeypots to observe how onion name directories were spied by unknown adversaries. In publication IV we showed how it is an ongoing challenge for Tor to provide practical online anonymity. Fortunately, two million client users, 60 000 onion services, and 7000 voluntary Tor relays create a massive degree of anonymity. Still, many criminals are arrested because operation security is difficult.

Another notable anonymous criminal phenomenon is hacktivism. In publication V we studied what motivates anonymous hacktivist groups. We showed how they carry out their attacks according to their public manifests which are widely motivated by political reasons. Clear patterns of their attacks indicate that the targets could be predicted before the attacks.

Finally, we contribute to the body of knowledge by reviewing and categorising research methodologies. In spite of the private and anonymous nature of the Tor network there are several methods to gather data. We hope that the future research of anonymous online behaviour finds these summaries and descriptions of methodologies valuable.

6 DISCUSSION

The preface of this study came from Roman emperor Marcus Aurelius. In same spirit, the final words are coming from another gigantic Roman stoic philosopher, Seneca. The richness and subtleness of Seneca's satire of the human condition seems to fit in multiple situations, even to scientific context. He wrote in Epistulae Morales (65 AD) *"It is easier to understand parts, than to understand the whole"*.

We grasped the surface of understanding the usage of onion services in the Tor anonymity network. We explained how to measure and determine a few phenomenon in the Tor anonymity network. It is better to do a little with certainty and this is only a tiny part of Tor usage and future research will explain and measure it more precisely. There may be more unknown outside the boundaries of the current research of anonymity.

Online anonymity is here to stay. Although, the Tor network is the most popular, it is not the only option and there will be online anonymity in the future with or without the Tor network. Practical online anonymity can be obtained with leverage of several anonymity networks, such as I2P, Freenet, GNUNet, and a few more; English Wikipedia's category of *Anonymous file sharing networks* lists 24 different anonymity networks [102].

The digitalization shift of the drug markets is very difficult to tackle and it is a growing trend in the future. All services in the society are digitalizing themselves, including illegal activities, and law enforcement is unable to prevent this shift. Certainly there will be enormous problems with the Digital Revolution.

These enormous troubles will be discussed outside of the scope of technologies. Obviously illegal drugs are harmful so how could we reduce harmful impacts of this inevitable shift?

Anonymity causes special issues. Privacy is a part of security and this makes anonymity a part of safety. Take a look at the following example of how we could implement novel anonymous safety measures.

For instance, how to verify the age of whom you communicate anonymously? In several contexts this would be one of those useful technical features which could be solved with cryptography. A seller could demand an anonymous age verification. A buyer could send an anonymous cryptographic age verification. In addition, this could be a useful mechanism for adult entertainment websites as well: a visitor could prove the age without identification.

This particular problem could be solved with cryptography. Similar technical solutions could be studied and developed to mitigate problems in society during the Digital Age.

6.1 Author's final words

After wide scale information manipulation in leverage of targeted content, including fake news and advertising, and attempts to consume user's attention and time as commodity, new opportunities rise for online anonymity - extended to online solitude (Figure 6.1).

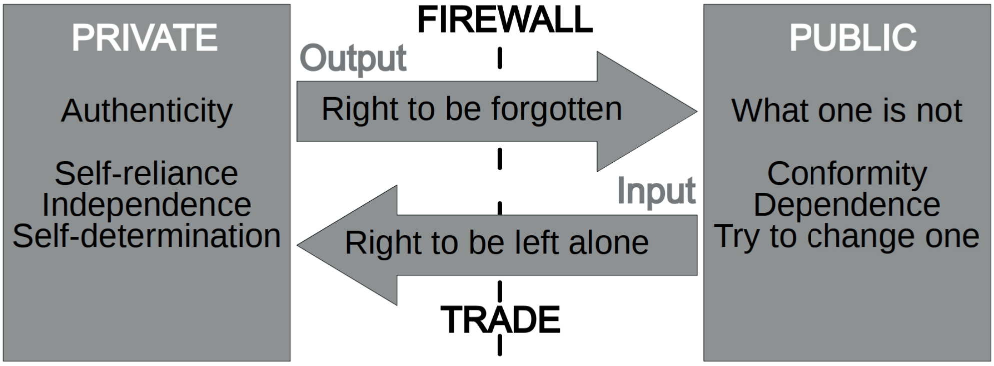


Figure 6.1 Online anonymity can be extended to online solitude. Internet user has full control over receiving or not receiving information and what information he is sharing. During the digital age a smart personal information firewall is needed to protect authenticity and values.

There is a demand for a smart personal information firewall that enables full control over outgoing and incoming information. People have a **right to be forgotten** and a **right to be left alone**. They can select profiles from pseudonyms to full anonymity and prevent targeted content and information manipulation according to their values.

This kind of system would require explicit permission from the user, for example, “Receive election manipulation and fake news” or “Personalise the news to reflect my views”. Not many would like to give the permission if online services would ask them directly.

In addition, time is valuable. We do not want to waste our time because we are eventually going to die. Neither we want to give away our privacy. Eventually economy itself is always a reflection of a certain value system. If we are not setting a cash value to our valuable assets, such as our time or concentration, we will loose them. When corporations try to monetise our privacy and time we notice the abysmal feeling that life is slipping through our fingers.

A smart system could give a transparent price for these valuable assets. A user could set an exact price for a certain labor and information. For instance, a Facebook user could demand hourly salary for using the service in exchange for private information, labor time and attention. The author would be ready to create a Facebook account if Facebook would be ready to pay him two thousand U.S. dollars per hour for clicking labor.

REFERENCES

[1] J. Nurmi and T. Kaskela. Silkkitie. Päihteiden suomalaista nappikauppaa. *Yhteiskuntapolitiikka* 80.4 (2015), 387–394.

[2] J. Nurmi, J. Kannisto and M. Vajaranta. Observing Hidden Service Directory Spying with a Private Hidden Service Honeynet. *11th Asia Joint Conference on Information Security (AsiaJCIS)*. 2016, 55–59.

[3] J. Nurmi, T. Kaskela, J. Perälä and A. Oksanen. Seller's reputation and capacity on the illicit drug markets: 11-month study on the Finnish version of the Silk Road. *Drug & Alcohol Dependence* 178 (2017), 201–207.

[4] J. Nurmi and M. S. Niemelä. Tor De-anonymisation Techniques. *11th International Conference on Network and System Security*. 2017, 657–671.

[5] J. Nurmi and M. S. Niemelä. PESTEL Analysis of Hacktivism Campaign Motivations. *Nordic Conference on Secure IT Systems 2018*. 2018, 323–335.

[6] J. Gantz and D. Reinsel. The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the future* 2007.2012 (2012), 1–16.

[7] S. M. Bellovin. Security problems in the TCP/IP protocol suite. *ACM SIGCOMM Computer Communication Review* 19.2 (1989), 32–48.

[8] C. Diaz. Anonymity and privacy in electronic services. PhD thesis. Katholieke Universiteit Leuven, 2005.

[9] G. Danezis. Better anonymous communications. PhD thesis. University of Cambridge, 2004.

[10] R. Dingledine. The free haven project: Design and deployment of an anonymous secure data haven. Master's thesis. Massachusetts Institute of Technology, 2000.

[11] I. Goldberg. A pseudonymous communications infrastructure for the Internet. PhD thesis. University of California, Berkeley, 2000.

[12] A. Serjantov. On the anonymity of anonymity systems. PhD thesis. University of Cambridge, 2004.

[13] K. Chatzikokolakis. Probabilistic and information-theoretic approaches to anonymity. PhD thesis. Ecole Polytechnique, 2007.

[14] S. J. Murdoch. Covert channel vulnerabilities in anonymity systems. PhD thesis. University of Cambridge, 2007.

[15] D. J. Kelly. A taxonomy for and analysis of anonymous communications networks. PhD thesis. Air Force Institute of Technology, 2009.

[16] K. Loesing. Privacy-enhancing Technologies for Private Services. PhD thesis. University of Bamberg Press, 2009.

[17] K. S. Bauer. Improving Security and Performance in Low Latency Anonymous Networks. PhD thesis. University of Colorado at Boulder, 2011.

[18] R. G. Jansen. Privacy preserving performance enhancements for anonymous communication networks. PhD thesis. University of Minnesota, 2012.

[19] D. M. Goldschlag, M. G. Reed and P. F. Syverson. Hiding routing information. *International Workshop on Information Hiding*. Springer. 1996, 137–150.

[20] R. Dingledine, N. Mathewson and P. F. Syverson. Tor: The Second-Generation Onion Router. *Proceedings of the 13th USENIX Security Symposium*. 2004, 303–320.

[21] Technical and legal overview of the Tor anonymity network. *NATO Cooperative Cyber Defence Centre of Excellence.* (2015).

[22] M. J. Barratt. Silk Road: EBAY FOR DRUGS. *Addiction* 107.3 (2012), 683–683.

[23] M. C. Van Hout and T. Bingham. Surfing the Silk Road: A study of users' experiences. *International Journal of Drug Policy* 24.6 (2013), 524–529.

[24] J. Martin. Lost on the Silk Road: Online drug distribution and the 'cryptomarket'. *Criminology & Criminal Justice* 14.3 (2014), 351–367.

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

[26] M. C. Van Hout and T. Bingham. 'Silk Road', the virtual drug marketplace: A single case study of user experiences. *International Journal of Drug Policy* 24.5 (2013), 385–391.

[27] G. E. R. Lloyd. *Early greek science: Thales to Aristotle*. Random House, 2012.

[28] I. Düring. *Aristotle in the ancient biographical tradition*. Garland Publishing, 1957.

[29] B. Gower. *Scientific method: A historical and philosophical introduction*. Routledge, 2012.

[30] E. Zalta. *The Stanford Encyclopedia of Philosophy: Newton's Philosophiae Naturalis Principia Mathematica*. Accessed: 2018-06-13. 2007. URL: `https://plato.stanford.edu/entries/newton-principia/`.

[31] E. Zalta. *The Stanford Encyclopedia of Philosophy: Scientific Method*. Accessed: 2018-06-12. 2015. URL: `https://plato.stanford.edu/entries/scientific-method/`.

[32] The Tor Project. *Tor Research Safety Board.* Accessed: 2018-06-12. 2016. URL: `https://research.torproject.org/safetyboard.html`.

[33] EU law - EUR-Lex. *GDPR: General Data Protection Regulation.* Accessed: 2018-06-15. 2016. URL: `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679`.

[34] M. Castells. The Information Age. *Media Studies: A Reader* 2.7 (2010). NYU Press, 152.

[35] DARPA. *ARPANET and the Origins of the Internet.* Accessed: 2018-06-18. URL: `https://www.darpa.mil/about-us/timeline/arpanet`.

[36] DARPA. *The Mother of all Demos.* Accessed: 2018-06-18. URL: `https://www.darpa.mil/about-us/timeline/the-mother-of-all-demos`.

[37] J. F. Heafner and E. F. Harslem. Large-Scale Sharing of Computer Resources. *Rand Corporation* (1972).

[38] P. Elmer-Dewitt. Twenty million strong and adding a million new users a month, the Internet is suddenly the place to be. *TIME International* 49 (1993).

[39] The Defense Data Network. *DDN Newsletter No. 26.* Accessed: 2018-06-19. 1983. URL: `https://www.rfc-editor.org/rfc/museum/ddn-news/ddn-news.n26.1`.

[40] S. Feinstein. *The 1960s*. Enslow Publishing, LLC, 2015.

[41] The Defense Data Network. *ARPANET newsletter 1 July.* Accessed: 2018-06-19. 1980. URL: `https://www.rfc-editor.org/rfc/museum/ddn-news/ddn-news.n1.1`.

[42] W. J. Croft and J. Gilmore. Bootstrap Protocol. *RFC* (1985), 1–12.

[43] The Electronic Frontier Foundation. *A History of Protecting Freedom Where Law and Technology Collide.* Accessed: 2018-06-20. 2007. URL: `https://www.eff.org/about/history`.

[44] J. P. Barlow. The economy of ideas. *Wired* (1994).

[45] J. P. Barlow. A Declaration of the Independence of Cyberspace. *The Humanist* 56.3 (1996), 18.

[46] S. Nakamoto. *RE: Bitcoin P2P e-cash paper.* Accessed: 2018-06-20. 2008. URL: `https://www.mail-archive.com/cryptography@metzdowd.com/msg09971.html`.

[47] L. Wittgenstein. *Philosophical investigations*. John Wiley & Sons, 2009.

[48] M. Tribus and E. C. McIrvine. Energy and information. *Scientific American* 225.3 (1971), 179–190.

[49] D. Goldschlag, M. Reed and P. Syverson. Onion routing. *Communications of the ACM* 42.2 (1999), 39–41.

[50] Y. Gilad and A. Herzberg. Spying in the Dark: TCP and Tor Traffic Analysis. *International symposium on privacy enhancing technologies symposium*. Springer. 2012, 100–119.

[51] P. Winter, T. Pulls and J. Fuss. *ScrambleSuit: A Polymorph Network Protocol to Circumvent Censorship*. 2013. URL: `http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-29031`.

[52] B. Schneier. Attacking Tor: how the NSA targets users' online anonymity. 4 (2013). Published by The Guardian.

[53] S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system. (2008).

[54] R. Grinberg. Bitcoin: An innovative alternative digital currency. *Hastings Sci. & Tech. LJ* 4 (2012), 159.

[55] S. Foley, J. R. Karlsen and T. J. Putniii. Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed Through Cryptocurrencies?: *Review of Financial Studies, Forthcoming* (Jan. 2018).

[56] D. McCoy, K. Bauer, D. Grunwald, T. Kohno and D. Sicker. Shining light in dark places: Understanding the Tor network. *International Symposium on Privacy Enhancing Technologies Symposium*. Springer. 2008, 63–76.

[57] A. Biryukov and I. Pustogarov. Bitcoin over Tor isn't a good idea. *Security and Privacy*. IEEE. 2015, 122–134.

[58] K. Bauer, D. McCoy, D. Grunwald, T. Kohno and D. Sicker. Low-resource routing attacks against Tor. *Proceedings of the 2007 ACM workshop on Privacy in electronic society*. ACM. 2007, 11–20.

[59] A. Biryukov, I. Pustogarov and R.-P. Weinmann. TorScan: Tracing long-lived connections and differential scanning attacks. *European Symposium on Research in Computer Security*. Springer. 2012, 469–486.

[60] G. Danezis. The traffic analysis of continuous-time mixes. *International Workshop on Privacy Enhancing Technologies*. Springer. 2004, 35–50.

[61] A. Serjantov and P. Sewell. Passive attack analysis for connection-based anonymity systems. *European Symposium on Research in Computer Security*. Springer. 2003, 116–131.

[62] A. Back, U. Möller and A. Stiglic. Traffic analysis attacks and trade-offs in anonymity providing systems. *International Workshop on Information Hiding*. Springer. 2001, 245–257.

[63] B. N. Levine, M. K. Reiter, C. Wang and M. Wright. Timing attacks in low-latency mix systems. *International Conference on Financial Cryptography*. Springer. 2004, 251–265.

[64] G. D. Bissias, M. Liberatore, D. Jensen and B. N. Levine. Privacy vulnerabilities in encrypted HTTP streams. *International Workshop on Privacy Enhancing Technologies*. Springer. 2005, 1–11.

[65] Y. Zhu, X. Fu, B. Graham, R. Bettati and W. Zhao. On flow correlation attacks and countermeasures in mix networks. *International Workshop on Privacy Enhancing Technologies*. Springer. 2004, 207–225.

[66] A. Panchenko, L. Niessen, A. Zinnen and T. Engel. Website fingerprinting in onion routing based anonymization networks. *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*. ACM. 2011, 103–114.

[67] S. J. Murdoch and G. Danezis. Low-cost traffic analysis of Tor. *Security and Privacy*. IEEE. 2005, 183–195.

[68] W. Yu, X. Fu, S. Graham, D. Xuan and W. Zhao. DSSS-based flow marking technique for invisible traceback. *Security and Privacy*. IEEE. 2007, 18–32.

[69] X. Wang and D. S. Reeves. Robust correlation of encrypted attack traffic through stepping stones by manipulation of interpacket delays. *Proceedings of the 10th ACM conference on Computer and communications security*. ACM. 2003, 20–29.

[70] X. Wang, S. Chen and S. Jajodia. Network flow watermarking attack on low-latency anonymous communication systems. *Symposium on Security and Privacy*. IEEE. 2007, 116–130.

[71] G. Owen and N. Savage. Empirical analysis of Tor hidden services. *IET Information Security* 10.3 (2016), 113–118.

[72] R. R. Rohrmann. Large Scale Anonymous Port Scanning. *University of Arizona* (2017).

[73] The Tor Project. *Tor Project Metrics.* Accessed: 2019-03-25. URL: `https://metrics.torproject.org`.

[74] S. Chakravarty, G. Portokalidis, M. Polychronakis and A. D. Keromytis. Detecting traffic snooping in Tor using decoys. *International Workshop on Recent Advances in Intrusion Detection*. Springer. 2011, 222–241.

[75] D. S. Dolliver. Evaluating drug trafficking on the Tor Network: Silk Road 2, the sequel. *International Journal of Drug Policy* 26.11 (2015), 1113–1123.

[76] M. C. Van Hout and T. Bingham. Responsible vendors, intelligent consumers: Silk Road, the online revolution in drug trading. *International Journal of Drug Policy* 25.2 (2014), 183–189.

[77] A. Celestini, G. Me and M. Mignone. Tor Marketplaces Exploratory Data Analysis: The Drugs Case. *International Conference on Global Security, Safety, and Sustainability*. Springer. 2017, 218–229.

[78] A. T. Zulkarnine, R. Frank, B. Monk, J. Mitchell and G. Davies. Surfacing collaborated networks in dark web to find illicit and criminal content. *Intelligence and Security Informatics (ISI)*. IEEE. 2016, 109–114.

[79] R. A. Hardy and J. R. Norgaard. Reputation in the Internet black market: an empirical and theoretical analysis of the Deep Web. *Journal of Institutional Economics* 12.3 (2016), 515–539.

[80] J. Bearman and T. Hanuka. The Rise and Fall of Silk Road,(Part 1): Ross Ulbricht's journey from libertarian idealist to savage kingpin. *Wired* 23.5 (2015), 90–97.

[81] S. Ackerman and D. Roberts. *NSA phone surveillance program likely unconstitutional, federal judge rules*. Accessed: 2018-08-01. 2013. URL: `https://www.theguardian.com/world/2013/dec/16/nsa-phone-surveillance-likely-unconstitutional-judge`.

[82] The NSA. *Advanced Open Source Multi-hop, top-secret presentation*. Accessed: 2018-08-01. 2013. URL: `https://edwardsnowden.com/wp-content/uploads/2013/10/tor-the-king-of-high-secure-low-latency-anonymity.pdf`.

[83] The NSA. *Tor Stinks, top-secret presentation*. Accessed: 2018-08-01. 2013. URL: `https://edwardsnowden.com/docs/doc/tor-stinks-presentation.pdf`.

[84] Ross Ulbricht. Bitcoin Forum. *Ross Ulbricht's message.* Accessed: 2018-08-02. 2011. URL: `https://bitcointalk.org/index.php?topic=47811.msg568744`.

[85] The Federal Bureau of Investigation. *Narcotic Trafficking Conspiracy. Sealed Complaint against Ross Ulbricht.* Accessed: 2018-08-02. 2014. URL: `https://www.documentcloud.org/documents/801103-172770276-ulbricht-criminal-complaint.html`.

[86] Attorneys For Plaintiff. *The United States district court complaint, case Alexandre Cazes*. Accessed: 2018-08-02. 2017. URL: `https://www.justice.gov/opa/press-release/file/982821/download`.

[87] Mozilla Foundation Security Advisory 2013-53. *Execution of unmapped memory through onreadystatechange event.* Accessed: 2018-08-03. 2013. URL: `https://www.mozilla.org/en-US/security/advisories/mfsa2013-53/`.

[88] The Federal Bureau of Investigation. *Affidavit Case 3:15-cr-05351-RJB Document 166-2. Playpen website exploit.* Accessed: 2018-08-02. 2016. URL: `https://regmedia.co.uk/2016/03/29/alfin.pdf`.

[89] Tor-talk mailing list. *JavaScript exploit.* Accessed: 2018-08-03. 2016. URL: `https://lists.torproject.org/pipermail/tor-talk/2016-November/042639.html`.

[90] Tails. *Tails Operating system - Privacy for anyone anywhere.* Accessed: 2018-08-10. URL: `https://tails.boum.org/`.

[91] Whonix. *Stay anonymous with Whonix Operating system.* Accessed: 2018-08-10. URL: `https://www.whonix.org/`.

[92] VLC - Ticket system. *VLC media player privacy leak due to –no-metadata-network-access not being respected.* Accessed: 2018-08-03. 2016. URL: `https://trac.videolan.org/vlc/ticket/17760`.

[93] Published by Der Spiegel. *The NSA TEMPORA documentation.* Accessed: 2018-08-10. 2013. URL: `http://www.spiegel.de/media/media-34103.pdf`.

[94] Naked Security. *Use of Tor pointed FBI to Harvard University bomb hoax suspect.* Accessed: 2018-08-03. 2013. URL: `https://nakedsecurity.sophos.com/2013/12/20/use-of-tor-pointed-fbi-to-harvard-university-bomb-hoax-suspect/`.

[95] A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. *International Workshop on Privacy Enhancing Technologies*. Springer. 2002, pp. 41–53.

[96] C. Diaz, S. Seys, J. Claessens and B. Preneel. Towards measuring anonymity. *International Workshop on Privacy Enhancing Technologies*. Springer. 2002, pp. 54–68.

[97] C. Dougherty. *Prometheus*. Routledge, 2006.

[98] J. W. Goethe. *Faust*. 1880.

[99] Judgment: defendant Kim Hjalmar Holviala. *Eastern Uusimaa District Court, Finland, R 18/3115/766, 19/106044* (2019).

[100] J. Gomez-Romero, M. D. Ruiz and M. J. Martin-Bautista. Open data analysis for environmental scanning in security-oriented strategic analysis. *Information Fusion (FUSION)*. IEEE. 2016, 91–97.

[101] UN. *The SOCTA Handbook — Guidance on the preparation and use of serious and organized crime threat*. United Nations Office on Drugs and Crime, 2010.

[102] Wikipedia, English, The Free Encyclopedia. *Anonymous file sharing networks*. Accessed: 2018-08-14. 2018. URL: `http://goo.gl/a0pGBv`.

APPENDIX: ENGLISH SUMMARY OF THE FIRST PUBLICATION

Nurmi, J., Kaskela, T. (2015). "Silkkitie. Päihteiden suomalaista nappikauppaa." *Yhteiskuntapolitiikka*, vol 80, no. 4, pp. 387–394.

This publication is in Finnish because it was written directly to the Finnish audience about early unique drug marketplace inside Finland. On 2014 Finnish online drug trade largely shifted to one special black market on Tor. Indeed, this marketplace was in Finnish, sellers and buyers were Finns, shipments proceeded through domestic mail.

At first, since 2013, the marketplace was known as Silkkitie (Finnish word for Silk Road) and on 2015 English version of it began to operate under the name Valhalla. Since then Silkkitie/Valhalla has a veteran reputation of being one of the oldest running marketplaces on Tor.

Language and country centric marketplace was unique and one of the first marketplaces on Tor. Using web-scraping tools the author answered how much of different types of illegal drugs Finnish people are trading on Silkkitie and described the function of the marketplace. The author understood that publishing this study as soon as possible in Finnish will have a large impact.

Accordingly, author contacted Finnish A-Clinic Foundation's researcher, Mr. Teemu Kaskela, and they published a research paper regarding the rise of drug marketplaces on Tor. This paper demonstrated web-scraping techniques to carry out quantitative research and contains the current estimations of the drug trade on Silkkitie. Selected Finnish journal *Yhteiskuntapolitiikka* is considered to be a top-level scientific journal.

As a result, this early publication has impact even beyond technical aspects and quantitative data. The author concluded that the digitalization shift to online drug markets is growing trend and any law enforcement is unable to prevent this shift.

Obviously illegal drugs are harmful so the real question is how to reduce harmful impacts of this inevitable shift. A-Clinic Foundation began to prepare an urgent project: The harms of drug use should be reduced using education, illness and injury prevention, and effective treatment. Finally, Muunto project started in 2016 to develop novel methods to respond to novel psychoactive substances and changing drug cultures.

A separate English translation of the abstract of the article is written for this thesis for the English speaking audience.

Silkkitie: Online drug marketplace in Finland

Abstract

Recent Internet service development has shifted major services to online. Currently, drug markets are also available on the Internet. Moreover, recent technologies and the users of the Internet take a stand against mass surveillance, censorship and other human rights violations. Leveraging same techniques illegal drug markets are shifting towards recent encryption technologies. In Finland the central place to buy illegal drugs is Silkkitie. It operates using the Tor network to hide its real location.

Especially during the year 2014 illegal drug online trade grew inside the Finnish borders. In other words, sellers send the products using domestic post service and there is no risk for custom check. Tor, Bitcoin and high-class online marketplace offers seamless way to sell and buy illegal drugs. In practice, this makes it very difficult for the law enforcement to interfere with the illegal drug trade. On the other hand, because there is now one public place for Finnish to buy illegal drugs we can automatically construct statistics that describes the markets. We can gain understanding how Finnish people use different types of substances.

Illegal online drug market is very difficult to tackle and indicates signs of growing in the future. However, we can follow the extent of trading and understand the technologies. We can also produce nearly real-time picture of the situation of how people buy and sell illegal drugs. As a result, statistics shed light on substance consumption habits in Finland. This is a novel approach. Furthermore, the technologies and how people interact using them is very interesting to research. In this paper, we examine how Finnish online marketplace Silkkitie operates and what we can learn form the trade data. We study both technical and social aspects of this phenomenon.

PUBLICATIONS

PUBLICATION

I

Silkkitie. Päihteiden suomalaista nappikauppaa
J. Nurmi and T. Kaskela

Yhteiskuntapolitiikka 80.4 (2015), 387–394

Silkkitie

Päihteiden suomalaista nappikauppaa

JUHA NURMI & TEEMU KASKELA

Taustaa

Vuodesta 2011 alkaen vakoilua vastaan perustettuun turvallisen viestinnän Tor-verkkoon nousi muutamia englanninkielisiä laittomien päihteiden myyntiin keskittyviä kauppapaikkoja. Kuuluisin oli alkuperäinen Silk Road -niminen kauppapaikka, joka suljettiin myöhemmin. Näissä kauppapaikoissa laittomien päihteiden myyjät ja ostajat kohtaavat yli valtioiden rajojen. (Barratt 2012;Van Hout & al. 2013.) Vuonna 2014 perustettiin suomalainen Silkkitie-kauppapaikka, joka lainasi nimensä ja bisnesideansa alkuperäiseltä kauppapaikalta.

Palvelujen käyttäjistä tiedetään hyvin vähän. Sekä ostaminen että myyminen vaativat kuitenkin jonkinlaista teknistä osaamista. Etenkin itse palvelun toteuttaminen vaatii teknologioiden syvällistä ymmärtämistä. (Van Hout & al. 2014.) Esittelemme kolme tärkeintä salaukseen perustuvaa teknologiaa, joihin Silkkitie-kauppapaikan toiminta nojautuu. Nämä ovat Tor-anonyymiverkko, PGP-salaus ja bitcoin-virtuaalivaluutta.

Tässä kirjoituksessa 1) esitellään keskeisiä käsitteitä, 2) kuvataan suomalaisen Tor-verkossa toimivan Silkkitie-palvelun toimintaperiaatteita, 3) kerrotaan alustavia tuloksia myynnistä ja myytyjen tuotteiden hinnoista, 4) pohditaan, mitä Silkkitiestä kauppapaikkana voidaan päätellä nykyisen tiedon perusteella, ja 5) esitellään ajatuksia huumeiden verkkokauppaan liittyvistä jatkotutkimuksista.

Keskeisiä käsitteitä

Tor-verkko on tarkoitettu suojaamaan luottamuksellista viestintää tietoliikennevalvontaa vastaan, kiertämään sensuuria, tarjoamaan anonymiteettiä ja turvaamaan erilaisten palveluiden toiminta monenlaisia hyökkäyksiä vastaan (Dingledine & al. 2004). Tor on saanut paljon kiitosta ja huomiota, koska se on auttanut kansalaisaktivisteja, sensuuria kiertäviä totalitarististen maiden asukkaita, tietovuotajia, journalisteja ja diktatuurien demokratialiikkeitä turvaamaan toimintansa. Verkkoa käytetään myös arkipäiväisen verkon käytön suojaamiseen valvonnalta. Laajat Tor-verkon turvallisuutta kartoittavat julkaisut pitävät Tor-verkkoa korkealaatuisena tietoturvatyökaluna, eikä edes Yhdysvaltain kansallinen turvallisuusvirasto NSA ole pystynyt murtamaan suojausta (The Guardian 2013). Tor-verkossa voi ylläpitää piilopalveluina (*hidden service*) samoja palveluita kuin internetissä (Dingledine & al. 2004). Tor-verkossa toimiessaan laittomien päihteiden kauppaa on vaikea sulkea, sensuroida tai käyttäjiä paljastaa.

Perinteisten valuuttojen rinnalle ovat nousseet erilaiset verkossa toimivat virtuaalivaluutat, joista yleisemmin käytetty on *bitcoin* (Nakamoto 2008). Nämä valuutat ovat internetissä hajautetusti toimivia maksuvälineitä. Niiden arvo määräytyy kuten perinteisillä valuutoilla, kysynnän ja tarjonnan perusteella. Käteistä rahaa voi vaihtaa bitcoineiksi esimerkiksi LocalBitcoins-palvelun (LocalBitcoins 2012) kautta tai fyysisillä bitcoin-automaateilla, joita Suomestakin löytyy. Bitcoin-automaatista käteisellä ostettua virtuaalivaluuttaa on vaikea jäljittää. Erityisesti Tor-verkossa toimivat monenlaiset kauppapaikat ovat omaksuneet bitcoinin maksuvälineeksi, koska maksutapahtumia on käteisen rahan tavoin hyvin hankala seurata verrattuna perinteisten pankkitilitapahtumien seurantaan.

PGP, Pretty Good Privacy, on julkiseen ja yksityiseen avaimeen perustuva epäsymmetrinen salausjärjestelmä. Perusajatus on se, että on julkista avainta voi nimensä mukaisesti jakaa avoimesti ja tällä avaimella voidaan salakirjoittaa viestejä. Näiden viestien purkaminen onnistuu käytännössä vain yksityisellä avaimella. PGP-salausjärjestelmän kehitti Philip Zimmermann, joka halusi luoda ilmaisen ja tehokkaan tavan ihmisille suojata viestintäänsä vakoilua vastaan. (Zimmermann 1994.)

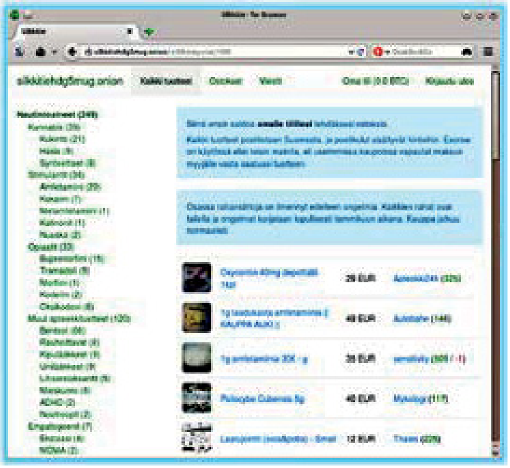
Yksinkertaistava analogia voisi olla, että on olemassa kuljetusarkku, jonka voi lukita kaikkien käytössä olevalla avainkopiolla, mutta avata vain yksityisen avaimen kanssa. Kuka tahansa voi lukita kuljetuksen arkkuun, mutta vain vastaanottajalla on sellainen avain, jolla lippaan saa auki.

Riittävää avainten pituutta käyttävää PGP-salausta pidetään yleisesti käytännössä murtumattomana, kunhan käyttäjä pitää yksityisen avaimensa turvassa (Thomas 2003).

Silkkitie

Silkkitie-sivusto perustettiin kuudes tammikuuta 2014. Se on julkinen verkkosivu, joka kertoo käyttäjälle osoitteen http://silkkitiehdg5mug.onion/ olemassaolosta (Kapteeni 2014). Tälle sivustolle pääsemiseksi vaaditaan Tor-ohjelmistoa käyttävä yksityisyydensuojatyökaluksi kehitetty Tor-selain (Tor Project).

Sivulla on kaupankäyntialusta, jonka kautta voi myydä ja ostaa tuotteita bitcoineilla. Suurin osa kauppatavarasta on laittomia päihteitä (kuvio 1). Palvelun perustajaksi on ilmoittautunut Kapteeni-nimellä esittäytyvä henkilö tai ryhmä. Tor-verkko kätkee palvelun oikean olinpaikan, ja sen sulkeminen tai sensurointi on hyvin hankalaa. On myös mahdoton valvoa, kuka sivustolla vierailee. Silkkitie-kauppapaikka ei myöskään näe sivustolla vierailevan henkilön IP-osoitetta. Sekä sivusto että sivustolla kävijä säilyttävät korkean anonymiteetin.



Kuvio 1. Silkkitie-kauppapaikka avattuna Tor-selaimessa. Palvelunäkymä kirjautumisen jälkeen. Ylhäällä oman bitcoin-tilin saldo, tehdyt ostokset ja viestintäjärjestelmä. Vasemmalla lista tarjolla olevista tuotteista.

Silkkitie-verkkopalvelua käyttääkseen täytyy ensin tehdä käyttäjätunnus luomalla nimimerkki ja salasana palveluun sekä erillinen maksusalasana tiliä varten, johon siirretään bitcoineja. Käyttäjätyyppejä on kahdenlaisia: myyjät ja käyttäjät (tässä ostajat). Vain myyjät voivat asettaa tuotteita myyntiin. Myyjä kuvaa tuotteensa sekä asettaa sille hinnan, kertoo myynnissä olevien tuotteiden varastomäärän ja toimitusehdot. Myyjien maine kehittyy kauppojen myötä, kun ostajat kertovat tuotteen perilletulosta painamalla vihreää (tuote perillä) tai punaista (kauppa epäonnistui) ja kommentoivat kirjallisesti saamiaan tuotteita. Toisaalta myös myyjä voi arvioida ostajan luotettavuutta hänen tekemiensä kauppojen määrän, niiden kokonaisarvon ja ongelmitta onnistuneiden kauppojen kokonaismäärän perusteella.

Myyjiltä vaaditaan toimiva PGP-salausjärjestelmän julkinen avain, jolla voidaan tehokkaasti salata viestintää. Ostajan osoite välitetään tällä julkisella avaimella salattuna myyjälle. Vain myyjällä on yksityinen avain, jolla viesti voidaan purkaa. Näin vältetään tilanne, että Silkkitie-palvelusta vuotaisi ostajien osoitetietoja ulkopuolisten tietoon.

Silkkitie on kauppapaikkana selvästi suunniteltu laittomaan kaupankäyntiin. Tästä kertovat Silkkitien käyttöehdot seuraavasti:

> Silkkitiellä kiellettyä on vain sivullisten vahingoittaminen. Kiellämme siis lapsipornon, luottokorttitiedot, väärennetyn rahan, väkivaltapalvelut ja räjähdysaineet. Sallittuja ovat muunmuassa kaikki nautintoaineet, aseet

itsepuolustus- tai harrastustarkoitukseen, rahanpesupalvelut ja reseptiväärennökset. Asiasta äänestettiin keskustelualueella. (Silkkitie 2015.)

Myyjän tilin luomiseksi täytyy maksaa 100 euron pantti. Pantin saa takaisin, kun myyjä on tehnyt vähintään kymmenen kauppaa 500 euron kokonaisarvosta. Kauppapaikan ansaintamalli on se, että Silkkitie ottaa jokaisesta kaupasta viiden prosentin osuuden. Rahaliikenne kulkee Silkkitie-palvelun kautta, ja käytössä on niin sanottu välitili. Käytännössä se tarkoittaa, että oston jälkeen rahat siirtyvät kolmannen osapuolen eli Silkkitien välitilille eikä suoraan myyjälle. Näin ostaja voi vapauttaa maksun myyjälle vasta sitten, kun on tyytyväinen kaupan ehtojen mukaiseen tapahtumaan, esimerkiksi saadessaan tuotteen. Myyjä voi asettaa ostotapahtumalle ehtoja. Erityisesti uusilta ostajilta myyjä usein vaatii maksun vapauttamista jo ennen tuotteen lähetystä. Sama koskee erityisen suuria tilauksia. Mikäli ostaja ei saa tuotetta tai esiintyy erimielisyyksiä, Silkkitien asiakaspalvelu ratkaisee tilanteen.

Silkkitie maksaa palvelun mainostamisesta. Tästä esimerkkejä ovat tarrojen levittäminen kaupungeissa, Silkkitien toiminnan jakaminen sosiaalisessa mediassa, juttuvinkkien jako lehdistölle ja jopa mainosgraffitin maalaaminen. Myös tietoturvahaavoittuvuuksien tiedottamisesta asiakaspalvelulle luvataan maksaa 1 000–5 000 euroa.

Kauppapaikassa on myös sisäinen viestintäjärjestelmä, jolla voi lähettää viestejä asiakaspalvelulle, ylläpidolle ja myyjille.

Testasimme Silkkitien toimintaa monella tavalla ja toteutus vaikuttaa hyvin suunnitellulta. Erityisesti palvelu on rakennuttu niin, ettei se ei vahingossakaan virhetilanteessa vuoda palvelun olinpaikkaa valottavia tietoja.

Vaikka Silkkitie suojelee tehokkaasti ostajien ja myyjien tietoja muun muassa salaamalla osoitteen luovutuksen myyjän julkisella avaimella, keskitetty kauppapaikka on edelleen monessa suhteessa riski. Kapteenin henkilöllisyys tai lopulliset motiivit eivät ole tiedossa. Maailmalla vastaavat palvelut ovat joutuneet yllättävien käänteiden kautta usein poliisiviranomaisten haltuun. Syitä on ollut monia: ylläpitohenkilöt ovat tehneet inhimillisiä virheitä, he ovat joutuneet poliisitutkinnan kohteeksi muiden rikosepäilyjen takia, palveluissa on ollut vakavia teknisiä toteutusvirheitä ja palveluihin on jopa soluttautunut työntekijöiksi Yhdysvaltain liittovaltion agentteja (Greenberg 2015).

Myös myyjille saattaa kertyä laaja lista heiltä tilanneiden ihmisten osoitteita. Mikäli myyjä joutuu esimerkiksi itse myyntitavaraa hankkiessaan poliisin haaviin, tällainen osoitelista saattaa löytyä kotietsinnässä ja todennäköisesti herättää poliisin huomion. Useat myyjät lupaavat tosin hävittää ostajien osoitteet välittömästi hallustaan.

Alkuvuosi on ollut epäonnen aikaa monelle vanhalle ja uudelle silkkitieläiselle. Rahansiirroissa on ollut sekoituspalvelimen vaihdossa alkaneita ongelmia ja Bitcoinin kurssi on romahtanut alimmaksi pitkästä aikaa. Ostajilta on jäänyt tärkeitä tilauksia saamatta ajoissa kun talletuksissa on kestänyt ja myyjät ovat menettäneet rahaa kurssilaskun aikana hitaiden nostojen takia. Jotkut myyjät ovat joutuneet lopettamaan myymisen kokonaan, siirtymään katukauppaan tai vaihtamaan toisille kauppiaille. –Kapteeni 26.1.2014

Palvelu on täysin riippuvainen Kapteeni-nimimerkkiä käyttävän henkilön toiminnasta. Tammikuun 2015 toisella viikolla Silkkitien maksujärjestelmä hyytyi tuntemattomasta syystä, ja samaan aikaan Kapteeni oli asiakaspalvelun mukaan lomalla. Seurauksena oli kaupankäynnin pysähtyminen noin kahdeksi viikoksi.

Palvelu on myös täysin riippuvainen Tor-verkon ja bitcoinin toiminnasta. On toki mahdollista korvata nämä teknologiat toisilla, mutta mikäli näissä teknologioissa esiintyisi vakava tietoturvaongelma, myös luottamus lähimpiin korvaaviin järjestelmiin, anonyymiverkko I2P:hen ja virtuaalivaluutta litecoiniin, kärsisi huomattavasti.

Huumeiden myyntimäärät Silkkitie-sivustolla

Päihdekaupankäynnin tilastoinnin kannalta Silkkitie on mielenkiintoinen. Koska suurin osa kauppatavarasta liikkuu rajoja ylittämättä Suomen sisäisessä postissa, esimerkiksi tullivalvonnan kautta ei saada käsitystä kaupankäynnin laajuudesta. Tähän voidaan kuitenkin käyttää automaattisia tiedonlouhinnan keinoja. Lähestymistapaa käytettiin jo ulkomaalaisen Silk Road -palvelun kaupankäynnin tilastoinnissa (Christin 2013).

Laajasti ajateltuna julkisten tietomassojen läpikäynti on Open Source Intelligence -toimintaa (OSINT) (Glassman & al. 2012). Menetelmällä voidaan luoda laadukas tilannekuva erilaisista ilmiöistä, tässä tapauksessa suomalaisten käymästä huumausainekaupasta internetissä. Muut verkkotiedustelun keinot eivät tässä tapauksessa edes toimisi, koska Tor-verkon vahva salaus ja anonymiteetti estävät tietojen keräämisen esimerkiksi tietoliikennetiedustelulla (Dingledine & al. 2004). OSINT-mallin mukainen tiedustelu on tässä tapauksessa edullista, yksinkertaista ja sitä voi laillisesti tehdä kuka tahansa siihen teknisesti kykenevä henkilö.

Teimme alustavan kokeilun rakentamalla ohjelman, joka käy läpi kaikki tuotesivut ja poimii niistä talteen oleelliset tiedot. Nämä tiedot ovat tuotteen otsikko, hinta, kuinka monta myyntiartikkelia varastossa on, myyjän nimimerkki, myyjän positiiviset palautteet, myyjän negatiiviset palautteet, lähetysmaa, myyjän sallimat kohdemaat ja tuotteen yksilöivä URL-osoite. Tiedot haetaan kerran päivässä ja ne tallennetaan tietokantaan. Haku kohdistuu vain Silkkitien kategoriaan "nautintoaineet", joka sisältää palvelussa myytävät päihteet. Päihteet muodostavat lähes kaiken kaupankäynnin palvelussa: muut tuotteet eivät ole yhtä suosittuja ja niitä ei ole niin paljon tarjolla.

Tietokantahakujen avulla voidaan tehdä laskentoa kaupankäynnistä. Käytännössä jokaisen tuotteen kohdalla voidaan vertailla sitä, miten varastossa olevien myyntiartikkeleiden määrä muuttuu. Laskentaan otettiin mukaan vain ne myyjät, joilla onnistuneita kauppoja oli 20 enemmän kuin epäonnistuneita kauppoja. Siis esimerkiksi vähintään 20 onnistunutta kauppaa ja ei epäonnistuneita kauppoja. Näin otetaan pois laskuista myyjät, jotka eivät juuri tee kauppaa tai eivät oikeasti lähetä mitään tuotteita.

Otimme laskentaan mukaan päivät ajalta 5.11.–5.12.2014 ja saimme tältä kuukauden jaksolta seuraavanlaiset tulokset: euromääräinen myynti pääkategorioiden mukaan (taulukko 1) ja euromääräinen myynti erilaisten tuotteiden mukaan (taulukko 2). Lisäksi tarkastelimme yksittäisten myyjien liikevaihtoa.

Kategorioina käytettiin Silkkitie-palvelun lajittelua eri päihteistä. Empatogeenit on nimitys MDMA:sta ja sen kaltaisista stimulanteista, jotka lisäävät empaattisia kokemuksia toisia ihmisiä kohtaan. Dissosiaatit sisältävät ketamiinin ja DXM:n sekä niiden kaltaiset psykedeelit, jotka aiheuttavat voimakkaan kokemuksen tietoisuuden erkanemisesta muusta todellisuudesta.

Päihderyhmittäin tarkasteltuna empatogeenien, kannabiksen ja stimulanttien myynti oli suurinta. Kannabista myytiin 5 810 eurolla, opiaatteja 1 932 eurolla, empatogeeneja 7 425 eurolla, dissosiaatteja 794 eurolla, depressantteja 1 120 eurolla, stimulantteja 7 788 eurolla, psykedeelejä 4 379 eurolla ja muita päihteitä 2 719 eurolla. Yhteensä kuukauden aikana myynti oli siis noin 32 000 euroa. Päihdekohtaisesti tarkasteltaessa eniten myyntiin ekstaasia, sen jälkeen amfetamiinia, kukintoa, LSD:tä, hasista, MDMA:ta, rauhoittavia reseptilääkkeitä ja kokaiinia. Muiden tuotteiden myyntimäärät ovat hyvin vähäisiä.

On huomattavaa, että kannabistuotteiden käyttöannos maksaa huomattavasti vähemmän kuin muiden päihteiden, joten viihdekäytössä käytetyillä käyttöannosmäärillä mitattuna kannabistuotteet ovat kaikista suosituimpia. Stimulantteihin kuuluu myös kallista kokaiinia, joka nostaa kategorian myyntimäärää euroissa (taulukko 2). Huomattavaa on, että paljon puhuttujen muuntohuumeiden myynti on hyvin vähäistä.

Myynnin kokonaismäärä oli vuoden 2014 lopulla noin 1 000 euroa päivässä eli kauppa oli suhteellisen pienimuotoista ottaen huomioon, että myyjiä oli useita. Matalat myyntimäärät kertonevat siitä, että lähes kaikki myyjät myivät päihteitä Silkkitie-palvelun kautta harrastuneisuuttaan tai oheiskauppana. He eivät saaneet merkittävää rahallista voittoa myyntitoiminnasta. Koska palvelu ottaa itse noin viiden prosentin välityspalkkion, Silkkitien ylläpitäjät tienaavat palvelusta noin 50 euroa päivässä. Ottaen huomioon työmäärän ja ylläpidon pakolliset juoksevat kulut, palvelu ei ollut kaupallisesti kovin kannattava.

Laskimme tietokannasta muutaman yleisen tuotteen keskihinnat. Hinnat on pyöristetty kokonaisiksi euroiksi. Tarkastelimme lisäksi käyttöannoksen hintaa. Käyttöannoksen referenssinä käytimme Erowid-säätiön tarjoamia viitearvoja (Erowid 1995). Tiedot on esitetty taulukossa 3.

Laskimme käsin LSD:n keskimääräiseksi hinnaksi 15 euroa / 100–125 mikrogrammaa, joka on tavallinen yksi käyttöannos. Vastaavasti voitaisiin

Taulukko 1. Myyntimäärät euroissa Silkkitien käyttämien pääkategorioiden mukaisesti.

	Nimike							
	Stimulantit	**Empatogeenit**	**Kannabis**	**Psykedeelit**	**Opiaatit**	**Depressantit**	**Dissosiatiivit**	**Muut**
Euroa	7 788	7 425	5 810	4 379	1 932	1 120	794	2 719

Taulukko 2. Päihdekohtaiset myyntimäärät euroissa kuukauden ajalta.

Nimike	Euroa
Ekstaasi	5 717
Amfetamiini	5 418
Kukinto	4 639
LSD	2 686
MDMA	1 603
Bentsot	1 551
Kokaiini	1 312
Buprenorfiini	911
GBL	887
Hasis	753
DMT	735
Oksikodoni	532
Kipulääkkeet	409
Tramadoli	371
Synteettiset	366
25x-NBOMe	333
Sienet	320
MDPV	290
Salvia	290
Mieskunto	254
GHB	233
A-PVP	230
2-FMA	205
Ketamiini	170
ADHD	155
MXE	155
Katinonit	140
Nootroopit	130
Unilääkkeet	126
Ayahuasca	120
Kodeiini	118
DOx	117
Metyloni	105
Leivonnaiset	52
2C-x	47
Rauhoittavat	25
Lihasrelaksantit	22
DXM	19
AMT	10
Nuuska	10
Muut	401

laskea ja vertailla, kuinka paljon erilaisia päihdeaineita myydään, kuinka paljon eri myyjät myyvät, kuinka monta käyttöannosta on keskimääräinen ostoskerta sekä miten ostokäyttäytyminen muuttuu ajan mittaan. Lisäksi voitaisiin laskea, kuinka monta yksittäistä ostoskertaa tapahtuu.

Taulukko 3. Taulukossa on kerätystä tietokannasta automaattisesti tunnistetut grammahinnat eräille myynnissä oleville päihteille. Taulukossa on myös arvio yhden käyttöannoksen hinnasta.

	Hinta	
Päihteen kauppanimi	euroa/ gramma	euroa/käyttöannos
Metyloni	79	8 euroa / 100 mg
DMT	113	2 euroa / 20 mg
Sienet	10	10 euroa / gramma
Kukinto	26	3 euroa / 100 mg
Kokaiini	128	13 euroa / 100 mg
MDMA	65	7 euroa / 100 mg
Hasis	25	3 euroa / 100 mg
Amfetamiini	31	3 euroa / 100 mg

Virhelähteet kerätyssä tilastossa

Kerätyssä tietokannassa sekä sen perusteella tapahtuvassa laskennassa on tiettyjä virhelähteitä.

Ensinnäkin itse keräys tapahtui joka päivä hieman eri aikaan. Vaihteluväli on jopa 12 tuntia, koska toisinaan Silkkitie-palvelu toimii hitaasti tai ei vastaa lainkaan tiedonkeruuohjelman pyyntöihin. Tämä ei kuitenkaan vääristä kerättyä tietoa pitkällä aikavälillä.

Toisena virhelähteenä on laskentatapa, jossa kerättyjen tietojen muutosta verrataan aina edelliseen päivään. Jos kappalemäärä on laskenut, niin muutos tulkitaan myytyjen tuotteiden määräksi. Tässä on kaksikin ongelmaa. Jos kauppias myy tuotteita ja lisää tuotteita saman vuorokauden aikana, niin laskettu muutos edelliseen päivään ei kerro oikeaa myynnin määrää. Samoin käy, jos kauppias laskee varastossa olevien tuotteiden määrää.

Jotta myynti huomioitiin laskennassa, myyjällä täytyi olla vähintään 20 onnistunutta kauppaa. Toisin sanoen tätä vähemmän myyneet myyjät jäivät huomiotta. Samoin huomioitta jäivät myyjät, joiden huono maine viittasi siihen, että he eivät oikeasti myyneet mitään. Lisäksi otimme mukaan vain ne kaupat, joissa kauppias on Suomen rajojen sisäpuolella. Tämä rajaus tehtiin siksi, että tarkoituksena oli tarkastella nimenomaan Suomen sisäistä kauppaa.

Koska kauppiaisiin voi ottaa yhteyttä suoraan viestillä, on mahdollista, että kauppoja sovitaan Silkkitie-palvelun ohi suoraan ostajan ja myyjän välillä. Näin tehdyt kaupat ovat kuitenkin ongelmallisia kaupan osapuolille, koska sekä ostaja että myyjä koettavat kerätä hyvää mainetta juuri Silkkitie-palvelussa tapahtuvien onnistuneiden kauppojen kautta.

Johtopäätökset

Alustuvat tutkimustulokset osoittavat Suomen sisäisen kaupankäynnin olevan pienimuotoista ja harrastuspohjaista. Toisaalta Suomen sisäistä, verkossa tapahtuvaa huumekauppaa käydään muuallakin kuin Silkkitiellä, vaikka sivusto on tällä hetkellä laajin kauppapaikka. Myyjien tulot eivät ole suuria. Voidaankin ennemmin puhua laittomista päihteistä kiinnostuneiden ihmisten harrastustoiminnasta tai oheistoiminnasta kuin liiketoiminnasta. Palvelun ylläpitokin vaikuttaa olevan ylläpitäjälle ideologista toimintaa, ei niinkään liiketoimintaa. Verrattuna Silkkitien kansainvälisen esikuvan, Silk Road -palvelun toimintaan kaupankäynti on häviävän pientä. Silk Roadilla kaupankäynti oli vuonna 2012 luokkaa 1,2 miljoonaa dollaria joka kuukausi. Palvelun ylläpitäjä tienasi 92 000 dollaria kuukaudessa (Christin 2013.)

Silkkitie-palvelun kokonaismyyntimäärä euroissa on varsin pieni, noin 1 000 euroa päivässä. Tämä summa jakautuu kymmenien myyjien kesken; vain kahdeksan myyjää ylittää 1 000 euron kuukausimyyntimäärään.

Silkkitie-verkkopalvelu on toteutettu huolellisesti käyttäen uusimpia teknologioita. Kapteeni-nimimerkkinen henkilö tai ryhmä on toteuttanut verkkopalveluita aikaisemminkin ja tuntee Tor-verkon ja bitcoinin toiminnan tarkkaan. PGP:tä hyödyntävän verkkopalvelun toteutus on vaatinut aikaa ja huomattavaa suunnittelua. Silti palvelun tuoma rahallinen hyöty on keskimäärin 50 euroa päivässä.

Osalla käyttäjistä tietotekninen osaaminen on huomattavan korkea. He hallitsevat Tor-verkon käytön ja hyödyntävät PGP-salausta. Heidän kommenteistaan käy selvästi ilmi korkea tietämys siitä, miten nämä teknologiat toimivat ja oikein käytettyinä suojaavat heitä viranomaisia vastaan. Myyjät kertovat muun muassa salaavansa tietokoneensa kiintolevyt mahdollisia kotietsintöjä vastaan ja näin suojelevansa ostajien osoitetietoja.

Myyjien tietämys erilaisista päihteistä on myös täsmällistä. Myyjät kertovat avoimesti kokeilleensa itse tuotteitaan ja mitanneensa erilaisilla testeillä tuotteiden vahvuuksia. Osa myyjistä varoittaa tuotteiden aiheuttavan riippuvuutta ja korostavat kohtuukäyttöä. Suurin osa myyjistä jopa varoittaa ensi kertaa päihdettä kokeilevia käyttöannosrajoista.

Tällaisten tilastojen julkaiseminen lisää tietoisuutta Silkkitien toiminnasta. Tämä on eettisesti ongelmallista, koska salaus, Tor-verkko ja bitcoin-virtuaaliraha liitetään julkisuudessa rikolliseen toimintaan. On kuitenkin selvää, että Silkkitien toiminta on häviävän pieni osa kaikesta Tor-verkon ja bitcoinin käytöstä, josta suurin osa on täysin laillista arkipäiväistä toimintaa. Toinen ongelma julkisuudessa on, että Silkkitie voi näin saada uusia asiakkaita. Laittomia päihteitä internetistä etsivät suomalaiset löytävät kuitenkin Silkkitien muutenkin. Tutkimustiedosta taas hyötyvät sekä päihdepolitiikka että haittojen minimointia suunnitteleva tutkimusyhteisö.

Uusien teknologioiden käytön rajoittaminen on teknisesti lähes mahdotonta. Rajoituksia olisi mahdotonta valvoa ja lakia rikkovat henkilöt olisivat valmiita uhmaamaan rajoituksia. Rajoitukset aiheuttaisivat haittaa vain tavallisille netin käyttäjille, jotka pyrkisivät parantamaan tietoturvaansa tai tekemään bitcoinilla laillisia ostoksia. Bitcoin, Tor ja tietoturvateknologiat ovat yleiskäyttöisiä teknologioita, vaikka niitä voidaan käyttää myös laittomien päihteiden myyntiin.

Tutkimuksen kannalta nykymuotoiset kauppapaikat mahdollistavat kaupankäynnin seurannan ja tilastoinnin. Saatavilla on aivan uudenlaista tutkimustietoa myytyjen aineiden määristä ja myyjien saamista tuloista. Lisäksi kaupankäyntimalli on sikäli uusi, että myyjä on ostajalle täysin tuntematon ja anonyymi.

Miten anonymiteetti muuttaa huumekaupan luonnetta? Ostajan ei tarvitse tuntea myyjää. Se voi vähentää väkivaltaa, tarvetta tuntea tuottajia tai olla tekemisissä varsinaisen myyjän kanssa, ja tutustumista käyttäjäpiireihin. Samalla tähän sisältyy omat riskinsä: aloitteleva ostaja ei opi muilta kokeneemmilta käyttäjiltä turvallisia tapoja käyttää aineita.

Huumekaupalle tyypillinen luottamus ostajan ja myyjän välillä (Perälä 2011) rakentuu Silkkitiellä sille, että ostaja näkee palvelusta myyjän maineen ja toteutuneiden kauppojen kommentit. Suurimmilla myyjillä on useita satoja toteutuneita kauppoja ja lähes kaikista kaupoista on annettu positiivinen palaute. Näin ollen ostaja voi hyvällä syyllä luottaa saavansa haluamansa tuotteen. Palautejärjestelmä vähentää ostajan riskiä saada erilaista tuotetta kuin luulee ostavansa.

Jatkotutkimus

Aineisto vaikuttaa alustavan analyysin perusteella käyttökelpoiselta. Tutkimusdataa Silkkitien kaupankäynnistä kerätään edelleen päivittäin. Näin syntyvästä tietokannasta voidaan selvittää monenlaisia ilmiöitä sekä seurata pitkittäistutkimuksilla, miten kaupankäynti kehittyy. Aikomuksemme on julkaista lisää tuloksia tietojen analysoinnin pohjalta. Lisäksi julkaisemme tutkimusyhteisölle hyödyllisiä tilastotietopaketteja, joista voidaan varmistaa tekemämme tutkimuksen paikkansapitävyys. Tilastoinnit ruokkivat jatkotutkimusta ja ovat helposti hyödynnettävissä erilaisissa tutkimusasetelmissa. Alla esitämme joitakin ajatuksia jatkosta.

Automaattinen tilastointi menetelmänä huumekaupan arvioinnissa. Internetissä toimivassa kauppapaikassa tuotteet ja kaupankäynti ovat näkyvillä kaikille sivustolla kävijöille. Tätä kaupankäyntiä voidaan seurata ja tilastoida. Seurannan voi tehdä perinteisesti siten, että joku merkitsee päivittäin talteen tiedot siitä, miten tuotteita myydään. Tämä tehtävä kannattaa kuitenkin automatisoida tietokoneelle.

Tällainen ohjelmisto käy Silkkitie-sivuston läpi joka päivä ja tallentaa kaupankäyntitilanteen. Tiedot tallennetaan päiväkohtaisesti tietokantaan. Tästä tietokannasta voidaan esimerkiksi laskea, paljonko tuotteita on myyty eri viikonpäivinä. Automatisointia voi laajentaa muihin verkkolähteisiin ja yhdistellä muihin olemassa oleviin tietoihin.

Tällaisilla menetelmillä saadaan mielenkiintoisia ja hyödyllisiä tilannekuvia huumausainekaupan nykytilasta internetissä. Erityisesti Suomen kohdalla tilastointia ei ole vielä tehty, joten sen automaattinen tuottaminen olisi tarkoituksenmukaista ja hyödyllistä.

Kyselytutkimus palvelun käyttäjille. Olisi myös mielekästä toteuttaa kysely Silkkitie-palvelun käyttäjille. Näin voitaisiin kartoittaa heidän näkemyksiään Silkkitiestä ja saada tietoa heidän taustastaan ja asemasta yhteiskunnassa sekä päihteiden käyttötottumuksia. Kyselyn voisi toteuttaa esimerkiksi itse rakennetussa Tor-verkon hidden service -palvelussa, jotta vastaajat saisivat parhaan mahdollisen anonymiteetin. Samalla menetelmällä voisi kysyä Kapteenilta anonyymisti hänen motiiveistaan kehittää ja ylläpitää Silkkitietä.

Silkkitien huumemyynti vuosina 2015–2016. Vaikka dataa on tätä kirjoittaessa kerätty vasta muutamia kuukausia, vaikuttaa siltä, että huumekauppa Silkkitiellä kasvaa. Aikomuksena on kirjoittaa jatkoartikkeli, jossa kaupankäyntiä seurataan pidempi ajanjakso sekä katsotaan trendejä ja muutoksia pidemmällä ajanjaksolla.

Silkkitien huumemyynti verrattuna trendeihin katukaupassa. Silkkitiellä tapahtuvaa myyntiä voisi verrata muualta saataviin (Vinkki-pisteet, NOPSA-verkosto, hoitopaikat, kyselyt, huumetestaus) tietoihin trendeistä suomalaisessa huumeidenkäytössä.

Kansainväliset vertailut. Myyntiä voisi verrata ainakin myytävien tuotteiden laadun suhteen joihinkin muihin maihin, joissa on vastaavia kauppapaikkoja. Myydäänkö Suomessa esimerkiksi tiettyjä huumeita suhteessa enemmän verrattuna muihin Pohjoismaihin? Kauppapaikat eivät edusta koko internetissä tapahtuvaa kauppaa, joten määrien vertailun sijaan kannattanee keskittyä siihen, mihin aineisiin kauppa keskittyy.

KIRJALLISUUS

Barratt, Monica J.: Silk road: eBay for drugs. Addiction 107 (2012): 3, 683–683.

Christin, Nicolas: Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2013.

Dingledine, Roger & Mathewson, Nick & Syverson, Paul: Tor: The second-generation onion router. Naval Research Lab Washington DC, 2004.

Erowid: Erowid Center, 1995. https://www.erowid.org/ (luettu 26.01.2015)

Glassman, Michael & Min Ju Kang: Intelligence in the internet age: The emergence and evolution of Open Source Intelligence (OSINT). Computers in Human Behavior 28 (2012): 2, 673–682.

Greenberg, Andy: Undercover Agent Reveals How He Helped the FBI Trap Silk Road's Ross Ulbricht Wired, 2015. http://www.wired.com/2015/01/silk-road-trial-undercover-dhs-fbi-trap-ross-ulbricht/

(luettu 5. 3.2015)

Kapteeni: Portti Silkkitielle – Olet pian ostoksilla. Silkkitie, 2014. http://silkkitie.net/ (luettu 26.01.2015)

LocalBitcoins: Buy and sell bitcoins near you. LocalBitcoins Oy, 2014. https://localbitcoins.com/ (luettu 14.01.2015)

Nakamoto, Satoshi: Bitcoin: A peer-to-peer electronic cash system. Bitcoin.org, 2008. https://bitcoin.org/bitcoin.pdf (luettu 13.7.2015)

Perälä, Jussi: ”Miksi lehmät pitää tappaa?” Etnografinen tutkimus 2000-luvun alun huumemarkkinoista Helsingissä. Tutkimus: 56. Helsinki: Terveyden ja hyvinvoinnin laitos, 2011.

Silkkitie: Asiakaspalvelu, 2015. http://silkkitiehdg5mug.onion/asiakaspalvelu (luettu 13.7.2015)

The Guardian: Tor: 'The king of high-secure, low-latency anonymity', 2013. http://www.theguardian.com/world/interactive/2013/oct/04/tor-high-secure-internet-anonymity (luettu 25.1.2015)

Thomas, Ryan: Attacks on PGP: A Users Perspective. SANS Institute, 2003. http://www.sans.org/reading-room/whitepapers/vpns/attacks-pgp-users-perspective-1092 (luettu 13.7.2015)

Tor Project: Tor Browser. The Tor Project, Inc. https://www.torproject.org/projects/torbrowser.html.en (luettu 18.1.2015)

Van Hout, Marie Claire & Bingham, Tim: Silk Road, the virtual drug marketplace: a single case study of user experiences. International Journal of Drug Policy 24 (2013): 5, 385–391

Van Hout, Marie Claire & Bingham, Tim: Responsible vendors, intelligent consumers: Silk Road, the online revolution in drug trading. International Journal of Drug Policy 25 (2014): 2, 183–189.

Zimmermann, Philip: Pretty good privacy: public key encryption for the masses. Teoksessa

Lance J. Hoffman (toim.): Building in big brother: The Cryptographic Policy Debate. New York: Springer-Verlag, 1995.

TIIVISTELMÄ

Juha Nurmi & Teemu Kaskela: Silkkitie. Päihteiden suomalaista nappikauppaa

Tietotekniikan ja tietoverkkojen alati kiihtyvä kehitysvauhti on siirtänyt monet palvelut internetiin. Näin on käynyt myös päihdyttävien aineiden kaupalle. Tämän lisäksi nykyiset teknologiat ja teknologioiden käyttäjät ovat entistä paremmin valveutuneita erilaisia tietoturvaongelmia, massavakoilua, sensuuria ja ihmisoikeusloukkauksia vastaan. Turvallisuutta parantamaan valmistettuja teknologioita osaavat hyödyntää myös laittomia päihteitä myyvät ja ostavat henkilöt. Keskeinen esimerkki tästä on suomalainen Silkkitie-niminen Tor-verkossa sijaitseva nettikauppa.

Erityisesti vuonna 2014 kaupankäynti siirtyi Suomen rajojen sisäiseksi. Toisin sanoen tilaukset kulkevat Suomessa postin välityksellä ja tullivalvonnan kaltaista riskiä ostajalle ei enää ole. Tor, bitcoin sekä laadukkaasti toteutettu verkkokauppa tarjoavat yhdessä saumattoman turvallisen tavan ostaa myös laittomia päihteitä. Käytännössä viranomaisilla ei ole näköpiirissä keinoja puuttua näin toimivaan kaupankäyntiin. Toisaalta, koska kaupankäynti keskittyy yhteen julkisesti toimivaan kauppapaikkaan, kaupankäyntiä voidaan tilastoida automatisoidusti. Ymmärrys päihteiden käytöstä kasvaa.

Laittomien päihteiden verkkokauppaan on erittäin hankala puuttua. Voimme kuitenkin seurata sivusta kaupankäynnin laajuutta ja ymmärtää käytettyjä teknologioita. Voimme myös tuottaa lähes reaaliaikaista tilannekuvaa siitä, kuinka paljon ihmiset ostavat ja myyvät päihteitä. Näin syntyvä tilastointi valottaa aivan uudella tavalla päihteiden kulutustottumuksia Suomessa. Myös itse teknologiat, joita ihmiset ovat valjastaneet turvalliseen kaupankäyntiin, ovat erittäin kiinnostavia ja uudet sosiaaliset käyttökontekstit tutkimattomia. Tässä tutkimuksessa tarkastellaan suomalaista Silkkitie-nimistä kauppapaikkaa. Tutkimme tällaista kaupankäyntiä niin teknisenä ja sosiaalisena ilmiönä.

PUBLICATION
II

Observing Hidden Service Directory Spying with a Private Hidden Service Honeynet

J. Nurmi, J. Kannisto and M. Vajaranta

11th Asia Joint Conference on Information Security (AsiaJCIS). ed. by 2016, 55–59

2016 11th Asia Joint Conference on Information Security

Observing Hidden Service Directory Spying with a Private Hidden Service Honeynet

Juha Nurmi, Joona Kannisto and Markku Vajaranta
Tampere University of Technology
Email: juha.nurmi@ahmia.fi, joona.kannisto@tut.fi, markku.vajaranta@tut.fi

***Abstract*—Tor's location hidden services (HS) are a tool for anonymous publishing, with the feature that the sites cannot be brought down without taking down the whole Tor network. People run HSs for a multitude of reasons. Some like them to be public, but others want to keep them their existence as private. We have run private unannounced HSs to detect whether the HS directory is spied on. Our results show that the hidden service directory is monitored for new addresses. This paper details the observations made from the scanning activity.**

I. Introduction

Location Hidden Services (HS) [1] are TCP services offered over the Tor network. The services are implemented in such a way that the Internet Protocol address of the publishing server is not visible to the clients. The clients know the service only by its pseudonymous identifier. The HS identifiers are called onion addresses because they are under a reserved special use domain called *.onion*.

HSs are made for many different purposes. Anyone can start and operate a hidden service, there are sites for forbidden content, drug and other illegal item market places, basically information that would not be tolerated anywhere else. The most prominent examples of cases where the existence of these services is beneficial for us all are leak sites that are operated by journalists. These services need to remain online, and out of reach of local governments.

The HS operator can decide whether they want to make their service public – by including it in a search engine, such as Ahmia [2] – or to keep it as a an unlisted one. Yet, if the nature of the hidden services is to publish, why would anyone run a hidden service and not tell anyone about it?

- content is very sensitive
- site is not finished
- does not want random traffic
- does not know how to password protect the HS

Exclusion from public HS directory as a protection mechanism does not make much sense, other than to limit traffic coming to the HS. For instance, SSH management connection can use a hidden service address.

For instance, Onionshare [3] lets users share files between each other. It authenticates the user by a knowledge of a short string at the end of the URI. Yet, such strings can be easily brute forced [4], and the files could end up in the wrong hands. In addition, the password protection requires the user to remember two random strings.

Furthermore, password protection using Tor's inbuilt mechanism *HiddenServiceAuthorizeClient* needs a strong password against an attacker who can get the HS descriptor. The password protected HS descriptors include an encrypted list of introduction points. The encryption uses AES in CTR mode with 128-bit key, which can be considered to be strong. Yet, if the password used to derive the encryption key is weak, it can be dictionary attacked using the directory information.

It is possible to discover descriptors for previously unknown onion addresses by running a Tor relay [5]. Attacker motives are many. State actors could be scanning the Tor network, there could be curious hobbyists, and someone may be trying to find valuable information. This paper sets to find out whether this happens in practice.

The paper is organized as follows. In the background section we outline the operation of Tor network's HSDir. In the third section we present related work on Tor and honeypots. Then in fourth section we describe our research questions and our experimental setup. In the fifth section results are presented. Discussion about our results along with future work and improvement ideas are presented in the sixth section. Finally, the seventh section draws the conclusions.

II. Background

Clients connect to the HS through introduction points that the service has listed. The HS has Tor circuit open to the introduction point, over which the introduction point relays the setup parameters (rendezvous point address and a shared secret) to the hidden service. The introduction points for a specific onion address are listed in a record called the service descriptor [1]. Tor network stores them in the global HS directory (HSDir), which is a distributed hash table (DHT). Originally the HS design included directory authorities, but in the second version any relay with HSDir flag can hold directory contents in a circular fashion [1].

Tor implements DHT (Distributed Hash Table) for onion address information. Every Tor hidden service directory (HSDir) node has a subset of hidden service descriptors. These are short signed messages created by hidden services. Most important information in the message is a list of introduction points, where the client can send a short introduction message. The hidden service then publishes its descriptor to a set of 6 responsible HSDirs once per hour. The responsible HSDirs are regular Tor nodes on the Tor network which have been online longer than 92 hours and which have received the HSDir flag

978-1-5090-2285-4/16 $31.00 © 2016 IEEE
DOI 10.1109/AsiaJCIS.2016.31

from the directory authorities. The set of responsible HSDirs is based on their position of the current descriptor-id in a list of all current HSDirs ordered by their node fingerprint.

Because the onion address identifies a public key, the hidden service can prove its ownership of the address for the connecting client. The service descriptors are signed with the public key as well. This makes the HSDir's operation less critical. Only availability of the HSDir has to be guaranteed. Regular Tor node may work as HSDir and anyone who deploys such node can also harvest onion addresses from it. The attacker may find addresses that have not been publicly shared ever.

The next generation hidden service proposal for Tor [6] makes it impossible for a HSDir operator to connect to the HS based on only the HS descriptor. The descriptor contents are encrypted with a key derived from the hidden service's identifier and other parameters such as time. However, this specification is not in the use yet. The new proposal will also change the hidden service naming to include a 256 bit ECC public key, instead of the current RSA key truncated hash. The descriptor-ids will not be predictable to the future either [5], [6].

Honeypot is a monitored network of decoy computers, whose purpose is to be attacked [7]. Traffic coming to the honeypot can be analyzed in order to find information about possible attacks against real systems or to profile the attacker. Sometimes the honeypot is designed to keep the attacker busy and away from real systems, particularly if the attack requires human effort.

The attackers goal is to detect that they are running in a honeypot, so that they will not reveal information about their attack tools or waste effort [8]. Low interaction honeypots [9], such as Kippo, usually seek to understand the behavior of automated or low sophistication attack traffic, such as SSH brute force [10]. Such a honeypot only emulates a real environment, and lets the attacker run different commands, which are then logged by the honeypot. High interaction honeypots [9] that are real but controlled systems are more laborous for the defender, yet, they do not reveal themselves as easily.

These systems appear either vulnerable or valuable systems, depending on what the focus of the honeypot is. If the focus is to gather information about general automated attacks, appearing vulnerable could be a good strategy. Contrary, if the idea is to gather information about targeted attacks [11], the honeypot should be, for example, named to appear very valuable.

Honeynets, or more specifically darknets are monitored network regions [12] that should not receive traffic in normal circumstances. Any traffic coming to the darknet addresses can be considered to be either malicious scanning or backscatter traffic from attacks like SYN-flood.

III. Related Work

Honeypots have been built for special purposes. For example, honeypots disguised as IoT devices [13]. Other examples are honeypots that appear as industrial control systems [14], [15], [16], [17]. Further, honeypots that target the mobile malware and attackers are presenting themselves as Android phones [18].

Yet, there is a distinct lack of HS honeypot research, even though it can be estimated that many organizations are actively scanning HSs. The behavior of Tor exit nodes has been explored with honeypots by Chakravarty et al. [19]. Tor hidden services have been used as part of a Internet facing setup to bridge traffic towards hosts in the Tor network [20]. However, traffic coming to HSs has not been under investigation previously.

Honeypot research in the HS case is complicated because the hidden service does not see the IP address of the connecting client, unless the client uses Tor2Web mode [21]. This makes the attacker attribution hard. Indeed, previous Tor HS research has revolved around the deanonymization of HS operators [5], [22] and Tor users [23].

IV. Description

In this section we outline the design choices and rationale for our experiment.

A. Research Questions

Because there is the possibility of spying the HSDir, we wanted answers for the following questions:

- Is someone monitoring the HS directory?
- How large parts of HS directory are under the monitoring?
- How frequent is the behavior, and what is the level of automation?
- What services are targeted?
- What kind of observers visit private HSs?

The first question is interesting since, there is a certain expectation of privacy, when setting up an unlisted hidden service. It is also considered to be unethical to harvest onion addresses from the HSDir while acting as a volunteer operator.

We were also interested to know whether all onion addresses are revealed to the attacker as fast as they are generated. And if they are scanned only periodically, what kind of interval the attacker has (i.e. what is the mean time of exposure).

We wanted also information about what kind of services were targeted. For example, if someone would be port scanning the onion network. Related to this would be information about automated vulnerability scans.

What can we tell about the headers that the attacker is sending to us? Even though the attacker can easily change the protocol headers that their client is using, careless attackers may reveal information about themselves.

B. Honeynet Setup

In total 100 honeypot onions were set up for the experiment. All of them had three ports open, port 21 (FTP), port 22 (SSH) and port 80 (HTTP). The addresses were randomly generated, and were not told to anyone. The experiment was run for 42 days.

All requests were logged. Only successful protocol data was logged. In other words, if an attacker would have sent an HTTP query to a SSH port, the query would have been silently discarded. Therefore, pure port scans would not have been seen.

The used HTTP daemon was Nginx [24]. The configured web site was very plain, and contained only a static front page. Later designs of the same experiment should perhaps contain more extensive setup.

V. Results

The only port that received any traffic was the HTTP. For instance, SSH brute forcing attacks, which are common on the normal Internet were not visible.

From the logs we first analyzed the first accesses to each of the onions. These are visualized in the Figure 1. The first access was usually by an automated script, which was not trying to hide its nature but included the Curl or Wget user agents.

The second line in the figure 1 is the time when the first *Tor Browser* user agent accesses the honeypot. This most probably means that the site is listed in a service that lists these unannounced hidden services.

The first request came on 12 days after the experiment was started. This means that the onion address space is either not under constant scanning, or the scanning is done with very low resources. The HSDir is distributed which should mean that the attacker cannot get all the addresses, but only a fraction of them. In order to know all onion addresses the attacker has to run a sizable operation.

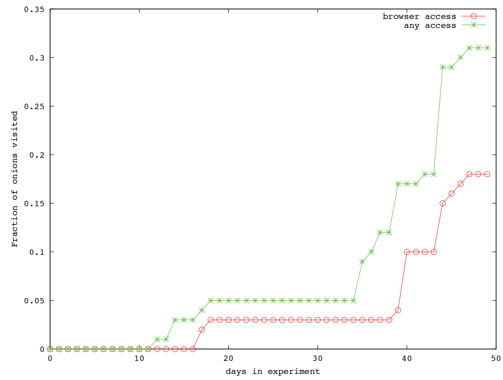


Fig. 1. Cumulative accesses of onion addresses as function of time

While the number of probed onions shows an estimate for the breadth of the activity, we also analyzed the characteristics of the revealed onions. The Figure 2 shows the logarithmic numeric value of the revealed onion addresses as a function of time. It is easy to see from the figure that most of the revealed addresses are in two bands. However, the onion addresses should be distributed over the HSDir nodes evenly, as the outer Hash function in equation 1 will mix them [1].

Fig. 2. Onion address high order bits versus time

$$\begin{aligned}&\text{descriptor-id}\\&= H(\text{permanent-id}|H(\text{time-period}|\text{descriptor-cookie}|\text{replica}))\end{aligned} \tag{1}$$

Yet, the seeming discrepancy can be explained. What is not independent of the onion address, however, is the time period of the hidden service descriptors validity (equation 2). The validity time period is divided to 256 intervals over a day, and these intervals are then indexed by the highest order byte of the onion address. This is to ensure all the HSs do not upload new descriptors at the same time.

$$\text{time-period} = \lfloor \{\text{permanent-id}\}_8/256 + \frac{\text{unix-time}}{86400} \rfloor \tag{2}$$

Indeed, using this interpretation on figure 2, we note that actually the connecting feature is that the update time-period of the onion addresses is the same. This most probably means that the attacker has been running a HSDir only at a couple specific moments. Accessing these gathered onion addresses is then distributed more evenly over the following days.

The volume of the traffic coming to the honeypot sites is seen in the figure 3. The traffic grows exponentially towards the end of the experiment. If the experiment would have continued for a longer period, the traffic might have stabilized to some value. The observation is that the majority of traffic came after listings in public directories.

The number of people surfing with normal browsers was surprisingly high. In total, 82% of all the requests came from Mozilla browser user agents. This could also mean that some of the crawlers are faking their user agents. Many visitors with normal browsers came to the site with referrer information http://skunksworkedp2cg.onion.

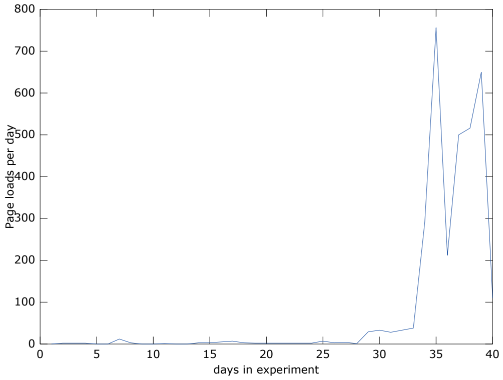


Fig. 3. Daily page loads

VI. Discussion and Future Work

The conducted study has its limitations. One is the lack of logging on the connection level. Logging also the connections would have allowed us to see possible port scans.

We did not have polling for our hidden services, so we cannot be sure whether they were reachable for the whole time. The polling was left out as to not interfere with the experiment itself. However, now that the phenomenon has been observed independently, the subsequent experiments could include monitoring.

The HSDir operators are not assumed to be honest. Therefore, scanning operations like the one seen in this paper are quite probable. They do not cause major problems for the protocol, and could even work as an incentive for a relay operator. However, what is seen in practice is that the relay operator gives out capacity only temporarily. Any wanting to be HSDir node has required uptime of 92 hours, this was increased from the previous 24 hours [5].

We can conclude that the attacker is not very systematic nor powerful. Out of our 100 HSs, only 30% were revealed to the attacker. These were highly correlated with time. The main threat was privacy of the HS, as the attacker added the sites to a directory that they operate at http://skunksworkedp2cg.onion. While the scanning was very infrequent, we can interpolate from our results that half of the hidden services would be revealed to the attacker observed in our experiment after 80 days. Likewise, a hidden service operational for five days should have a five percent chance of getting listed by our attacker.

The subject is necessary to talk about, as HSDir manipulation has a related phenomenon, which is the purposeful deployment of Onion Routers to specific locations in the DHT, detailed by Biryukov et al. [5]. Even though, the rate of the revealed onions was low, it was higher than what would be expected by a small operation of only one relay. Therefore, we expect that our attacker was using some of the methods outlined in the paper.

In the future we plan to look at HSDir attacks in more detail. For instance, it would be possible to detect whether someone is generating nodes to be HSDirs for specific onions [25], [5]. One simple detection mechanism would be to check for HSDir nodes that have abnormal public keys, i.e. RSA keys that do not have exponents 3 or 65537. Yet, making indistinguishable keys is only a computational nuisance. Therefore in the long run, statistical methods could be more effective.

Another line of future research is to make dictionary guessable password protected service descriptors, and monitor traffic coming onto them. If there is traffic, we can confirm that someone is attacking vulnerable password protected stealth hidden services.

Improving our honeypots is also a next term target. While we do not believe that HS honeypots should be the same as their clear net brethren, our current design has lots to improve. Therefore, interesting follow up study would be to build honeypots that have more features and more interesting content. Both approaches to honeypots could be tried where some honeypots could pretend to be vulnerable and others to contain valuable information.

VII. Conclusions

Our investigation into private HS publishing revealed some risks for individuals who may not take into account the possibility of HSDir monitoring. The knowledge of the HS address should not be used as an authentication measure. Neither should the services use low entropy passwords (or tokens).

We have shown that the Tor HSDir is under monitoring by parties that operate Tor Relays. Results indicate the operation is not very large scale. Based on our observations short-lived onion services should remain unnoticed for a few days with a high probability. The upcoming move to the newer HS specification will make these attacks impossible altogether [6]. We hope that our findings will give motivation for this transition.

References

[1] R. Dingledine, N. Mathewson, A. Lewman, K. Loesing, S. Hahn, R. Ransom, J. Bobbio, D. Goulet, and D. Johnson, "Tor rendezvous specification," Tor Project, Tech. Rep., 2015. [Online]. Available: https://gitweb.torproject.org/torspec.git/tree/rend-spec.txt

[2] J. Nurmi. Search for tor hidden services. [Online]. Available: http://msydqstlz2kzerdg.onion/

[3] M. Lee. Securely and anonymously share a file of any size. [Online]. Available: https://github.com/micahflee/onionshare

[4] M. Georgiev and V. Shmatikov, "Gone in six characters: Short urls considered harmful for cloud services," *arXiv preprint arXiv:1604.02734*, 2016.

[5] A. Biryukov, I. Pustogarov, and R. Weinmann, "Trawling for tor hidden services: Detection, measurement, deanonymization," in *Security and Privacy (SP), 2013 IEEE Symposium on*. IEEE, 2013, pp. 80–94.

[6] N. Mathewson and G. Kadianakis, "Next-generation hidden services in tor," Tor Project, Tech. Rep., 2014. [Online]. Available: https://gitweb.torproject.org/torspec.git/tree/proposals/224-rend-spec-ng.txt

[7] N. Provos *et al.*, "A virtual honeypot framework." in *USENIX Security Symposium*, vol. 173, 2004, pp. 1–14.

[8] S. Mukkamala, K. Yendrapalli, R. Basnet, M. Shankarapani, and A. Sung, "Detection of virtual environments and low interaction honeypots," in *Information Assurance and Security Workshop, 2007. IAW'07. IEEE SMC*. IEEE, 2007, pp. 92–98.

[9] D. Watson. Low interaction honeypots revisited. [Online]. Available: https://www.honeynet.org/node/1267

[10] C. Valli, "Ssh–somewhat secure host," in *Cyberspace Safety and Security*. Springer, 2012, pp. 227–235.

[11] K. G. Anagnostakis, S. Sidiroglou, P. Akritidis, K. Xinidis, E. P. Markatos, and A. D. Keromytis, "Detecting targeted attacks using shadow honeypots." in *Usenix Security*, 2005.

[12] D. Inoue, M. Eto, K. Yoshioka, S. Baba, K. Suzuki, J. Nakazato, K. Ohtaka, and K. Nakao, "nicter: An incident analysis system toward binding network monitoring with malware analysis," in *Information Security Threats Data Collection and Sharing, 2008. WISTDCS'08. WOMBAT Workshop on*. IEEE, 2008, pp. 58–66.

[13] Y. M. P. Pa, S. Suzuki, K. Yoshioka, T. Matsumoto, T. Kasama, and C. Rossow, "Iotpot: analysing the rise of iot compromises," in *9th USENIX Workshop on Offensive Technologies (WOOT 15)*, 2015.

[14] J. P. Disso, K. Jones, and S. Bailey, "A plausible solution to scada security honeypot systems," in *Broadband and Wireless Computing, Communication and Applications (BWCCA), 2013 Eighth International Conference on*. IEEE, 2013, pp. 443–448.

[15] P. Simões, T. Cruz, J. Gomes, and E. Monteiro, "On the use of honeypots for detecting cyber attacks on industrial control networks," in *proc of 12th European Conf. on Information Warfare and Security (ECIW 2013)*, 2013.

[16] S. M. Wade, "Scada honeynets: The attractiveness of honeypots as critical infrastructure security tools for the detection and analysis of advanced threats," 2011.

[17] A. V. Serbanescu, S. Obermeier, and D.-Y. Yu, "Ics threat analysis using a large-scale honeynet," in *Proceedings of the 3rd International Symposium for ICS & SCADA Cyber Security Research*. British Computer Society, 2015, pp. 20–30.

[18] C. Mulliner, S. Liebergeld, and M. Lange, "Poster: Honeydroid-creating a smartphone honeypot," in *IEEE Symposium on Security and Privacy*, 2011.

[19] S. Chakravarty, G. Portokalidis, M. Polychronakis, and A. D. Keromytis, *Recent Advances in Intrusion Detection: 14th International Symposium, RAID 2011, Menlo Park, CA, USA, September 20-21, 2011. Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, ch. Detecting Traffic Snooping in Tor Using Decoys, pp. 222–241. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-23644-0_12

[20] K. Anagnostakisi, S. Antonatos, and E. Markatos, "Honey@ home: A new approach to large-scale threat monitoring," in *The Proceedings of the 5th ACM Workshop on Recurring Malcode, WORM*, 2007.

[21] S. Aaron and G. Virgil. Tor2web: Browse the tor onion services. [Online]. Available: https://tor2web.org/

[22] R. Jansen, F. Tschorsch, A. Johnson, and B. Scheuermann, "The sniper attack: Anonymously deanonymizing and disabling the tor network," DTIC Document, Tech. Rep., 2014.

[23] A. Johnson, C. Wacek, R. Jansen, M. Sherr, and P. Syverson, "Users get routed: Traffic correlation on tor by realistic adversaries," in *Proceedings of the 2013 ACM SIGSAC Conference on Computer &#38; Communications Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 337–348. [Online]. Available: http://doi.acm.org/10.1145/2508859.2516651

[24] W. Reese, "Nginx: the high-performance web server and reverse proxy," *Linux Journal*, vol. 2008, no. 173, p. 2, 2008.

[25] O. Donncha. Trawling tor hidden service – mapping the dht. [Online]. Available: http://donncha.is/2013/05/trawling-tor-hidden-services/

PUBLICATION

III

Seller's reputation and capacity on the illicit drug markets: 11-month study on the Finnish version of the Silk Road

J. Nurmi, T. Kaskela, J. Perälä and A. Oksanen

Drug & Alcohol Dependence 178.(2017), 201–207

Drug and Alcohol Dependence 178 (2017) 201–207

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

journal homepage: www.elsevier.com/locate/drugalcdep

Full length article

Seller's reputation and capacity on the illicit drug markets: 11-month study on the Finnish version of the Silk Road

Juha Nurmi[a], Teemu Kaskela[b], Jussi Perälä[c], Atte Oksanen[d,*]

[a] Laboratory of Pervasive Computing, Tampere University of Technology, Finland
[b] A-Clinic Foundation, Finland
[c] Department of Social Research, University of Helsinki, Finland
[d] Faculty of Social Sciences, University of Tampere, 33014, Finland

ARTICLE INFO

Keywords:
Illicit drugs
Cryptomarket
Internet
Tor network
Anonymity
Longitudinal
Web crawling

ABSTRACT

Aims: This 11-month study analyzed illicit drug sales on the anonymous Tor network, with a focus on investigating whether a seller's reputation and capacity increased daily drug sales.
Design and setting: The data were gathered from Silkkitie, the Finnish version of the Silk Road, by web crawling the site on a daily basis from (November 2014 to September 2015). The data include information on sellers (n = 260) and products (n = 3823).
Measurements: The measurements include the sellers' reputation, the sale amounts (in euros), the number of available products and the types of drugs sold. The sellers' capacity was measured using their full sales potential (in euros). Fixed-effects regression models were used to estimate the effects of sellers' reputation and capacity; these models were adjusted for the types of drugs sold.
Findings: Overall, illicit drug sales totalled over 2 million euros during the study, but many products were not sold at all, and sellers were active for only a short time on average (mean = 62.8 days). Among the products sold, stimulants were most widely purchased, followed by cannabis, MDMA, and psychedelics. A seller's reputation and capacity were both associated with drug sales.
Conclusion: The Tor network has enabled a transformation in drug sales. Due to the network's anonymity, the seller's reputation and capacity both have an impact on sales.

1. Introduction

Illicit online drug sales first attracted major media attention with the rise and fall of the Silk Road in 2011–2013 (Barratt et al., 2014; Martin, 2014a; Martin, 2014b). The Federal Bureau of Investigation shut down the Silk Road in October 2013 and its successor, the Silk Road 2, in November 2014 (Dolliver, 2015). The Silk Road 3 was launched almost immediately after but was overrun by other sites, such as Agora and Evolution, which adopted some features of the original Silk Road (Dolliver and Kenney, 2016; Van Buskirk et al., 2016). The common feature of all of these websites is that they protect the anonymity of their users; they are hence referred to as "cryptomarkets" (Aldridge and Décary-Hétu, 2016; Bancroft and Reid, 2017; Demant et al., 2016; Martin, 2014b; Van Buskirk et al., 2016). This article focuses particularly on the Finnish online service called Silkkitie.

Silkkitie operates using the Tor network to hide its real location, just as Agora and various Silk Road versions did. The Tor (which stands for The Onion Router) network and software are designed to defend human rights (Dingledine et al., 2004). This network can be used for censorship circumvention, online anonymity and high-level online privacy. It is possible to host a website inside the Tor network (a so-called onion service) and to hide the physical location of the site's server. Tor provides anonymity by routing the user's traffic through three separate relay servers so that it is difficult to reveal the user's physical location or IP address. This means that Tor protects users by encryption to ensure privacy, authentication between clients and relays and signatures to ensure that all clients know the same set of relays.

Tor is considered very secure and resilient against surveillance: The U.S. National Security Agency called Tor "the king of high-secure, low-latency Internet anonymity" (The Guardian, 2013) in its top-secret documentation, in which it discusses its futile attempts to spy on Tor users. Despite this, different police operations have targeted sites and have at least been partially successful in reducing the activity in cryptomarkets (Décary-Hétu and Giommoni, 2017). Also, the identity of the original Silk Road developer, Ross William Ulbrich, was discovered after the FBI followed his behaviour pattern and gathered cumulative

[*] Corresponding author.
E-mail addresses: Juha.nurmi@ahmia.fi (J. Nurmi), teemu.kaskela@a-klinikka.fi (T. Kaskela), Jussi.perala@helsinki.fi (J. Perälä), atte.oksanen@uta.fi (A. Oksanen).

http://dx.doi.org/10.1016/j.drugalcdep.2017.05.018
Received 3 January 2017; Received in revised form 10 May 2017; Accepted 10 May 2017
Available online 21 June 2017

critical information for long enough (FBI, 2014).

The anonymity of online drug markets is guaranteed by digital currencies and payment systems that significantly increase the difficulty of tracking down buyers and sellers. Bitcoin is one digital currency and a payment system in one; it was invented by an unknown entity who calls himself Satoshi Nakamoto (Nakamoto, 2008; Bohannon, 2016; Kristoufek, 2015). Bitcoin is a peer-to-peer application that allows digital money transfers between users. Its monetary system is decentralised and designed to work without central banks, governments and regulations. Although Bitcoins are used in multiple legal ways, these anonymity properties also make Bitcoin attractive to criminals.

The work on illicit cryptomarkets has expanded rapidly since 2015. Many of the studies have been based on the Silk Road (e.g., Aldridge and Décary-Hétu, 2016; Barratt and Maddox, 2016) and their successors, such as the Silk Road 2.0, Agora, Alphabay and Valhalla (e.g., Van Buskirk et al., 2016 Décary-Hétu and Giommoni, 2017; Demant et al., 2016; Dolliver and Kuhns, 2016; Van Hout and Hearne, 2017). Studies have employed traditional research methods such as surveys and interviews (Barratt et al., 2014; Van Hout and Bingham, 2013a, 2013b) and newer approaches employing web crawlers (Aldridge and Décary-Hétu, 2016; Demant et al., 2016; Dolliver, 2015; Dolliver and Kenney, 2016; Christin, 2013; Hardy and Norgaard, 2015; Munksgaard et al., 2016). Some studies have also investigated the quality of drugs sold in the cryptomarkets (Caudevilla et al., 2016; Rhumorbarbe et al., 2016). There is also currently a need to understand that cryptomarkets are also used very locally, as in our Finnish case. So far country cases have involved, for example, Canada (Broséus et al., 2016) and Switzerland (Rhumorbarbe et al., 2016), and some studies have involved cross-national comparison (Van Buskirk et al., 2016).

Studies on the Silk Road indicate that users were typically males in their 20s and prioritised the Silk Road over street markets for quality reasons and for personal safety (Barratt et al., 2014; Van Hout and Bingham 2013b; Barratt et al., 2016a,b). Cryptomarket users have been characterised as a "technological drug subculture" (Van Hout and Bingham, 2013b), which also represents a form of online activism underlining individual freedoms based on libertarian ethos (Maddox et al., 2016; Munksgaard and Demant, 2016). Users' right to choose is combined with the expressed joy of having the opportunity to choose their preferred drugs like "kids in a candy store" (Barratt et al., 2016b).

As cryptomarkets are organised to guarantee the full anonymity of the users, they also bring challenges for the users. Social psychologists have underlined that anonymous behaviour can be highly regulated on the Internet (Keipi and Oksanen, 2014; Spears et al., 2002). In cryptomarkets, anonymity is enforced by instability, which may even enforce the anonymity effects of online behaviour. Instability is one of the central features of online drug markets, as most items are sold quickly and the majority of sellers disappear within a few months (Christin, 2013). Hence, there is a need for building "reputation systems" (Resnick et al., 2000; Houser and Wooders, 2006) or "trust systems" (Lusthaus, 2012) on cryptomarkets. Public user feedback is one of the common parts of such reputation or trust systems.

A previous study by Hardy and Norgaard (2015) showed that seller reputation had a positive impact on cannabis sales. They argue that users' feedback is a central feature in cannabis sales on the Silk Road because sellers and buyers do not know each other. Positive reputation gives the impression that the seller is trustworthy, which leads to increased sales (Hardy and Norgaard, 2015). It has also been found that vendors who had poor user ratings were more likely to take the risk of shipping drugs internationally (Décary-Hétu et al., 2016). Besides user feedback, it is often important to communicate well with the customer and create an impression of trustworthiness. According to Décary-Hétu and Leppänen, successful online criminals have communicative skills that lead to higher rewards (Décary-Hétu and Leppänen, 2013).

In the absence of visual cues that are typically available in an offline context, online buyers must rely on what information is available, which also separates online cryptomarkets from the traditional drug trade. We argue here, in addition to the seller's reputation, that seller's apacity is another potential factor of successful sales. Seller's capacity refers to sellers' promises regarding the variety and quantity of high-quality products available. This is part of communication with the customers. Seller's capacity or promise of a lot of drugs in stock will create positive expectations that will lead to higher sales. It is common for cybercriminals to construct an online identity and create trustworthiness to attract criminal partners (Lusthaus, 2012). Seller's capacity is part of this communication, but it has not been previously studied with regard to cryptomarkets. However, economic studies on advance selling have shown that seller capacity is positively associated with consumer behaviour (Yu et al., 2015).

In this study, our aim was to research both the seller's reputation and capacity in Silkkitie. We will first provide descriptive results on the general drug sales on Silkkitie. The main focus of the article lies, however, on the reputation and capacity that represent different facets of the seller that are available to buyers. Anonymity directs users' attention to the available features and typically also enforces group behaviour (Keipi and Oksanen, 2014; Spears et al., 2002). Hence, in cryptomarket, trust system is based on both direct feedback and successful communication (Hardy and Norgaard, 2015; Resnick et al., 2000; Houser and Wooders, 2006; Décary-Hétu and Leppänen, 2013; Lusthaus, 2012). Our hypotheses were grounded on these premises, and we expected that a positive seller reputation would be positively associated with that entity's daily sales. In addition, we expected that an increase in a seller's capacity would increase that entity's sales.

2. Methods

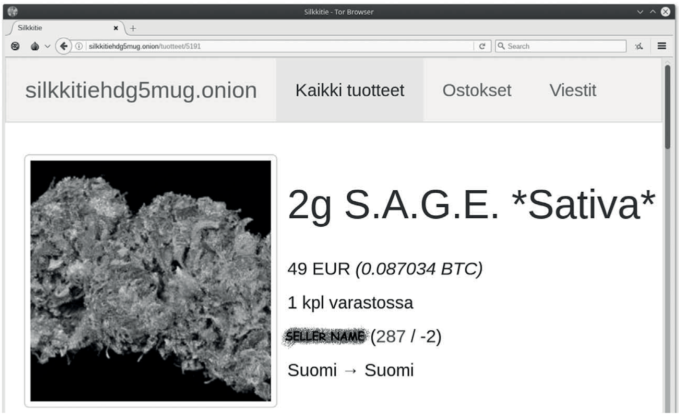
2.1. Silkkitie as a marketplace

Silkkitie was opened in the Tor network on 6 January 2014, and it was originally intended to be used specifically for illicit drug sales in Finland. Because both the seller and the buyer are located within Finland, even detecting the drugs is difficult because there is no customs process for screening domestic shipments. For these reasons, virtually all of the Finnish sellers mention that the shipments are posted via domestic mail. As Silkkitie is the main online marketplace for Finnish drug buyers, it is a useful measure of the overall Finnish online drug market. In 2015, Silkkitie published the English translation of the site to attract user space outside Finland. The original Finnish version (http://silkkitiehdg5mug.onion/) now has an English translation called Valhalla (http://valhallaxmn3fydu.onion/). Bitcoin currency and the Bitcoin wallet system are used for payment, and Silkkitie requires sellers to use PGP encryption (Pretty Good Privacy, see Zimmermann, 1995) and offer their public keys to buyers for encrypted and safe text communication. Silkkitie shows prices in euros and Bitcoins using the latest exchange rate. We have collected transaction data in terms of euros.

After Silkkitie users have created their user accounts on the service, they may select a role as a seller, which enables them to offer products on sale. The sellers may then set the price and describe both the product and the terms of the delivery. The seller's reputation is a central feature on Silkkitie (see Fig. 1). Buyers can give feedback by rating the product and the transaction (good: + or bad: −). Buyers see this feedback, so sellers have to keep up their reputations. More importantly, Silkkitie offers escrow service by holding the Bitcoins used for purchase and returning them to buyers if there is a lot of negative feedback. This makes it very important for sellers to build good reputations.

2.2. Web-crawling process

We employed automatic web-crawling and web-scraping techniques to extract information from the Silkkitie site every day between 5 November 2014 and 23 September 2015. This method is similar to the one that was developed to study the Silk Road's transactions (Christin,



silkkitiehdg5mug.onion Kaikki tuotteet Ostokset Viestit

2g S.A.G.E. *Sativa*

49 EUR *(0.087034 BTC)*

1 kpl varastossa

SELLER NAME (287 / -2)

Suomi → Suomi

Fig. 1. Example screenshot from Silkkitie, with the seller's positive and negative feedback in parentheses following the seller's name.

2013). From a wider perspective, we are using the Open Source Intelligence methodology (Glassman and Kang, 2012). Our method is simple and very precise. Our web-scraping software extracted the product information from every product page on Silkkitie. The collected fields were title, price, number of items in stock, seller's nickname, seller's positive and negative feedback, seller's country, buyer's country and the product page's unique URL.

We executed the software once a day over a period of almost 11 months (322 days). Once a day is enough to study day-to-day drug trade. During our data collection period, 36 days are missing due to the Silkkitie server being offline. Half of these days took place in June and July 2015. During this period, Silkkitie users raised concern that the administrator of Silkkitie, who calls himself Kapteeni ("Captain in English"), had shut down Silkkitie and run off with the remaining Bitcoins. The service continued, however, after a short break.

2.3. Measures

For this study we only used sales that took place within Finland by removing items where the seller's location was abroad or seller was willing to sell abroad. The full data includes 93 878 observations, and the data covers 286 days of transactions from 260 sellers offering 3823 products. The data includes information on products' prices, the amount in stock and additional details. Drug sales were calculated using the information on available stock. For analytic purposes, the data was organised into panel format and included 16 815 daily observations of sellers, including seller's daily drug sales, reputation, capacity and the drugs that the seller had sold during the specific day.

Seller's *daily drug sales* are the total value of all the products sold in a single day. The median amount of daily sales was 0, but the mean daily sales were 129 euros/seller. The excess of zero-value days is caused by the fact that many sellers did not make sales on a daily basis. This value is used as an outcome variable in the regression models.

Silkkitie sellers have categorized their products as stimulants, cannabis, opiates, empathogens, psychedelics, opiates, dissociatives, depressants and other pharmaceuticals (see Table 1 for the specific information). We decided to use this categorization, which seems to represent experiences that buyers try to gain from the drugs, instead of categorizations based on the chemical formations and legal status.

A seller's *reputation* quantifies the seller's combined positive and negative feedback (reputation = positive comments − negative comments). The sellers' reputations ranged from -19 to 3418. However, out of 16 815 observations, only 101 (0.6%) were under zero. Based on the 16 815 daily observations of sellers, the median capacity was 78 euros. The seller reputation figures for different types of drugs are provided in Table 2.

A seller's *capacity* is calculated for each day for each seller. It is based on the total value of the seller's reported stock in euros in a single day (capacity $= \sum_{i=0}^{n} price_i stock_i$), where n is the total number of products that the seller is offering. In other words, the seller's capacity is calculated on the basis of the daily product information of each seller. Based on the 16 815 daily observations of sellers, the median capacity was 1602 euros. Capacity figures for different types of drugs are provided in Table 2.

2.4. Statistical analyses

We first ran descriptive statistics using the full data with a total of 93 878 observations. We used Stata 12.0 to run fixed-effects Poisson regression models to analyse the daily drug sales (in euros). These models accounted for both the sellers and the day when the data was gathered, ranging from 1 (5 November 2014) to 268 (23 September 2015).

Poisson regression was selected due to the skewed continuous outcome variable. Although there are several alternatives for treating such outcomes (including a linear regression with log-transformed outcomes), Poisson regression has been argued to provide the best alternative (Santos Silva and Tenreyro, 2006). We ran the models using Huber-White standard errors (i.e., robust standard errors) due to overdispersion (Palmer et al., 2007). Another option would have been using a fixed-effects negative binomial regression, but there is risk of incidental parameter bias that does not occur with Poisson regression models (Allison, 2012).

The models first report the unadjusted effects of both reputation and availability. We used a square-root transformation for the seller's reputation and a natural logarithmic transformation for the seller's capacity to correct the skewness of these variables. The final models include drug categories as dummy variables (0 = no sales in that category, 1 = at least one sale in that category). Models were checked for multicollinearity; no problems were found in this respect, and sellers' reputations and capacities are both clear and distinct predictors of their daily sales. The results section also reports the daily sales based on the percentiles of the seller's reputation and capacity.

3. Results

The total value of the illegal drugs sold on Silkkitie during the 11 months of the study was 2 171 387 euros, and 41 131 items were sold during this period (see Table 2). Stimulants were the most valuable class of drug (538 354 euros), followed by cannabis products (487 111 euros) and empathogens (mostly MDMA; 396 127 euros). Depressants

Table 1
Drug categories in Silkkitie.

Drug category in Silkkitie	Drugs included
Stimulants	Amphetamine, cocaine, metamphetamine, A-PVP, cathinones, crack, MDPV, 2-FMA/4-FA, 4-MePPP, phenidates, nicotine, coffeine, snuff
Cannabis	Weed, hash, edibles, concentrates, leaf, synthetics
Emphatogens	Ecstasy, MDMA, MDA/MDAI, methylone
Psychedelics	LSD, DMT/AMT, mushrooms, 25x-NBOMe, ayahuasca, mescaline, 2C-x, 2C-T-x, 5-MeO-DiPT, 5-MeO-MiPT, bromo-dragonfly, LSA/LSZ, AL-LAD, Ibogaine, Dox
Other pharmaceutics	Benzodiazepines, erectile dysfunction drugs, sleeping pills, ADHD medication, pain, sedatives, muscle relaxants, MAOIs, nootropics
Opiates	Heroin, buprenorphine, oxycodone, fentanyl, tramadol, codeine, morphine, opium, metadon
Dissociatives	Ketamine, salvia, MXE/MXP, datura
Depressants	GHB/GBL, alcohol

were the least sold, with only 4975 euros of sales. Fig. 2 shows the sales values of different substances during the study. We can see that stimulants, cannabis and empathogens combine to make up over half of all sales during each month.

Out of 3823 products listed, only 51% were sold. Psychedelics had the highest proportion of sold products (64%), and opiates had the lowest (38%). The mean price of the sold products was 84 euros, and the mean daily sales amount per seller was 129 euros, but the largest daily sale by a seller was 85 000 euros. High figures in daily sales were, however, very unusual, as 5000 euros/day were exceeded only 10 times, and by only 7 sellers out of a total of 260 sellers. Considering all the sales taking place during the follow-up period of the study, almost 15% of the sellers did not have sales at all, while the highest selling 15% were responsible for 82% of the total sales in Silkkitie.

We used fixed-effects regression models to analyse the associations between a seller's daily sales and his/her reputation and capacity (Table 3). Model 0 reports the unadjusted effects on daily sales due to a seller's reputation and capacity. The models were run separately for both reputation and capacity, which were both positively associated with sales. Model 1 includes both the seller's reputation and capacity in the same model. Model 2 adds the daily sold substances as dummy variables. None of these dummy variables were statistically significant in the model at the level of $p < 0.05$; hence, they were not included in Table 3. Model 2 included 222 sellers and 16 243 observations in total, for an average of 72.9 daily observations per seller. The model is statistically significant; both seller reputation and capacity were positively associated with daily sales.

Fig. 3 shows the changes in daily sales on Silkkitie based on the percentile values of the seller's reputation and capacity (adjusted based on a fixed-effects regression model). The model includes the seller's reputation and capacity as predictors; it is adjusted for the daily sales of different substances. We can see that the (lowest) 10th-percentile group in terms of reputation has daily sales of only 62 euros on average and that the (highest) 90th-percentile group has sales of 222 euros on average. A similar linear increase is seen for seller's capacity. We can also see that the sales are especially high for the 80th and 90th percentile in terms of reputation.

4. Discussion

Our results showed that the illicit drug sales on Silkkitie were worth over 2 million euros within Finland during the 11-month study. Our main goal in this study was to analyse the impact of the seller's reputation and capacity on daily drug sales. Theoretically, the article was grounded on previous work done on trust systems (Houser and Wooders, 2006; Décary-Hétu and Leppänen, 2013; Lusthaus, 2012) and previous research showing that the seller's reputation has a positive impact on sales in drug markets (Hardy and Norgaard, 2015). We also consider the seller's daily capacity (i.e., the reported sales potential), which has been shown to have an interaction with consumer behaviour (Yu et al., 2015), but this effect has not previously been studied in a cryptomarket setting.

Our Silkkitie findings provide further empirical evidence on the importance of trust systems on cryptomarkets. Higher reputations and capacities indicated higher daily sales. Reputation and capacity represent different facets of information available for the cryptomarket consumers, and they both can be considered an integral part of the trust system in the cryptomarket. The most successful sellers have both a high reputation and a high capacity. Specifically, the capacity figures are dependent on what the sellers have reported as the available stock and we consider them as part of communication with the buyers. It has been also indicated in previous studies that cybercriminals are expected to attract companions in crime through successful communication (Décary-Hétu and Leppänen, 2013).

It is possible that a seller's reputation and capacity have become even more important due to the instability of cryptomarkets. Within our data, we could see that daily drug sales varied considerably during the study period, many products were not sold at all and that many sellers remained active for only a short period of time. Specifically, risk-averse online consumers might avoid such sellers. At least in our data, this was

Table 2
Descriptive information about sales on Silkkitie during the 11-month study period.

	Total sales (€)	Items sold	Different products on sale	% of sold products	Median product price	Median price of products sold	Median seller reputation	Median seller reputation (during the day the sale took place)	Median seller capacity per product
Stimulants	538 354	7 242	529	52	80	48	153	203	1 165
Cannabis	487 111	8 878	807	62	70	55	116	131	648
Emphatogens	396 127	8 424	236	53	99	70	368	476	3 645
Psychedelics	336 010	5 826	261	64	50	40	537	617	900
Other pharmaceuticals	180 728	6 244	1 113	50	32	23	63	67	140
Opiates	181 316	3 734	804	38	70	43	368	476	388
Dissociatives	46 766	632	38	47	70	68	1732	1778	1 878
Depressants	4 975	151	35	49	85	60	156	539	630
Total	2 171 387	41 131	3 823	51	55	39	120	139	488

Note: figures are based on 93 878 observations.

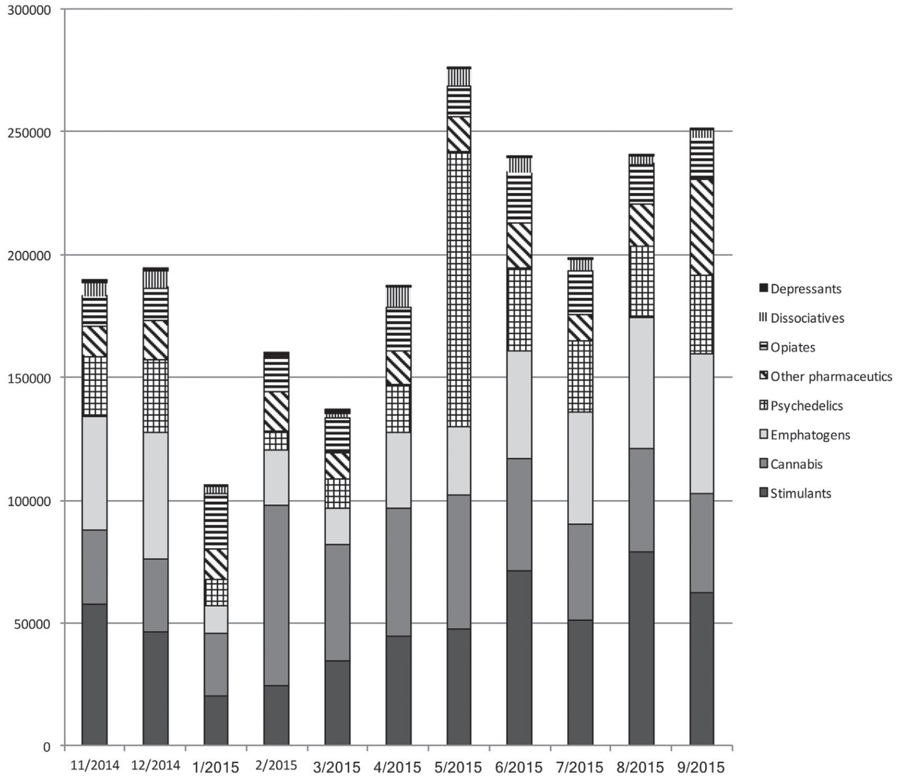


Fig. 2. Sales of different substances on Silkkitie between 5 November 2014 and 26 September 2015.

Table 3
Reputation and availability explaining daily sales on Silkkitie (based on fixed-effects Poisson regression models).

	Reputation				Capacity			
Model	Coeff.	Robust SE	Z	*P*	Coeff.	Robust SE	Z	*P*
0	0.029	0.012	2.51	0.012	0.209	0.047	4.49	< 0.001
1	0.024	0.011	2.27	0.023	0.199	0.043	4.65	< 0.001
2	0.025	0.010	2.52	0.012	0.120	0.055	2.18	0.029

Note: Model 0 shows unadjusted effects for both reputation and capacity separately. Model 1 includes reputation and capacity in the same model, and Model 2 adds the daily sold substances as controls. The final model 2 included 222 sellers and 16 243 observations (observations per seller: minimum 2, maximum 267, mean 72.9).

shown as a concentration of sales towards those sellers who had higher reputations and capacity. Results on reputation and capacity are, however, also valid for the smaller sellers and sales of a variety of drugs.

Silkkitie also provides a very interesting national case that drug markets are changed by the cryptomarkets. Although stimulants, cannabis products and empathogens (e.g., MDMA) were among the most sold products during the study, a wide variety of substances were sold. This follows the general user pattern in Finland. The drug use problem in Finland is characterised by the absence of heroin and by the high prevalence of amphetamine and buprenorphine use (EMCDDA, 2016). Cannabis, amphetamines and MDMA are the most widely used illegal substances in Finland (Hakkarainen et al., 2015). Compared to other widely popular drugs, these substances are more associated with recreational use; opioids and depressants are more popular among marginalised drug users (Pitkänen et al., 2016).

Based on the wastewater analyses, there is an estimate that the size of the Finnish drug market for stimulants (amphetamine, cocaine and methamphetamine) was around 80 million euros in 2014 (Kankaanpää et al., 2016). The cannabis market can be estimated to be somewhere between 20 million and 100 million based on estimations of cannabis use (c.f. Hakkarainen et al., 2006), and there are markets for buprenorphine and benzodiazepines, used most frequently but not only by disadvantaged drug users (Tammi et al., 2011). Considering this, Silkkitie is a major marketplace in Finland. However, most of the drugs in Finland are sold through traditional routes. It is also possible that Silkkitie provides a forum for new types of drug users who would not necessarily buy their drugs from the normal offline markets (Barratt et al., 2016a, 2016b).

According to the Finnish law enforcement, the largest vendor of Silkkitie, called Douppikauppa, was caught in April 2016 when a Finnish Customs unit intercepted a vehicle that was smuggling drugs between the Netherlands and Finland (Finnish Customs, 2016). The car contained several kilos of illegal drugs, including ecstasy tablets, LSD tabs, amphetamine and methamphetamine (Finnish Customs, 2016). This did not disrupt the drug sales on Silkkitie and Valhalla, and hence the Silkkitie case demonstrates how difficult it is to control and prevent sales in cryptomarkets. Although law enforcement could successfully capture some of the key actors involved, the markets are quickly re-established (Décary-Hétu and Giommoni, 2017).

Despite the fact that we were able to collect data for almost 11 months, we would like to mention some limitations of our study. We performed data extraction once a day, but on some occasions, we failed to get this data because the Silkkitie server was offline. In addition, we

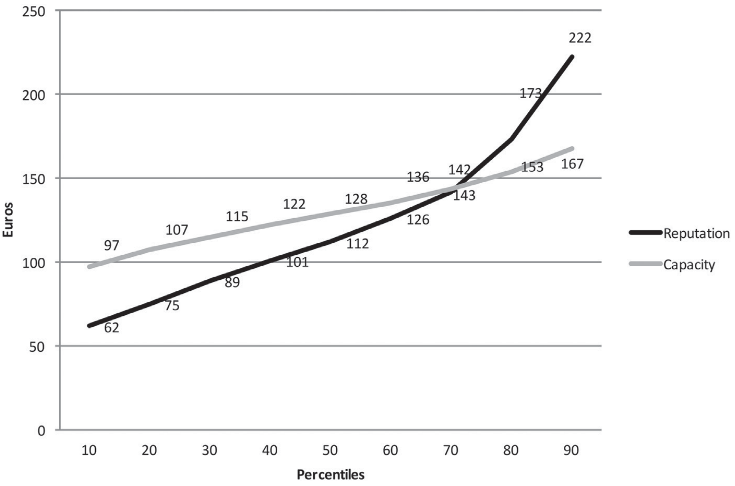


Fig. 3. Change in daily sales on Silkkitie based on percentile values of reputation and availability (adjusted predictions based on fixed-effects regression models).

might have missed some information if a seller removed a product from the marketplace between the measurements. Similarly, it is possible that an old product could be registered as new because the seller changed the product title. We believe, however, that these limitations are not critical for our main analyses, which concentrated on seller reputation and capacity. Furthermore, our data is extensive, so the missing data is unlikely to have an impact on the analyses. To our knowledge, our data set is one of the largest one that has been collected from the cryptomarkets. Furthermore, our data set is not vulnerable to the challenges concerning the collection of online data from other platforms, such as the Silk Road. In the Silk Road, for example, incomplete crawls may lead to misleading results (Munksgaard et al., 2016).

In this study, we were able to follow the extent of trading and produce almost real-time statistics about how people buy and sell illegal drugs. Our method is highly reliable, and the long period of data collection guarantees this reliability. Hence, this study contributes to the increasing body of literature studying cryptomarkets with web crawlers (e.g., Aldridge and Décary-Hétu, 2016; Christin, 2013; Demant et al., 2016; Hardy and Norgaard, 2015). Most important, our results help reveal the consumer interactions between sellers and buyers on online cryptomarkets, and our findings underline the relevance of both reputation and capacity as parts of the cryptomarket trust system. Future studies should continue to collect longitudinal data sets.

Conflict of interest

Authors report no conflict of interest.

Contributors

All the authors are responsible for the reported research, and all have contributed to the article (concept, research design, analysis). All the authors have drafted and revised different versions of the article and have approved the final manuscript as submitted. Juha Nurmi collected the data, designed the study with Atte Oksanen and wrote significant parts of the article. Teemu Kaskela took part in writing of the article and he run part of the analysis with Juha Nurmi. Jussi Perälä took part in commenting and writing of the article. Atte Oksanen designed the study with Juha Nurmi, run analyses and wrote significant parts of the article.

Acknowledgements

We would like to thank Dr. Monica Barratt for constructive comments, which helped us to improve the manuscript.

References

Aldridge, J., Décary-Hétu, D., 2016. Hidden wholesale: the drug diffusing capacity of online drug cryptomarkets. Int. J. Drug Policy 35, 7–15.

Allison, P., 2017. Beware of software for fixed effects negative binomial regression [Article on the Internet]. City: Publisher; Year [updated 8 June 2012; cited 4 August 2016]. Available from: http://statisticalhorizons.com/fe-nbreg.

Bancroft, A., Reid, P.S., 2017. Challenging the techno-politics of anonymity: the case of cryptomarket users. Inf. Commun. Soc. 20, 497–512.

Barratt, M.J., Maddox, A., 2016. Active engagement with stigmatised communities through digital ethnography. Qualit. Res. 16, 701–719.

Barratt, M.J., Ferris, J.A., Winstock, A.R., 2014. Use of Silk Road, the online drug marketplace, in the United Kingdom, Australia and the United States. Addiction 109, 774–783.

Barratt, M.J., Ferris, J.A., Winstock, A.R., 2016a. Safer scoring? Cryptomarkets, social supply and drug market violence. Int. J. Drug Policy 35, 24–31.

Barratt, M.J., Lenton, S., Maddox, A., Allen, M., 2016b. 'What if you live on top of a bakery and you like cakes?'—drug use and harm trajectories before, during and after the emergence of Silk Road. Int. J. Drug Policy 35, 50–57.

Bohannon, J., 2016. Bitcoin busts. Sci. 351, 1144–1146.

Broséus, J., Rhumorbarbe, D., Mireault, C., Ouellette, V., Crispino, F., Décary-Hétu, D., 2016. Studying illicit drug trafficking on Darknet markets: structure and organisation from a Canadian perspective. Forens. Sci. Int. 264, 7–14.

Caudevilla, F., Ventura, M., Fornís, I., Barratt, M.J., Vidal, C., Quintana, P., Lladanosa, C.G., Muñoz, A., Calzada, N., 2016. Results of an international drug testing service for cryptomarket users. Int. J. Drug Policy 35, 38–41.

Christin, N., 2013. Traveling the Silk Road: a measurement analysis of a large anonymous online marketplace. In: WWW '13 Proceedings of the 22nd International Conference on World Wide Web. ACM. New York. pp. 213–224.

Décary-Hétu, D., Giommoni, L., 2017. Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of Operation Onymous. Crime Law Soc. Change 67, 55–75.

Décary-Hétu, D., Leppänen, A., 2013. Criminals and signals: an assessment of criminal performance in the carding underworld. Security J. 29, 442.

Demant, J., Munksgaard, R., Houborg, E., 2016. Personal use, social supply or redistribution? Cryptomarket demand on Silk Road 2 and Agora. Trends Organized Crime. http://dx.doi.org/10.1007/s12117-016-9281-4.

Dingledine, R., Mathewson, N., Syverson, P., 2004. Tor: The Second-Generation Onion Router. Naval Research Lab, Washington, DC.

Dolliver, D.S., Kenney, J.L., 2016. Characteristics of drug vendors on the Tor network: a cryptomarket comparison. Vict. Offenders(May (2)). http://dx.doi.org/10.1080/15564886.2016.1173158.

Dolliver, D.S., Kuhns, J.B., 2016. The presence of new psychoactive substances in a Tor network marketplace environment. J. Psychoact. Drugs 48, 321–329. http://dx.doi.org/10.1080/02791072.2016.1229877.

Dolliver, D.S., 2015. Evaluating drug trafficking on the tor network: silk road 2, the sequel. Int. J. Drug Policy 26, 1113–1123.

European Monitoring Centre for Drugs and Drug Addiction. Finland country overview. Lisbon: EMCDDA [cited 4 August 2016]. Available from: http://www.emcdda.

europa.eu/countries/finland.

FBI, 2014. Sealed Complaint against Ross Ulbricht. Narcotic Trafficking Conspiracy. https://www.documentcloud.org/documents/801103-172770276-ulbricht-criminal-complaint.html.

Finnish Customs, 2016. Pre-Investigation Reports of the Douppikauppa Case. 9010/R/6369/15; 9010/R/7463/15; 9010/R/2757/16; 9010/R/3004/16. Finnish Customs, Tampere.

Glassman, M., Kang, M.J., 2012. Intelligence in the internet age: the emergence and evolution of open source intelligence (OSINT). Comp. Hum. Behav. 28, 673–682.

Hakkarainen, P., Kainulainen, H., Perälä, J., 2006. Suomen kannabismarkkinat – paljonko pilveä palaa?: [Measuring the size of the cannabis market in Finland]. Yhteiskuntapolitiikka 71, 583–595.

Hakkarainen, P., Karjalainen, K., Ojajärvi, A., Salasuo, M., 2015. Huumausaineiden ja kuntodopingin käyttö ja niitä koskevat mielipiteet Suomessa vuonna 2014 [Drug use, doping and public opinion in Finland: results from the 2014 Drug Survey]. Yhteiskuntapolitiikka 80, 319–333.

Hardy, R.A., Norgaard, J.R., 2015. Reputation in the internet black market: an empirical and theoretical analysis of the deep web. J. Inst. Econ. 12, 515–539.

Houser, D., Wooders, J., 2006. Reputation in auctions: theory, and evidence from eBay. J. Econ. Manag. Strateg. 15, 353–369.

Kankaanpää, A., Ariniemi, K., Heinonen, M., Kuoppasalmi, K., Gunnar, T., 2016. Current trends in Finnish drug abuse: wastewater based epidemiology combined with other national indicators. Sci. Total Environ. http://dx.doi.org/10.1016/j.scitotenv.2016.06.060.

Keipi, T., Oksanen, A., 2014. Self-exploration, anonymity and risks in the online setting: analysis of narratives by 14–18-year olds. J. Youth Stud. 17, 1097–1113.

Kristoufek, L., 2015. What are the main drivers of the Bitcoin price? Evidence from wavelet coherence analysis. PLoS One 10, e0123923.

Lusthaus, J., 2012. Trust in the world of cybercrime. Global Crime 13, 71–94.

Maddox, A., Barratt, M.J., Allen, M., Lenton, S., 2016. Constructive activism in the dark web: cryptomarkets and illicit drugs in the digital 'demimonde'. Inf. Comm. So. 19, 111–126.

Martin, J., 2014a. Drugs on the Dark Net: How Cryptomarkets Are Transforming the Global Trade in Illicit Drugs. Palgrave Macmillan, New York, NY.

Martin, J., 2014b. Lost on the Silk Road: online drug distribution and the 'cryptomarket'. Criminol. Crim. Justice 14, 351–367.

Munksgaard, R., Demant, J., 2016. Mixing politics and crime–The prevalence and decline of political discourse on the cryptomarket. Int. J. Drug Policy. 35, 77–83.

Munksgaard, R., Demant, J., Branwen, G., 2016. Replication and methodological critique of the study 'Evaluating drug trafficking on the Tor Network'. Int. J. Drug Policy 35, 92–95.

Nakamoto, S., 2008. Bitcoin: A Peer-to-peer Electronic Cash System. [cited 4 August 2016]. Available from: http://bitcoin.org/bitcoin.pdf.

Palmer, A., Losilla, J.M., Vives, J., Jimenez, R., Llorens, N., 2007. Overdispersion in the Poisson regression model a comparative simulation study. Methodology 3, 89–99.

Pitkänen, T., Perälä, J., Tammi, T., 2016. Huumeiden käyttäjiä on monenlaisia: kahdensadan helsinkiläisen huumeiden aktiivikäyttäjän elämäntilanne ja päihteiden käyttö [Different kinds of drug users: life situation and substance use of two hundred active drug users in Helsinki]. Tietopuu, Tutkimussarja, pp. 1–10 (Finnish).

Resnick, P., Kuwabara, K., Zeckhauser, R., Friedman, E., 2000. Reputation systems. Comm. ACM. 43, 45–48.

Rhumorbarbe, D., Staehli, L., Broséus, J., Rossy, Q., Esseiva, P., 2016. Buying drugs on a Darknet market: a better deal? Studying the online illicit drug market through the analysis of digital, physical and chemical data. Foren. Sci. Int. 267, 173–182.

Santos Silva, J.M., Tenreyro, S., 2006. The log of gravity. Rev. Econ. Stat. 88, 641–658.

Spears, R., Postmes, T., Lea, M., Wolbert, A., 2002. When are net effects gross products? The power of influence and the influence of power in computer-mediated communication. J. Soc. Issues 58, 91–107.

Tammi, T., Pitkänen, T., Perälä, J., 2011. Stadin nistit – huono-osaisten helsinkiläisten huumeidenkäyttäjien päihteet sekä niiden käyttötavat ja hankinta [Disadvantaged drug users in Helsinki: what drugs do they use, how do they use them and how do they get them]. Yhteiskuntapolitiikka 76, 45–54.

The Guardian, 2013. Tor: 'the King of High-secure, Low-Latency Anonymity' [updated 4 Oct 2013; Cited 4 August 2016]. Available from: http://www.theguardian.com/world/interactive/2013/oct/04/tor-high-secure-internet-anonymity.

Van Buskirk, J., Naicker, S., Roxburgh, A., Bruno, R., Burns, L., 2016. Who sells what? Country specific differences in substance availability on the Agora Cryptomarket. Int. J. Drug PolicyJul 20. http://dx.doi.org/10.1016/j.drugpo.2016.07.004.

Van Hout, M.C., Bingham, T., 2013a. 'Silk Road', the virtual drug marketplace: a single case study of user experiences. Int. J. Drug Policy 24, 385–391.

Van Hout, M.C., Bingham, T., 2013b. 'Surfing the Silk Road': a study of users' experiences. Int. J. Drug Policy 24, 524–529.

Van Hout, M.C., Hearne, E., 2017. New psychoactive substances (NPS) on cryptomarket fora: an exploratory study of characteristics of forum activity between NPS buyers and vendors. Int. J. Drug Policy 40, 102–110.

Yu, M., Kapuscinski, R., Ahn, H.S., 2015. Advance selling: effects of interdependent consumer valuations and seller's capacity. Manag. Sci. 61, 2100–2117.

Zimmermann, P., 1995. Pretty Good Privacy: Public Key Encryption for the Masses. Building in Big Brother. Springer-Verlag, New York, NY.

PUBLICATION

IV

Tor De-anonymisation Techniques
J. Nurmi and M. S. Niemelä

11th International Conference on Network and System Security. Ed. by 2017, 657–671

Publication reprinted with the permission of the copyright holders

Tor De-anonymisation Techniques

Juha Nurmi[1,2]( ) and Mikko S. Niemelä[1]

[1] Kinkayo Pte Ltd, Singapore, Singapore
mikko@kinkayo.com
[2] Laboratory of Pervasive Computing, Tampere University of Technology,
Tampere, Finland
juha@kinkayo.com

Abstract. Tor offers a censorship-resistant and distributed platform that can provide easy-to-implement anonymity to web users, websites, and other web services. Tor enables web servers to hide their location, and Tor users can connect to these authenticated hidden services while the server and the user both stay anonymous. However, throughout the years of Tor's existence, some users have lost their anonymity. This paper discusses the technical limitations of anonymity and the operational security challenges that Tor users will encounter. We present a hands-on demonstration of anonymity exposures that leverage traffic correlation attacks, electronic fingerprinting, operational security failures, and remote code execution. Based on published research and our experience with these methods, we will discuss what they are and how some of them can be exploited. Also, open problems, solutions, and future plans are discussed.

1 Introduction

Anonymity is considered an important right for supporting freedom of speech and defending human rights. An Internet user can use various tools to hide his or her identity [1]. Among these, the most popular tool is Tor. It is used by two million people every day, including ordinary citizens concerned about their privacy, corporations who do not want to reveal information to their competitors, and law enforcement and government intelligence agencies who need to carry out operations on the Internet without being noticed [2]. Furthermore, human rights activists and journalists communicate anonymously using Tor to protect their lives [3].

Tor provides anonymity by routing the user's traffic through three separate relay servers so that it is hard to reveal the user's physical location or IP address. This technique is called *onion routing* [4]; it means that Tor protects users through encryption to ensure privacy, authentication between clients and relays, and signatures to ensure that all clients know the same set of relays.

The Tor network is considered to be a well-studied and very secure communication network [2]. According to top secret National Security Agency (NSA) documents disclosed by whistleblower Edward Snowden, a former Central Intelligence Agency (CIA) employee and a former NSA contractor, the Tor network has been too difficult for the NSA and CIA to spy on. The NSA even wrote

Z. Yan et al. (Eds.): NSS 2017, LNCS 10394, pp. 657–671, 2017.
DOI: 10.1007/978-3-319-64701-2_52

in their top secret documents that Tor is "the King of high-secure, low latency Internet anonymity" [5].

However, these documents also reveal that some users can sometimes be de-anonymised. "We will never be able to de-anonymize all Tor users all the time. With manual analysis we can de-anonymize a very small fraction of Tor users, however no success de-anonymizing a user in response to a TOPI request/on demand." This means that when this document was written in June 2012, the NSA had not been able to discover the identity of Tor users that it wanted to specifically target.

In addition, using special networks such as the Tor network, it is possible to run web servers anonymously and without fear of censorship [4]. Servers configured to receive inbound connections through Tor are called hidden services (HSs); rather than revealing the real IP address of the server, an HS is accessed through the Tor network by means of the virtual top-level domain .onion [4]. As a result, the published content is diverse [6,7]. Undoubtedly, some HSs share pictures of child abuse or operate as marketplaces for illegal drugs, including the widely known black market Silk Road. These few services are obviously controversial and often pointed out by critics of Tor and anonymity, but a vast number of HSs are devoted to human rights, freedom of speech, journalism, and information prohibited by oppressive governments.

The Tor Browser is as easy to use as any common web browser and uses a Mozilla Firefox Extended Support Release (ESR). Similarly, deploying a hidden website is simple. Tor offers a censorship-resistant and distributed platform that can provide easy-to-implement anonymity to web users, websites, and other web services.

2 Background

In this chapter, we give a basic overview of how Tor protects anonymity and what the known design flaws of these techniques are. In particular, we study what research reveals to us about the design of Tor and applications on top of Tor.

2.1 Onion Routing

Onion routing was patented by U.S. Naval Research Laboratory (NRL) researchers Paul Syverson, Michael G. Reed, and David Goldschlag (US patent 6266704 2001-07-24) [1,7]. The first version of onion routing was deployed by researchers at the NRL in the 1990s to protect online intelligence activity; it developed further under the Defense Advanced Research Projects Agency (DARPA) [2,4]. The source code of Tor software was released under a so-called BSD open-source license, and a non-profit organization, the Tor Project, was founded in 2006. Since then, Tor has been freely available for everyone [3].

The Tor Project, along with Tor design papers [4], warns that onion routing is not protected against an attacker who can follow both traffic going into and coming out of the Tor network [8–11]. In this case, the research community knows no practical low-latency design that prevents traffic and timing correlation attacks [10,12–14].

2.2 Privacy-Aware Applications on Top of Tor

Exploiting network protocols is not a very common way to attack modern systems. After all, for instance, it is fairly easy to find exploits against web frameworks and applications but extremely hard to find an exploit against a TCP/IP stack. Similarly, it is hard to attack a Tor application or networking protocol. However, on top of Tor, people use applications that are likely to be targeted by an attacker [10,12].

The Tor Project offers applications on top of Tor, such as a special web browser and a messenger application [3]. Tor Messenger is a cross-platform chat program that aims to be secure by default and sends all of its traffic over Tor. Tor Browser is the main privacy-aware application produced by the Tor Project. Tor Browser is a modified Mozilla Firefox browser with best-practice default settings and extensions, such as NoScript and HTTPS Everywhere. Tor Browser automatically starts Tor background processes and routes all traffic through the Tor network. In addition, this browser takes a lot of effort to remove all possible fingerprinting methods; it fakes the information about operating system and hardware. Tor Browser does not save privacy-sensitive data, such as the browsing history, cache, or cookies.

According to security professional Bruce Schneier's analysis of the leaked NSA documents, the NSA seems to be individually targeting Tor users, by exploiting vulnerabilities in their Firefox-based Tor Browser, and not the Tor application directly [15]. Exploiting Tor Browser is difficult but much more straight forward than exploiting Tor itself. Web browsers are complex combinations of features and software libraries. Firefox has security flaws and rapid application development that reflects to Tor Browser. As a result, the NSA uses a series of native Firefox vulnerabilities to attack users of Tor Browser [15]. According to the training presentation provided by Snowden, the so-called EgotisticalGiraffe (The NSA code name) exploits an XML extension for JavaScript. This vulnerability existed in Firefox 11.0 (year 2012) [15]. According to another document, the vulnerability exploited by EgotisticalGiraffe was inadvertently fixed, but the NSA was confident that they would be able to find new exploits against the Firefox and Tor Browsers [15]. It is clear that the applications on top of Tor are under attack.

2.3 Pseudonyms and Operational Security

Operational security (OPSEC) is the process of protecting individual pieces of non-critical data that could be grouped together to reveal critical data. Over time, it is difficult to understand how much one's cumulative online behavior reveals through the eyes of an adversary. Tor users should follow a strict OPSEC process by protecting all information that could be used against their anonymity.

Sun Tzu wrote, "The art of war teaches us to rely not on the likelihood of the enemy's not coming, but on our own readiness to receive him; not on the chance of his not attacking, but rather on the fact that we have made our position unassailable."[1]

We are our own worst enemy. It is too easy to give away the advantage of perfect technical anonymity by sharing a combination of identifying information. Any information a Tor user shares in the public domain is also vulnerable to de-anonymisation. The attacker may follow a Tor user's behavior pattern and gather cumulative information. The user may think that it looks like he/she merely reads some random news and writes some random comments under a pseudonym, but, of course, this behavior is not random and could be linked to the real identity. Shared information may reveal language, probable time zone, interests, knowledge, and – as we will show later – clear links between the real name of the user and the user's pseudonym identity.

3 Tor De-anonymisation Techniques

In this chapter, we give a basic overview of how, throughout the years of Tor's existence, some users have lost their anonymity.

There is controversial content published using Tor – for instance, dark markets and child abuse material. As a result, law enforcement agencies have been using a range of state-of-the-art exploits to remove the cover from some users of the Tor network. These methods include exploitation of human errors as well as highly sophisticated mathematical methods that exploit software flaws. In addition, operational security failures have led to de-anonymisation.

3.1 Operational Security is Difficult

Here, we present the most famous example in which an attacker followed a Tor user's behavior pattern and gathered critical cumulative information. In this case, the OPSEC process failed to protect individual pieces of non-critical data. These data were grouped together to reveal the identity of Ross William Ulbricht, who was known under the pseudonyms "Dread Pirate Roberts" (DPR), "frosty", and "altoid" [16]. He was convicted of creating and running the Silk Road dark market onion site until his arrest in October 2013.

On 11 October 2011, a user called altoid posted publicly on the Bitcoin Talk forum [17]. The message, titled "a venture backed Bitcoin startup company", asked for help to build a Bitcoin startup company. Altoid asked people to contact rossulbricht@gmail.com. Moreover, altoid talked about the new market service Silk Road. Simultaneously, a user also going by the name of altoid advertised Silk Road on a forum at shroomery.org, which is a magic mushroom discussion board.

Ulbricht's Google Plus page and his YouTube profile both make multiple references to the Austrian economic theory site called the Mises Institute.

[1] The Annotated Art of War, Parts 8.3-11: Advantages.

On the Silk Road forums, DPR shared links to the Mises Institute. DPR cited Austrian economic theory and shared the Mises Institute's material. Furthermore, DPR mentioned that he is in the Pacific time zone.

Ulbricht posted the question "How can I connect to a Tor hidden service using curl in php?" to a popular site called stackoverflow.com. According to the criminal complaint, Ulbricht posted the question using his own real name. Less than one minute later, he changed his username to frosty [16].

Finally, on top of everything else, Ulbricht purchased nine counterfeit identification documents with his face but with different names. The package traveled from Canada to the US, and it was intercepted by US border customs. The package was addressed to Ulbricht's San Francisco apartment. Obviously, law enforcement suspected criminal activity. Simultaneously, technical investigations of Silk Road gathered evidence that DPR lives in San Francisco. We have gathered these events into one timeline in Table 1.

Table 1. The main OPSEC-related events that linked Ulbricht to Silk Road through the pseudonyms altoid, frosty, and DPR.

Date	OPSEC: leak of critical information
01/2011	Silk Road HS http://tydgccykixpbu6uz.onion is created
01/2011	Silk Road portal silkroad420.wordpress.com is created
01/2011	Silk Road portal starts to advertise Silk Road HS
01/2011	altoid advertises Silk Road on shroomery.org forum
01/2011	altoid advertises Silk Road on Bitcointalk forum
10/2011	altoid posts a job offer on Bitcointalk, rossulbricht email
03/2013	Question about Tor and PHP is posted on stackoverflow.com
03/2013	Ulbricht changes his real name on Stack Overflow to "frosty"
07/2013	A routine border search intercepts a package of fake IDs

Eventually, the FBI closed in on the suspect Ross Ulbricht. Using warrants and technical investigation and by following Ulbricht, the FBI arrested him and seized his open laptop in a public library. Access to this open laptop provided plenty of evidence to convict Ulbricht as the creator of Silk Road market place.

3.2 Attacks Against Tor Network-Affiliated Systems

Tor protocol is just one service that a client or server might be running. Tor network affiliated systems are still vulnerable to conventional cyberattacks. Depending on the configuration and exposure of the system, several techniques might be used to reveal the true identity of a server or actor operating in Tor. Actual de-anonymisation takes place after the attacker acquires relevant data from or takes full control of the target system. Depending on the information available, any identifier or the configuration files of the system can be used to de-anonymise.

Any visible service increases the exposure of the system and therefore the probability of a cyberattack. At the application level, typical attacks are input validation, session handling and access control attacks. At the operating system level, attacks generally target misconfiguration. System performance can also be compromised with denial-of-service attacks that can lead to system failure or crash. The purpose of performance attacks is to cause a change in the system's state that either reveals the details of the system or allows access because of a safe mode or other recovery measure.

Typical input validation attacks are based on injection and include cross-site scripting (XSS), buffer overflows and malicious file uploads [18].

Session handling attacks, which target tokens that are exchanged during communication to ensure a correct state in both endpoints, include token value eavesdropping, token value guessing and session fixation [19].

Access control attacks focus on privilege escalation, where a normal user will be upgraded to an administrator-level user or a user with other privileges [20].

Operating system-level attacks focus on misconfiguration using default user credentials, administrator interface disclosure and direct object reference [21].

Performance attacks use denial-of-service techniques and distributed denial-of-service techniques where more than one host participates to take down the target [22].

In August 2013, the FBI exploited a memory-management vulnerability in the Firefox/Tor Browser to turn Freedom Hosting sites into malware-spreading trackers [23]. A hosting operator, Freedom Hosting facilitated child abuse HS websites on a massive scale. The FBI accessed the servers of Freedom Hosting and injected a malicious JavaScript exploit code. The JavaScript code looks for a MAC address and hostname and sends them back as HTTP requests to the Virginia server to expose the user's real IP address [23]. Later, Firefox developer Mozilla patched the underlying vulnerability [24].

On 29 November 2016, an anonymous writer sent a warning to the popular tor-talk mailing list [25]. This email published findings from a Tor HS that was sharing child abuse material. A piece of malicious JavaScript code found from the site exposed the user's real IP address. This exploit was similar to the one that was used in 2013. It was able to call kernel32.dll on Windows operating systems and execute the attacker's commands. Mozilla quickly fixed this vulnerability [26].

Another option is to offer a file to the user to be opened outside Tor Browser. Web browsers are not able to open all types of files, and the server can offer malicious files to be downloaded. In this case, Tor Browser clearly warns that some types of files can cause applications to connect to the Internet without Tor. The attacker does not need to find new exploits against software; instead, the attacker can use the features of applications that are not designed for protecting privacy. For instance, many document viewers make connections to online sources to, for example, download images or stylesheets from the web.

Nevertheless, it is not clear which file formats and applications have no effect on Tor users' privacy. For example, a simple folder of mp3 files may include a popular m3u file, which enables many major media players to look up the album image from the online source pointed to from inside the m3u file. The attacker can share this kind of music album with targets and can see the incoming connections from their real IP addresses. The most popular free media players, such as VLC player, have been suffering from this kind of unintentional privacy leak [27].

3.3 Attacks on Hidden Services

In this section, we examine a few common mistakes and flaws that may reveal critical information about a hidden service. It is possible to deploy any TCP service on Tor using and onion address. We concentrate on websites and secure shell (SSH) services.

First, let us examine the SSH service that works through the Tor network using an onion address. An SSH service is a typical way to offer remote login to Linux or Unix machines. An SSH shows an unique fingerprint of the service before login. This makes it possible to check that the SSH service really is the one it should be and that there is no man-in-the-middle attack on the network. However, if the user offers the same SSH service on a public IP address and through an onion address, this reveals the IP address of the hidden service. Here is a demonstration of this de-anonymisation technique:

```
# torsocks ssh root@msydqstlz2kzerdg.onion
RSA key fingerprint a7:93:84:a6:97:fa:25:65:77:c9:58:bb:fe:8e:e2:2f
# ssh root@ahmia.fi
RSA key fingerprint a7:93:84:a6:97:fa:25:65:77:c9:58:bb:fe:8e:e2:2f
```

As a result, we can be sure that the SSH server on ahmia.fi and on msydqstlz2kzerdg.onion are the same. We have just revealed the real address of the hidden service. It is fairly easy to make this unintentional configuration and share the same service simultaneously through a public IP address and an onion address.

It is important to understand that Tor is listening to SOCKS connections on a localhost port. This means that any software that uses Tor is connecting to localhost. Because the software thinks that the connections are coming from localhost, a new danger for anonymity is exposed: many web frameworks treat localhost as a safe zone. A typical example is the very popular Apache HTTP Server with the Apache Server Status module, which comes activated by default to localhost connections http://127.0.0.1/server-status/. Normally, this is a safe configuration because localhost is usually a safe zone, and only the users with login access to the server can access this server-status page. However, with Tor and an onion address, the connection to this page through Tor comes in to Apache from the localhost, and Apache displays the page http://someHSaddress.onion/server-status/ publicly. These kinds of services can lead to de-anonymisation of the hidden service.

In addition, it is possible to exploit unintentional features of popular web frameworks that are not designed to be installed on Tor. A tiny mistake or software error may lead to a critical information leak. This happened to the notorious Silk Road marketplace. It leaked its real IP address in an error situation.

It is extremely hard to test every possible error situation and anticipate how web software may perform under malfunction circumstances. Tor adds an extra layer of complexity to web application security.

3.4 Traffic and Timing Correlation Attacks

Tor does not provide protection against end-to-end timing attacks. An attacker observing traffic to the first relay (entry guard) and traffic to the destination (onion site, exit relay) can use statistical analysis to discover that they are part of the same circuit. In this scenario, Tor does not provide absolute anonymity; the client address and destination address of the traffic are known to the adversary who, through correlation attacks, effectively de-anonymises the client [14]. Note that the attacker does not necessarily need to control the first and the last router in a Tor circuit to correlate streams observed at those relays. It is enough that the attacker is able to observe the traffic.

Sometimes, the de-anonymisation does not require sophisticated statistical analysis. For example, a 20-year-old Harvard University student was arrested and charged with allegedly sending hoax bomb threats using Tor to get out of a final exam [28]. According to the FBI affidavit [29], the investigators found that these emails came from Guerrilla Mail, a free email service that creates temporary email addresses. Guerrilla Mail embeds the sending IP address in every outgoing email, and in this case, this was the IP address of one of the Tor exit nodes. The FBI noted that one student had been using Tor from the university wireless network shortly before and while the emails were being sent. The correlation led the FBI to interrogate the student, who confessed and was arrested.

As we can see, a traffic and timing correlation attack is easy when the anonymity set (number of clients) is small. Several research papers have concluded that the degree of anonymity is small if the number of clients in the anonymity system is small [30,31]. The content or context of an anonymous message may reveal background information and reasons to suspect that the Tor user is using a certain network. If there are only few people using Tor (the anonymity set is small) on that network, then it is possible to suspect one of them. However, in the case of Tor, there is still plausible deniability, and common police work is needed to investigate the suspects.

More complex attacks require sophisticated statistical analysis of traffic and timing. Research shows that these methods may reveal some Tor users and HSs [10,12–14].

Finally, we demonstrate an example of a traffic correlation attack that can be executed on a real Tor network. To do this, we have created a set of onion services and selected the entry guards for them. For ethical reasons, these guard relays were installed just for us, and they are not real guard relays for anyone else.

Each onion service is serving a website to make normal HTTP traffic between the HS and Tor users possible.

We observe the network traffic of 10 Tor relays, which our test set of 100 HSs are using as the first hop to the Tor network. Simultaneously, we are shaping a distinguishable HTTP traffic pattern to the HSs. As expected, we can clearly see the traffic pattern between the HS and the entry guard traffic, and this reveals to us the real IP address of the HS. Figure 1 shows our test with the onions, the entry guards, and the point of passive traffic analysis.

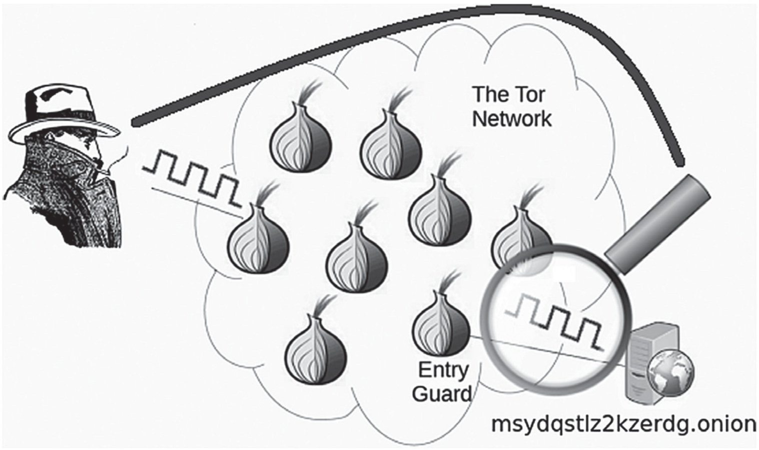


Fig. 1. The attacker sends a distinguishable traffic pattern to a hidden service. Simultaneously, the attacker is able to monitor the network between the hidden service's first hop to the Tor network, the so-called entry guard. As a result, it is possible to reveal the real IP address of the hidden service.

As a result, we detected without special effort the traffic pattern to the real IP address of the hidden service. This kind of traffic correlation is a known problem for the Tor network. The attacker does not need to operate the entry guard itself; only the traffic is relevant. Furthermore, the mechanisms of the connection, such as the length of the circuits and hidden service protocol, are not relevant to this attack. If any anonymity network provides TCP connectivity, there is always the possibility of this type of attack. This type of attack can be weaponized against HSs. It is reasonably possible that some intelligence services are already auditing connections to the Tor network. For instance, according to classified documents leaked by Edward Snowden, The British Government Communications Headquarters (GCHQ) extracts data from major fibre-optic cables to be processed and searched at a later time [31].

Obviously, intelligence agencies can catch only a fraction of the Tor entry traffic, and there are no technical limitations for monitoring. An intelligence service might monitor guard traffic and simultaneously send shaped traffic to all known HSs. As a result, an intelligence service would be able to de-anonymise a small random fraction of HSs.

Fortunately for the security of Tor, there are over two million Tor users every moment and almost 60 000 HSs. This large anonymity set makes it extremely hard to target a certain HS, because the attacker cannot know if the targeted HSs are connected to the entry guards that the attacker is able to monitor [14].

4 Results

In this paper, we have presented hands-on demonstrations of anonymity exposures that leverage traffic correlation attacks, electronic fingerprinting, operational security failures, and remote code executions. Based on published research and our experience with these methods, we showed how some of the security limitations of the Tor network can be exploited. Finally, we present analysis of these realistic de-anonymisation techniques. We recognize four different fields of security. Figure 2 shows a 22 matrix of these fields of security and their technical dimensions and responsibilities.

I. It is the responsibility of the Tor Project to deploy secure software,
II. and offer clear tutorial material with examples.
III. It is the responsibility of the user to read and follow these guidelines,
IV. and understand basic operation security principles.

I. We demonstrated anonymity exposures that leverage traffic and timing correlation attacks. The Tor Project tries to decrease the probability of this kind of technical de-anonymisation. Fortunately for anonymity, The Tor network has a large number of varied users, and possible correlation attacks usually require global network piercing monitoring for every major network.

II. We showed that enabling JavaScript or downloading files can cause Tor Browser or related software to leak identifying information about the user. The Tor Project should make it as easy as possible to follow safe usage of privacy-aware applications. In many cases, these are not technical but practical features for instance, clear user interface and notifications.

III. We demonstrated cases of off-hand usage of Tor that have led to de-anonymisation of hidden services and users. In particular, a user who installs services on an HS should carefully follow guidelines and understand the technical details of the setup. For instance, the same SSH service should not be available on a public IP address and through an onion address. Otherwise, technical de-anonymisation is possible.

IV. We presented operational security failures; in particular, the founder of Silk Road, Ross Ulbricht, did not have a clear OPSEC process to protect his real identity. This type of de-anonymisation is a non-technical process, and the user could have followed clear guidelines to separate his pseudonymous identities from his real identity.

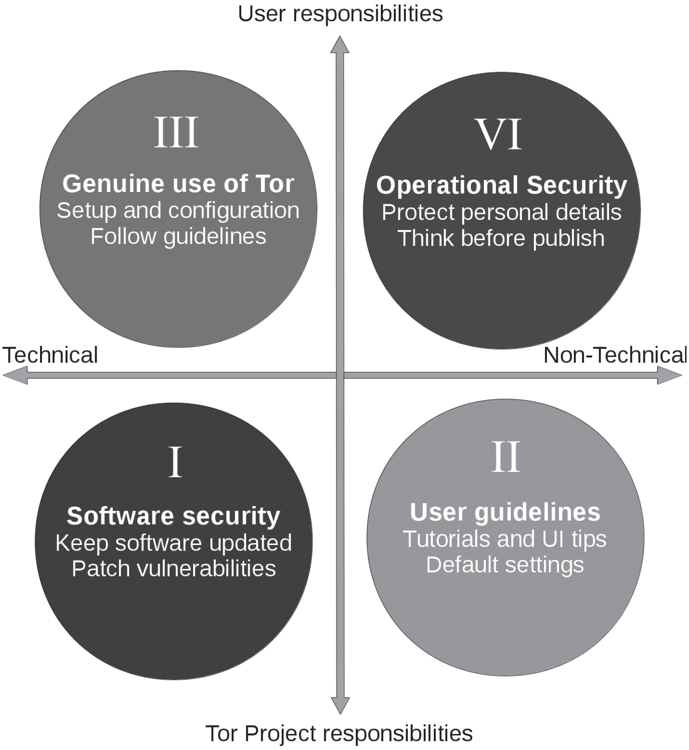


Fig. 2. The Tor Project develops software, including Tor and Tor Browser, and provides user-friendly applications for free. Tutorial materials help the user to protect privacy and anonymity. However, the Tor user must understand how Tor should be used and understand operational security.

5 Conclusion

Providing a usable anonymising network on the Internet today is an ongoing challenge for Tor. Fortunately for Tor's anonymity, the network has a great deal of various users: two million client users, 60 000 hidden services, and 7000 voluntary Tor relays creates a massive anonymity set and degree of anonymity.

If any type of anonymity network provides TCP connectivity, there is always the possibility of weaponized traffic and timing correlation attacks. There is no attempt in onion routing to remove the correlation between incoming and outgoing messages. However, this type of targeted attack (in which the attacker can select the user or hidden service) would require global piercing network monitoring for every major network. There is no hint that even the largest intelligence services have this capability.

Unintentional configurations and leaky web software are problematic. Most web frameworks are not tested or designed to be used as a hidden service. This may result in the leak of the real IP address of the server.

The weakest link of anonymity is the user. The Tor Project offers extensive tutorials and guidelines for online anonymity. However, we have seen that even technically talented users sometimes fail to follow obvious OPSEC rules or simply make mistakes and leak their real identity.

6 Discussion

Tor Browser is not the only method for using Tor; there are other security options. These tools could become very popular too, although they are more difficult to install and more complex to understand. Instead of using Tor Browser, a user can use Linux-based security-focused operating systems. Anonymous operating systems are significantly safer to use because the attacker cannot de-anonymise the user by exploiting the browser. All traffic of the operating system is routed through the Tor network.

The Amnesic Incognito Live System (Tails) is a security-focused Linux distribution aimed at preserving privacy and anonymity [32]. All of its outgoing traffic is forced through the Tor network, and non-anonymous connections are blocked. By default, Tails is designed to be booted as a live USB and does not save data to the computer hard disk.

Another security-focused Linux distribution is Whonix [33]. This operating system consists of two separated virtual machines, a desktop installation, and a Tor gateway; all desktop traffic is forced through the Tor gateway. As a result, even if the attacker is able to obtain root access to the desktop operation system, the attacker is unable to de-anonymise the user because the information about the real host IP address is not available inside the desktop machine.

7 Future Work

The Tor network and the diversity of the Tor users and hidden services are increasing steadily. It is possible that the use of Tor will become widespread and popular and, like the Internet itself, become a regular communication network for people. Beyond privacy, Tor supports many desirable, general, and versatile features, including unique network addressing (onion), connectivity behind firewalls or NAT, and end-to-end encryption by default. We believe that the global Tor user base is growing, and this opens up several new research questions.

Strong anonymity can be obtained in many ways. There are other anonymity software programs, such as I2P, Freenet, GNUNet, and many more; English Wikipedia's category of *Anonymous file sharing networks* lists 24 different anonymity networks [34]. We focused on Tor because it is the most popular online anonymity system.

We certainly need more research about anonymity systems. Needless to say, people will use Tor as long as it is easy to use and provides strong privacy features. If there are unfixable security problems on Tor, people will switch to other anonymity systems. For instance, an anonymous marketplace can be deployed without full real-time TCP requirements between the receiver and sender. As a result, a traditional mix of network architecture and an end-to-end encrypted messaging application could be used to sell illegal products anonymously. This kind of anonymity system is highly resilient against traffic and timing correlation attacks [35]. It appears that online anonymity will be a considerable part of the future digital world.

References

1. Goldschlag, D., Reed, M., Syverson, P.: Onion routing. Commun. ACM **42**(2), 39–41 (1999). doi:10.1145/1653662.1653708
2. Dingledine, R., Mathewson, N., Syverson, P.: Deploying low-latency anonymity: design challenges and social factors. IEEE Secur. Priv. **5**(5), 83–87 (2007). doi:10.1109/MSP.2007.108
3. The Tor Project Foundation. https://www.torproject.org/
4. Dingledine, R., Mathewson, N., Syverson, P.: Tor: the second-generation onion router. Technical report, DTIC Document (2004)
5. Guardian, T.: Tor: the king of high-secure, low-latency anonymity (2013). https://www.theguardian.com/world/interactive/2013/oct/04/tor-high-secure-internet-anonymity
6. Biryukov, A., Pustogarov, I., Thill, F., Weinmann, R.P.: Content and popularity analysis of tor hidden services. In: 2014 IEEE 34th International Conference on Distributed Computing Systems Workshops (ICDCSW), pp. 188–193. IEEE (2014). doi:10.1109/ICDCSW.2014.20
7. Semenov, A.: Analysis of services in tor network: finnish segment. In: Proceedings of the 12th European Conference on Information Warfare and Security: ECIW 2013, p. 252. Academic Conferences Limited (2013)
8. Edman, M., Syverson, P.: As-awareness in tor path selection. In: Proceedings of the 16th ACM Conference on Computer and Communications Security, pp. 380–389. ACM (2009). doi:10.1145/1653662.1653708
9. Murdoch, S.J., Zieliński, P.: Sampled traffic analysis by internet-exchange-level adversaries. In: Borisov, N., Golle, P. (eds.) PET 2007. LNCS, vol. 4776, pp. 167–183. Springer, Heidelberg (2007). doi:10.1007/978-3-540-75551-7_11
10. Johnson, A., Wacek, C., Jansen, R., Sherr, M., Syverson, P.: Users get routed: traffic correlation on tor by realistic adversaries. In: Proceedings of the 2013 ACM SIGSAC Conference On Computer and Communications Security, pp. 337–348. ACM (2013). doi:10.1145/2508859.2516651
11. Reed, M.G., Syverson, P.F., Goldschlag, D.M.: Anonymous connections and onion routing. IEEE J. Sel. Areas Commun. **16**(4), 482–494 (1998). doi:10.1109/49.668972
12. Chakravarty, S., Barbera, M.V., Portokalidis, G., Polychronakis, M., Keromytis, A.D.: On the effectiveness of traffic analysis against anonymity networks using flow records. In: Faloutsos, M., Kuzmanovic, A. (eds.) PAM 2014. LNCS, vol. 8362, pp. 247–257. Springer, Cham (2014). doi:10.1007/978-3-319-04918-2_24
13. Danezis, G.: The traffic analysis of continuous-time mixes. In: Martin, D., Serjantov, A. (eds.) PET 2004. LNCS, vol. 3424, pp. 35–50. Springer, Heidelberg (2005). doi:10.1007/11423409_3
14. Elahi, T., Bauer, K., AlSabah, M., Dingledine, R., Goldberg, I.: Changing of the guards: a framework for understanding and improving entry guard selection in tor. In: Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society, pp. 43–54. ACM (2012). doi:10.1145/2381966.2381973

15. Schneier, B.: Attacking Tor: how the NSA targets users' online anonymity. (2013). https://www.theguardian.com/world/2013/oct/04/tor-attacks-nsa-users-online-anonymity
16. The Federal Bureau of Investigation: Narcotic Tracking Conspiracy. Sealed Complaint against Ross Ulbricht (2014). https://www.documentcloud.org/documents/801103-172770276-ulbricht-criminal-complaint.html
17. Ulbricht, R., Forum, B.: Ross Ulbricht's message (2011). https://bitcointalk.org/index.php?topic=47811.msg568744
18. Fonseca, J., Vieira, M., Madeira, H.: Testing and comparing web vulnerability scanning tools for sql injection and xss attacks. In: 13th Pacific Rim International Symposium on Dependable Computing, PRDC 2007, pp. 365–372. IEEE (2007). doi:10.1109/PRDC.2007.55
19. Adida, B.: Sessionlock: securing web sessions against eavesdropping. In: Proceedings of the 17th International Conference on World Wide Web, pp. 517–524. ACM (2008). doi:10.1145/1367497.1367568
20. King, S.T., Tucek, J., Cozzie, A., Grier, C., Jiang, W., Zhou, Y.: Designing and implementing malicious hardware. LEET **8**, 1–8 (2008). doi:10.1145/1346281.2181012
21. Khandelwal, S., Shah, P., Bhavsar, M.K., Gandhi, S.: Frontline techniques to prevent web application vulnerability. Int. J. Adv. Res. Comput. Sci. Electron. Eng. (IJARCSEE) **2**(2), 208 (2013)
22. Mirkovic, J., Dietrich, S., Dittrich, D., Reiher, P.: Internet denial of service: attack and defense mechanisms (radia perlman computer networking and security) (2004)
23. The Federal Bureau of Investigation: Affidavit Case 3: 15-cr-05351-RJB Document 166–2. Playpen website exploit (2016). https://regmedia.co.uk/2016/03/29/alfin.pdf
24. Mozilla Foundation Security Advisory 2013–53: Execution of unmapped memory through onreadystatechange event (2013). https://www.mozilla.org/en-US/security/advisories/mfsa2013-53/
25. Tor-talk mailing list : JavaScript exploit (2016). https://lists.torproject.org/pipermail/tor-talk/2016-November/042639.html
26. Mozilla Foundation Security Advisory 2016–92: Firefox SVG Animation Remote Code Execution (2016). https://www.mozilla.org/en-US/security/advisories/mfsa2016-92/
27. VLC - Ticket system: VLC media player privacy leak due to -no-metadata-network-access not being respected (2016). https://trac.videolan.org/vlc/ticket/17760
28. Naked Security: Use of Tor pointed FBI to Harvard University bomb hoax suspect (2013). https://nakedsecurity.sophos.com/2013/12/20/use-of-tor-pointed-fbi-to-harvard-university-bomb-hoax-suspect/
29. The Federal Bureau of Investigation : Affidavit of special agent Thomas M. Dalton (2013). https://cbsboston.files.wordpress.com/2013/12/kimeldoharvard.pdf
30. Serjantov, A., Danezis, G.: Towards an information theoretic metric for anonymity. In: Dingledine, R., Syverson, P. (eds.) PET 2002. LNCS, vol. 2482, pp. 41–53. Springer, Heidelberg (2003). doi:10.1007/3-540-36467-6_4
31. Published by Der Spiegel: The NSA TEMPORA documentation (2013). http://www.spiegel.de/media/media-34103.pdf
32. Tails: Tails operating system - privacy for anyone anywhere. https://tails.boum.org/
33. Whonix: Stay anonymous with Whonix Operating system. https://www.whonix.org/

34. Wikipedia, The Free Encyclopedia (English): Anonymous file sharing networks (2017). http://goo.gl/aOpGBv

35. Díaz, C., Sassaman, L., Dewitte, E.: Comparison between two practical mix designs. In: Samarati, P., Ryan, P., Gollmann, D., Molva, R. (eds.) ESORICS 2004. LNCS, vol. 3193, pp. 141–159. Springer, Heidelberg (2004). doi:10.1007/978-3-540-30108-0_9

PUBLICATION

V

PESTEL Analysis of Hacktivism Campaign Motivations
J. Nurmi and M. S. Niemelä

Nordic Conference on Secure IT Systems 2018. Ed. by 2018, 323–335

PESTEL Analysis of Hacktivism Campaign Motivations

Juha Nurmi[1,2](✉) and Mikko S. Niemelä[1,3]

[1] Cyber Intelligence House Ltd., Singapore, Singapore
{juha,mikko}@cyberintelligencehouse.com
[2] Tampere University of Technology, Tampere, Finland
[3] Singapore Management University, Singapore, Singapore

Abstract. A political, economic, socio-cultural, technological, environment and legal (PESTEL) analysis is a framework or tool used to analyse and monitor the macro-environmental factors that have an impact on an organisation. The results identify threats and weaknesses which are used in a strengths, weaknesses, opportunities and threats (SWOT) analysis. In this paper the PESTEL framework was utilized to categorize hacktivism motivations for attack campaigns against certain companies, governments or industries. Our study is based on empirical evidence: of thirty-three hacktivism attack campaigns in manifesto level. Then, the targets of these campaigns were analysed and studied accordingly. As a result, we claim that connecting cyberattacks to motivations permits organizations to determine their external cyberattack risks, allowing them to perform more accurate risk-modeling.

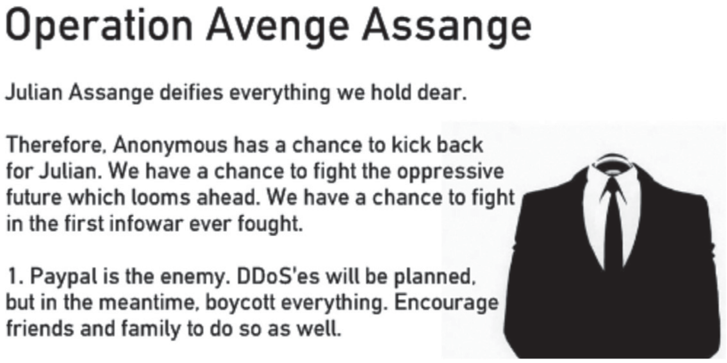
Keywords: PESTEL analysis · Security
Online anonymity · Hacktivism · Cyberattack · Political activism
Strategic management · Risk modeling

1 Introduction

In May 2007, the European Commission published a report "...towards a general policy on the fight against cyber crime..." where cybercrime is defined as "...criminal acts committed using electronic communications networks and information systems or against such networks and systems..." [4]. Furthermore, the report pointed out that cyber attacks are increasing and becoming more sophisticated and internationalised.

The motivations are not always economical. Instead, hacktivism is a way of protesting and it is motivated by ideology, religion, social causes or political opinions [18]. Even many local protests have an aspect of global cyber hacktivism [18]. For example, in 2012, the hacker collective, Anonymous, drew attention to the Anti-Homosexuality Bill in Uganda and attacked several government websites [18]. These protests had significant economical implications [18].

N. Gruschka (Ed.): NordSec 2018, LNCS 11252, pp. 323–335, 2018.
https://doi.org/10.1007/978-3-030-03638-6_20



Operation Avenge Assange

Julian Assange deifies everything we hold dear.

Therefore, Anonymous has a chance to kick back for Julian. We have a chance to fight the oppressive future which looms ahead. We have a chance to fight in the first infowar ever fought.

1. Paypal is the enemy. DDoS'es will be planned, but in the meantime, boycott everything. Encourage friends and family to do so as well.

Fig. 1. The Anonymous hacktivist campaign manifesto (2010), source https://www.undernews.fr/hacking-hacktivisme/avenge-assange-les-anonymous-s%E2%80%99app-retent-a-venger-julian-assange.html. Response to the financial companies which shutdown Wikileaks' accounts and froze personal assets of Julian Assange (the founder of Wikileaks).

Anonymous is a loosely-associated international hacktivist group, which only exists online [9]. The group launched activism operations or campaigns, through a series of distributed denial-of-service (DDoS) attacks on the government organisations and corporate online systems [9]. A study of these campaigns suggests that eighty-two percent were motivated by a defense of free speech or political causes [9].

For instance, in November 2010, Wikileaks (an international non-profit organisation that publishes secret information) released over 251,287 documents (the United States (U.S.) diplomatic cables leak) [13]. These classified documents had been sent to the U.S. State Department by its diplomatic consulates and embassies around the world [13].

In December, financial companies terminated Wikileaks donations [14]. Paypal closed the Wikileaks donation account, the Swiss bank, PostFinance, froze the assets of Julian Assange (the founder of Wikileaks), and MasterCard and Visa stopped payments to the organisation [14].

To protest this, the Anonymous group campaigned to assist WikiLeaks in their quest to release classified government documents. Each Anonymous campaign is accompanied by a manifesto. Figure 1 shows a part of the "Operation Avenge Assange" manifesto.

As a result, the Anonymous campaign produced DDoS attacks which disabled the PayPal website and disrupted the sites of Visa and MasterCard [12]. According to PayPal, the damage cost the company five million USD [12].

After a number of Anonymous protest attacks, hacktivism was weaponised by national states [2] and is no longer driven by well-meaning amateurs. Instead, it is increasingly militarised for geopolitical causes, such as to affect the United Kingdom European Union membership referendum (2016) and the United States presidential election (2016) [15]. These attackers are supported by government institutions to conduct highly specialized attacks with clear a strategy. Hillary Clinton, after losing the presidential election of 2016, even claimed that Vladimir Putin has been conducting a "cyber cold war" against the west [15].

2 Background

In this chapter, a basic overview of a PESTEL analysis framework is provided and a typical hacktivist campaign, manifesto and target list are described.

2.1 PESTEL

In this paper the political, economic, socio-cultural, technological, environment and legal (PESTEL) framework is employed to categorize hacktivism motivations for attack campaigns. PESTEL is a framework for strategic analysis [24], which is also known as PEST analysis [5] and STEPE analysis [16]. Figure 2 represents possible examples of factors under the PESTEL framework, which influences the strategy of analysis.

For governments and companies, PESTEL analysis offers two basic functions: first, it permits the identification of the operational environment and, second, it provides data and information that will enable the company to predict future situations and circumstances [24]. The factors examined in current literature are described.

Political factors indicate the methods through which a government intervenes in the economy, for example, through government policy, political stability, foreign trade policy, tax policy, labour law, environmental law and trade restrictions. Accordingly, these political factors impact business performance. Organizations must respond to the current and anticipated future legislation, and adjust their operations accordingly.

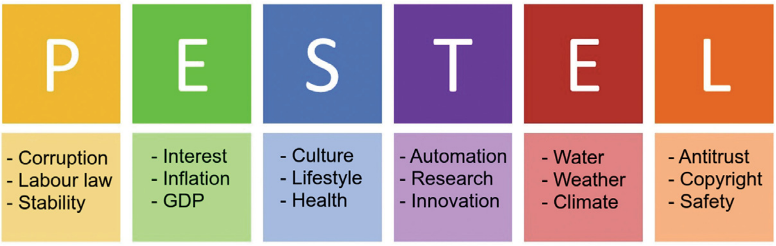


Fig. 2. Political: A government might influence the economy or a certain industry. **Economic:** Performance and patterns of the economy have direct long-term impact. **Social:** These factors are cultural trends, demographics and population analytics. **Technological:** Innovations in technology influence the operations of the industry. **Environmental:** Factors are determined by the surrounding natural environment. **Legal:** This includes regulations that affect the business environment and the market.

Economic factors include economic growth, interest rates, exchange rates, inflation, disposable income of consumers and businesses. On occasion, the factors are categorised into macro-economical and micro-economical factors. Macro-economical factors involve with the management of demand in a given economy and micro-economical factors involve, for instance, the amount of money that customers are able to spend. An economical environment impacts the business performance of an organisation.

Social factors, also known as socio-cultural factors, are the areas that involve shared beliefs and attitudes of the population. These factors include population growth, age distribution, health consciousness, and career attitudes. These factors are of interest as they permit the marketers to understand the motivational forces of their customers.

Technological factors change quickly and influence the markets and the management in three distinct avenues: firstly, in methods of producing services and products; secondly, in methods of service and product distribution; thirdly, in methods of communicating with the target markets.

Environmental factors have become important as a result of the increasing scarcity of raw materials, pollution target requirements, ethical and sustainable company practices and carbon footprint targets determined by governments. These are only a few of the issues the marketers face with respect to this factor. Increasingly, the consumers demand that the products are sourced ethically, and if possible, from a sustainable source.

Legal factors include health and safety, equal opportunities, advertising standards, consumer rights and laws, product labeling and product safety. Companies must be cautious of what is legally permissible in order to trade successfully. If an organization trades globally, this can become a very complex factor, as each country possesses its own rules and regulations.

Factors can be classified under multiple categories at the same time. For instance, carbon footprint targets are considered both political and environmental factors.

PESTEL has been applied to investigate factors of emergence of cloud computing and similar technologies [1]. Strategic analysis has been proposed as a method to follow macro-economic and social trends from online data sources in order to identify and monitor early indicators of security threats [7]. *The United Nations Office on Drugs and Crime The SOCTA Handbook - Guidance on the preparation and use of serious and organized crime threat assessments* recommends PESTEL analysis for criminal activities [21].

For any organization, PESTEL analysis provides macro-environmental predictions and risk analysis, which is utilized in strategic management [5,16]. Currently, however, there is no data available to calculate the risk of hacktivism against an organization. Without this data, it is difficult to perform accurate risk management and formulate a risk-based approach to strategy execution [17].

In the U.S., the average cost of cyber attack is estimated to be two million USD per organization and it is proposed that insurance be utilized as a tool for cyber-risk management related to information security [8]. It is important to understand and estimate the risks associated with this proposal.

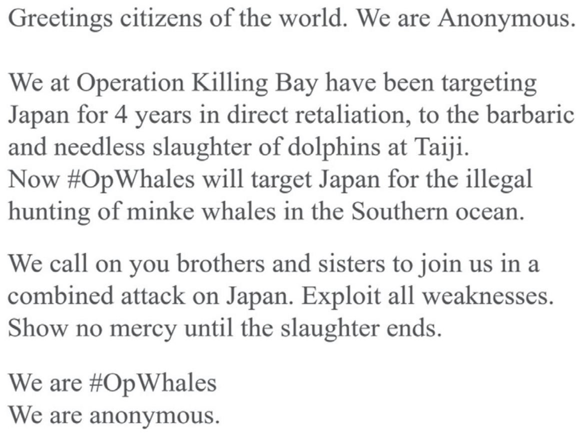
We selected the PESTEL framework because organizations are already performing PESTEL analyses and PESTEL is already utilized as a tool to assess organized crime threats. The discussion of this framework will be extended by mapping cyber attack risks by hacktivist campaigns under the PESTEL framework. This will be completed by compiling research regarding the motivations of these attacks and identify whether they fit under PESTEL. In addition, we will validate the identities of the targets selected according to the manifesto and discuss the impact of the attacks on the targets.

2.2 Hacktivist Campaign Manifestos and Target Lists

The first step of a hacktivist campaign is to publish a manifesto. A manifesto is a declaration of the intentions and motives of the campaign. Hacktivists publish their manifestos online and share them on social media. They explain their motives and intentions to achieve public acceptance for the protest. For instance, Fig. 3 demonstrates how Anonymous seeks public attention and followers to their causes to prevent Japan's whaling program in the Southern Ocean.

The second step of the campaign is to target online services with cyber and DDoS attacks. After these attacks interfere the online services, Anonymous publishes new targets. For example, after attacking online services in Japan they attacked online services in Iceland as well, for selling whale meat. For instance, the target list of the "OpWhales" attack is shown in Fig. 4.

The target list contains HTTP server domain names, HTTP software names, port numbers and IP addresses. The list is publicly available and shared on social media. Anonymous provided DDoS tools for others to participate in the DDoS



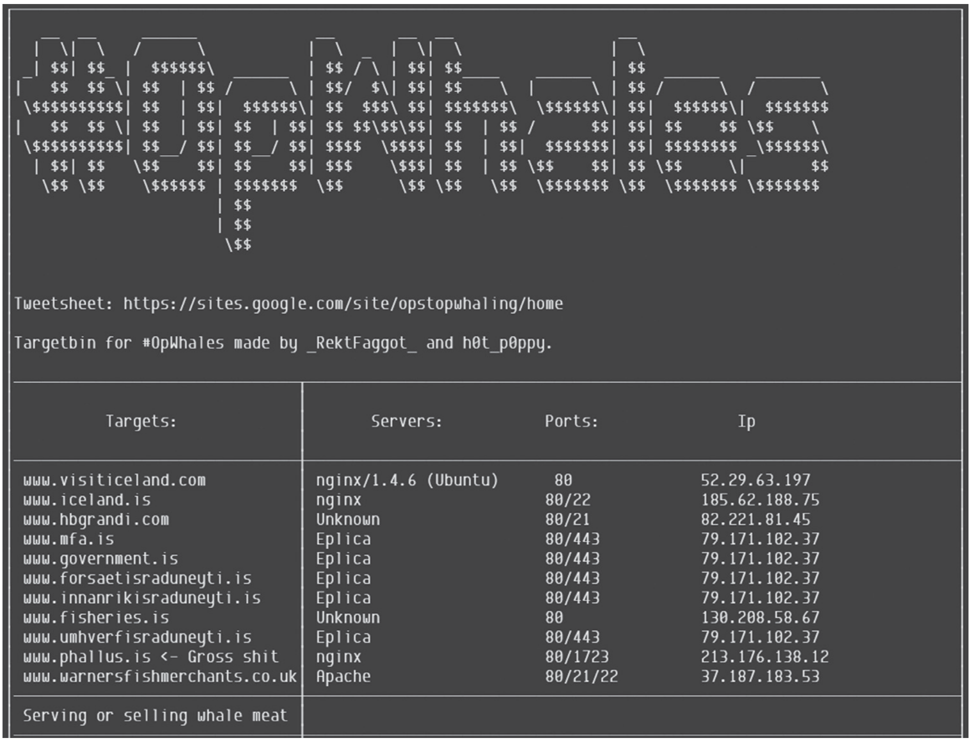
Greetings citizens of the world. We are Anonymous.

We at Operation Killing Bay have been targeting
Japan for 4 years in direct retaliation, to the barbaric
and needless slaughter of dolphins at Taiji.
Now #OpWhales will target Japan for the illegal
hunting of minke whales in the Southern ocean.

We call on you brothers and sisters to join us in a
combined attack on Japan. Exploit all weaknesses.
Show no mercy until the slaughter ends.

We are #OpWhales
We are anonymous.

Fig. 3. A part of the hacktivist campaign manifesto for "OpWhales", an operation in response to the hunting of whales for their meat. Anonymous commands people to participate in the attack against Japan for lifting the ban on whale hunting.



```
   __  __     ______          __    __  __               __
  |  \|  \   /      \        |  \  _  |  \|  \             |  \
 _| $$| $$_ |  $$$$$$\  ______ | $$ / \ | $$| $$____   ______ | $$  ______    _______
|   $$  $$ \| $$  | $$ /      \ | $$/  $\| $$| $$    \ |      \ | $$ /      \  /       \
 \$$$$$$$$$$| $$  | $$|  $$$$$$\| $$  $$$\ $$| $$$$$$$\  \$$$$$$\| $$|  $$$$$$\|  $$$$$$$
|   $$  $$ \| $$  | $$| $$  | $$| $$ $$\$$\$$| $$  | $$ /      $$| $$| $$    $$ \$$    \
 \$$$$$$$$$$| $$__/ $$| $$__/ $$| $$$$  \$$$$| $$  | $$|  $$$$$$$| $$| $$$$$$$$ _\$$$$$$\
  | $$| $$   \$$    $$| $$    $$| $$$    \$$$| $$  | $$ \$$    $$| $$ \$$     \|       $$
   \$$ \$$    \$$$$$$ | $$$$$$$  \$$      \$$ \$$   \$$  \$$$$$$$ \$$  \$$$$$$$ \$$$$$$$
                      | $$
                      | $$
                       \$$


Tweetsheet: https://sites.google.com/site/opstopwhaling/home

Targetbin for #OpWhales made by _RektFaggot_ and h0t_p0ppy.

          Targets:                    Servers:           Ports:              Ip

 www.visiticeland.com         nginx/1.4.6 (Ubuntu)      80              52.29.63.197
 www.iceland.is               nginx                    80/22            185.62.188.75
 www.hbgrandi.com             Unknown                  80/21            82.221.81.45
 www.mfa.is                   Eplica                   80/443           79.171.102.37
 www.government.is            Eplica                   80/443           79.171.102.37
 www.forsaetisraduneyti.is    Eplica                   80/443           79.171.102.37
 www.innanrikisraduneyti.is   Eplica                   80/443           79.171.102.37
 www.fisheries.is             Unknown                  80               130.208.58.67
 www.umhverfisraduneyti.is    Eplica                   80/443           79.171.102.37
 www.phallus.is <- Gross shit nginx                    80/1723          213.176.138.12
 www.warnersfishmerchants.co.uk Apache                 80/21/22         37.187.183.53

 Serving or selling whale meat
```

Fig. 4. A target list of an Anonymous hacktivist campaign. Anonymous published a list of online targets as a part of "OpWhales". There were multiple similar target lists during the operation. After targeting online services in Japan, Anonymous targeted companies in Iceland for selling whale meat.

attack. As a result, an overwhelming amount of HTTP traffic overwhelmed the sites and prevented these online services from operating normally.

Anonymous shared existing network stress-testing tools which could be utilized to perform a DDoS attack on a target site by overwhelming the server with HTTP traffic. Web-based tools can be utilized without installation as they involve a website that executes the attack through a JavaScript code that launches a flood of traffic from the user's machine.

People voluntary visited the attack tool site and selected targets from the target list, despite the fact that in many countries it is illegal to participate in DDoS attacks. These attacks caused a significant amount of web traffic to the targeted online services and resulted in an interruption for them.

Declaring participation illegal does not prevent these attacks [22]. There are multiple techniques available to hide the origin of the attack. Anonymity networks, such as the Tor network, hide the IP address of the machine of the user who participates to the DDoS attack [6,20]. Furthermore, there are anonymous discussion channels, including the Internet Relay Chat (IRC) program channels inside the Tor network. These anonymous communication channels are utilized to coordinate DDoS attacks against the targets. De-anonymisation of these users or communication channels is technically very difficult [11]. See Fig. 7 in the Appendix A as an example of an anonymous onion website that shares a tutorial to connect IRC channels that operate inside the Tor network.

3 PESTEL Analysis for Hacktivism Campaign Motivations

In this chapter, the motivations of hacktivist campaigns are studied and are classified according to the PESTEL framework.

3.1 Motivations of Hacktivist Campaigns

Hacktivist campaign manifestos and target list data were assembled and studied. Here, we present thirty-three campaigns between 2011–2018. The motivations of the campaigns were examined, and we validated the selected targets according to the motivations. We have gathered these events into one timeline in Table 1 of Appendix B.

Motivations are commonly clearly stated in the manifestos and other publications of the hacktivist groups [19]. An example of this is the "OpBahrain" manifesto in Fig. 8 of Appendix C.

We validated that the targets that were selected according to these motivations. This indicates that campaigns follow their stated motives: the target lists contains online services of industries, organization and governments which are, in the hacktivist world view, involved with operations that the they are against.

As a result, we can claim that campaigns have a clear motivation as dictated in the manifesto and that the campaigns follow their manifestos. The targets are selected according to manifesto motivation. Hactivism is motivation driven, indicating that it is reasonable to examine how these motivations are classified with the PESTEL model.

3.2 Fitting the Motivations to the PESTEL Framework

In Fig. 5, the motivations are placed under political, economical, social, technological, environmental and legal categories. We selected the most suitable category according to the campaign motivations, although, there is no absolute methodology to perform categorisation and a campaign might fit under more than one category.

For example, "OpSaveGaza" was against the Israeli bombing of Gaza (political), "OpIcarus" was against the dominance of the financial sector (economic), "OpNoDAPL" was regarding solidarity with Native American protests against the Dakota Access Pipeline (social), "OpWhales" protested against whale hunting (environmental), and "OpAbdiMohamed" protested against police violence in the U.S. (legal).

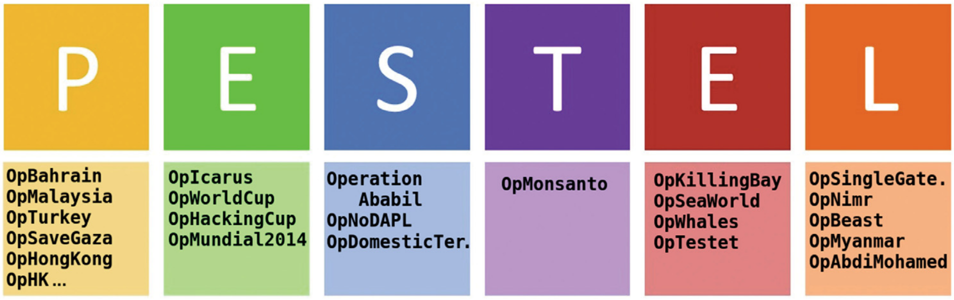


Fig. 5. Hacktivist follow their campaign motives and select targets that are, from their point of view, involved in unethical activities. These motives can be categorized under PESTEL: political, economical, social, technological, environmental and legal.

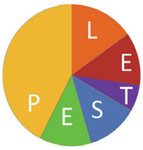


Fig. 6. We fit the motivations under political, economical, social, technological, environmental and legal categories. Political motivation is the most popular category.

The motivations of hacktivists targets cover all PESTEL categories, however, it appears as if technology is not often targeted by hacktivists. Instead, they are more motivated by political, economical, social, technological, environmental and legal causes (Pie diagram 6).

This does not indicate that there are no technological motivations for cyber attacks. Many operations have a distinct technological aspect. Anonymous published “OpSingleGateway” against Thailand after the passing of technical surveillance methods, which permitted the government to censor websites and intercept private communications without a court order or warrant. Notably there was a “Operation Monsanto” against carcinogenic chemicals in food, which are produced by Monsanto.

4 Results

In this paper, we have presented how attack campaigns fit under PESTEL at the level of manifestos. Moreover, we validated that the targets are selected according to motivations. Finally, we present analysis of these realistic cyber attack threads to different industries.

I. Political. We demonstrated that hacktivist groups target governments and companies if they provoke political activism against them. This is the most frequent cause of the attack.

II. Economical. We indicated that economical decisions and circumstances can cause hacktivist groups to attack companies and governments.

III. Social. We demonstrated that cases of social causes can result in a swift hacktivist response against companies and governments.

IV. Technological. We detected that technology itself is seldom the main reason for attacks.

V. Environmental. We found that environmental causes are common reasons to launch hacktivist campaigns.

VI. Legal. Our results indicate that a legal atmosphere has activated several hacktivism campaigns.

Our results demonstrate that governments and companies are able to consider the risk of cyber attacks when the PESTEL framework is employed in analysis. For instance, if a company provides services to a whaling industry, they should prepare to be targeted by hacktivist organisations. There is a significant price attached to a cyber attack that disturbs their online services [10]. Or, if their data is stolen during an attack [23].

5 Conclusion

Providing a usable framework to analyse the risk of cyber attack on the Internet is an ongoing challenge for any organization. Fortunately, because of strategical analysis, we are able to study the motivations of hacktivism. This permits a company to forecast whether it is doing something that could motivate attacks against its online systems.

Because organizations are already applying PESTEL analysis their to macro-environmental predictions and risk analysis, they could look their organization from the hacktivism point of view. Strategic management should ask two questions:

(1) What could cause our organization being targeted by hacktivists?
(2) What is the price of this risk?

After this, the strategic management is able to react accordingly.

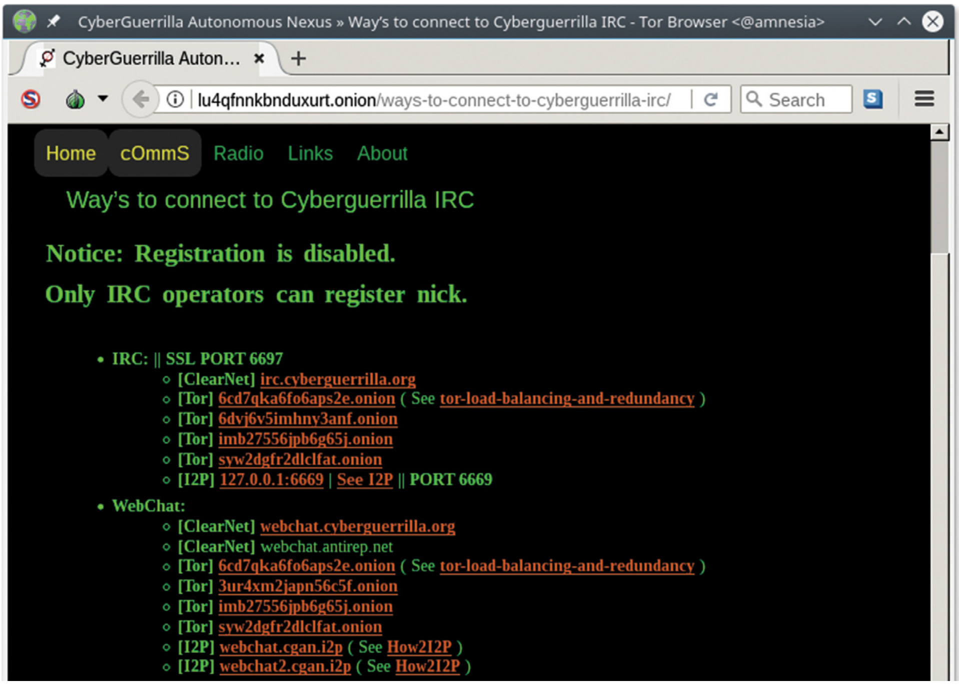
6 Discussion

A PESTEL framework is not the only method to analyse risks and opportunities. There are other frameworks available. These tools could be extended to map cyber security risks. Also, hacktivism is not the only cause of cyber attacks. Instead, hackers are increasingly supported by government institutions and conduct highly-specialised attacks. These attacks have strategic and geopolitical causes. Previously the actors have been analyzed by Intel which developed threat agent library (TAL) which describes the human agents that pose threats to IT systems [3].

The number of effective cyber attacks is increasing steadily. It is possible that new motivations for these attacks arise and these motivations could be organised under PESTEL. Clear limitation in our paper is that there is no absolute methodology for the categorisation of the motivations. Human motivations are various and complex. We need more research to improve understanding of the motivations to predict hacktivism.

We believe that the price of cyber attacks should be calculated and this leads to new research questions. More research is needed regarding the effects of cyber crime and hacktivism. In addition, novel methods to estimate what is the risk of being targeted by hacktivist groups and how to mitigate these risks are necessary.

A Discussion Channels Within the Tor Anonymity Network Are Used to Coordinate DDoS Attacks


CyberGuerrilla Autonomous Nexus » Way's to connect to Cyberguerrilla IRC - Tor Browser <@amnesia>
CyberGuerrilla Auton...
lu4qfnnkbnduxurt.onion/ways-to-connect-to-cyberguerrilla-irc/
Search
Home
cOmmS
Radio
Links
About

Way's to connect to Cyberguerrilla IRC

Notice: Registration is disabled.

Only IRC operators can register nick.

- IRC: || SSL PORT 6697
  - [ClearNet] irc.cyberguerrilla.org
  - [Tor] 6cd7qka6fo6aps2e.onion (See tor-load-balancing-and-redundancy)
  - [Tor] 6dvj6v5imhny3anf.onion
  - [Tor] imb27556jpb6g65j.onion
  - [Tor] syw2dgfr2dlclfat.onion
  - [I2P] 127.0.0.1:6669 | See I2P || PORT 6669
- WebChat:
  - [ClearNet] webchat.cyberguerrilla.org
  - [ClearNet] webchat.antirep.net
  - [Tor] 6cd7qka6fo6aps2e.onion (See tor-load-balancing-and-redundancy)
  - [Tor] 3ur4xm2japn56c5f.onion
  - [Tor] imb27556jpb6g65j.onion
  - [Tor] syw2dgfr2dlclfat.onion
  - [I2P] webchat.cgan.i2p (See How2I2P)
  - [I2P] webchat2.cgan.i2p (See How2I2P)

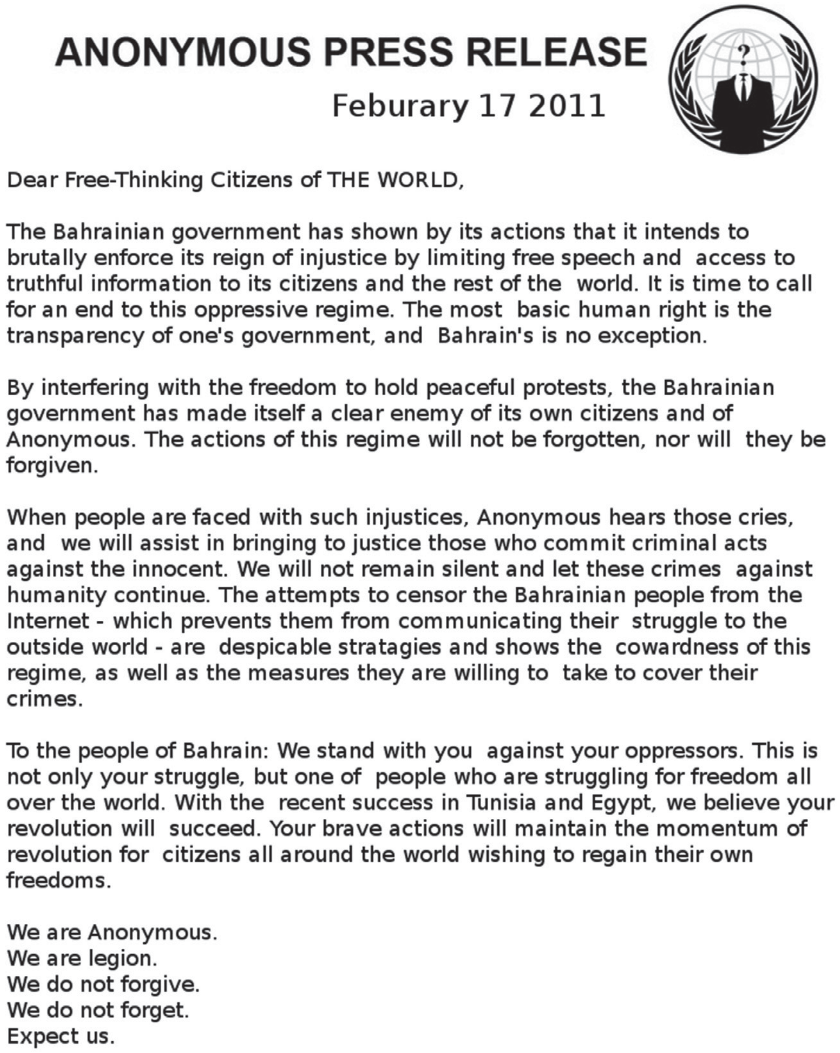
Fig. 7. An example of anonymous onion website that shares a tutorial to connect IRC channels. These services operate inside the Tor anonymity network.

B Hacktivist Campaigns, Motivations and Targets

Table 1. An approximated timeline of hacktivist campaigns. Thirty-three cases were examined for motivation and targets. The main target sectors and countries are listed here. The number of targets refers to unique sites and online services which were attacked during the campaign. Anon. represents Anonymous, CyFi represents Cyber Fighters of Izz ad-Din al Qassam, and NWH represents New World Hackers. Please note that timeline is not clear because many campaigns failed to start or were re-launched several times. The main target sectors and countries are listed here. Finally, there are categories of motivation under the PESTEL framework. Note that several campaigns could intuitively fit under more than one category. We selected the main category.

Began	Group	Campaign name	Motivations and reasons in manifesto	Main target countries and industries	Targets	Cat.
02/2011	Anon.	OpBahrain	Bahrain interfered peaceful protest	Saudi Arabia; government, aviation, education, media, financial, sport, medical and energy	22	Pol.
06/2011	Anon.	OpMalaysia	Against Internet censorship in Malaysia	Malaysia; government and law enforcement	2	Pol.
07/2011	Anon.	OpMonsanto	Against a seeds supply monopoly and harmful farming chemicals	Agricultural biotech giant Monsanto	2	Tec.
08/2012	Anon.	OpMyanmar	Myanmar refused to recognize the Rohingya minority as citizens	Myanmar; government, education, media, airlines, energy and telecommunications	71	Leg.
08/2012	CyFi.	Operation Ababil	An anti-Islamic short film were uploaded to YouTube in July 2012	USA; banking and financial	28	Soc.
10/2012	Red October	Red October	Data collection	Eastern Europe, central Asia, specifically government embassies, military installations, energy providers, research firms	N/A	N/A
06/2013	Anon.	OpTurkey	Response to the police crackdown of protests and Internet censorship	Turkey; government, media, political party, law enforcement, financial and telecommunications	3235	Pol.
01/2014	Anon.	OpWorldCup, OpHackingCup, OpMundial2014	Against corruption and inequality in Brazil	Brazil, USA; government, sport, airlines, financial, education, telecommunications, energy and sport	39	Eco.
01/2014	Anon.	OpKillingBay, OpSeaWorld, OpWhales	Protest against whale hunting	Faeroe Islands, Japan, China, Singapore, USA, Norway, Iceland, Turkey, Canada; restaurant, media, marine services, logistics, lodging, government, fishing, airlines and entertainment	401	Env.
07/2014	Anon.	OpSaveGaza	Protest against Israel bombing Gaza	Israel; government	168	Pol.
09/2014	Anon.	OpTestet	Protest to against Testet dam project to save the Sivens forest, France	France; government and construction industry	343	Env.
10/2014	Anon.	OpHK, OpHongKong	Riot police used tear gas and pepper spray on peaceful protesters	Hong Kong, China; media, aviation, government and party	181	Pol.
04/2015	Anon.	OpBeast	Demand of world wide ban on sexual intercourse with animals	USA, Hungary, Finland; zoophilia and government	438	Leg.
09/2015	Anon.	OpNimr	Calling on Saudi Arabia to halt the execution of Al-Nimr who participated Saudi Arabian protests as a teenager	Saudi Arabia, UAE, MENA; Government, financial, aviation, energy, media and education	132	Leg.
10/2015	Anon.	OpSingleGateway	Against Internet censorship in Thailand	Thailand; government, military and media	94	Leg.
11/2015	Turla	N/A	Russian state-sponsored group hacked over 100 websites	Governments and businesses	N/A	Pol.
12/2015	Phantom Squad	N/A	A DDoS attach on Microsoft's Xbox Live service	Microsoft's Xbox Live	1	Tec.
12/2015	Packrat	N/A	Targeting South American countries with malware	Governments and businesses	N/A	N/A
03/2016	NWH	OpAbdiMohamed	Protesting police violence: a 17-year-old Abdi Mohamed was shot by police while holding a broomstick	USA; Salt Lake City police, financial and airport	6	Leg.
05/2016	Anon.	OpIcarus	**O**peration **I Car**e, against the financial sector dominance	Most of the countries in the world; financial and stock exchange	390	Eco.
08/2016	Anon.	OpNoDAPL	Solidarity with Native American protest against Dakota Access Pipeline	Mainly USA; financial, defense, military and government	46	Soc.
01/2017	Fancy Bear	APT28	Infiltrating TV stations in the UK	TV stations in the UK	N/A	Pol.
01/2017	Gaza Cybergang	N/A	Cyber espionage campaign against governments in the Middle East Area	Governments in the Middle East area	N/A	Pol.
02/2017	North Korea	N/A	Malware campaign against South Korea	The South Korean government	N/A	Pol.
02/2017	Gamaredon	N/A	Cyber espionage campaigns against the Ukrainian law enforcement	The Ukrainian government	N/A	Pol.
02/2017	Anon.	Operation Darknet	Bringing down dark net websites that had child pornography	Freedom Hosting II servers	1	Soc.
03/2017	APT10	OperationCloudHopper	Access to several MSPs, a campaign that ran since 2016	Major MSPs	N/A	Pol.
12/2017	Anon.	OpDomesticTerrorism	Taking down 12 neo-Nazi sites. The official website of Charlottesville	Charlottesville city, Virginia, government	13	Soc.
08/2017	Anon.	N/A	Breach of 1.2 million patients in the UK National Health Service	the UK National Health Service	1	N/A
10/2017	Anon.	Operation Catalonia	The Catalan independence crisis	Spanish government institutions	N/A	Pol.
02/2018	Group 123	N/A	A total of six malicious campaigns focused on South Korean targets	South Korean industries	N/A	Pol.
02/2018	Dark Caracal	N/A	Targeting victims around the world to collect useful information	Governments, militaries, utility companies, financial institutions, manufacturing companies and defense contractors	N/A	N/A
02/2018	N/A	TopHat	Attacking Middle Eastern Internet users with malware	Internet users in the Middle East	N/A	N/A

C OpBahrain Manifesto by the Anonymous Hacktivist Group



ANONYMOUS PRESS RELEASE

Feburary 17 2011

Dear Free-Thinking Citizens of THE WORLD,

The Bahrainian government has shown by its actions that it intends to brutally enforce its reign of injustice by limiting free speech and access to truthful information to its citizens and the rest of the world. It is time to call for an end to this oppressive regime. The most basic human right is the transparency of one's government, and Bahrain's is no exception.

By interfering with the freedom to hold peaceful protests, the Bahrainian government has made itself a clear enemy of its own citizens and of Anonymous. The actions of this regime will not be forgotten, nor will they be forgiven.

When people are faced with such injustices, Anonymous hears those cries, and we will assist in bringing to justice those who commit criminal acts against the innocent. We will not remain silent and let these crimes against humanity continue. The attempts to censor the Bahrainian people from the Internet - which prevents them from communicating their struggle to the outside world - are despicable stratagies and shows the cowardness of this regime, as well as the measures they are willing to take to cover their crimes.

To the people of Bahrain: We stand with you against your oppressors. This is not only your struggle, but one of people who are struggling for freedom all over the world. With the recent success in Tunisia and Egypt, we believe your revolution will succeed. Your brave actions will maintain the momentum of revolution for citizens all around the world wishing to regain their own freedoms.

We are Anonymous.
We are legion.
We do not forgive.
We do not forget.
Expect us.

Fig. 8. A manifesto of a hacktivist campaign. Anonymous published this manifesto before it launched "OpBahrain" attacks against the Bahrainian government. The manifesto describes clear motivation for the attacks.

References

1. Bakri, N.A.M., et al.: Pestle analysis on cloud computing
2. Caldwell, T.: Hacktivism goes hardcore. Netw. Secur. **5**, 12–17 (2015)
3. Casey, T.: Threat agent library helps identify information security risks. Intel White Paper (2007)

4. Commission, E.: Towards a general policy on the fight against cyber crime. Technical report, COM (2007) 267 final (2007). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0267:FIN:EN:PDF
5. Dale, C.: The uk tour-operating industry: a competitive analysis. J. Vacation Mark. **6**(4), 357–367 (2000)
6. Dingledine, R., Mathewson, N., Syverson, P.: Deploying low-latency anonymity: design challenges and social factors. IEEE Secur. Privacy **5**(5), 83–87 (2007). https://doi.org/10.1109/MSP.2007.108
7. Gómez-Romero, J., Ruiz, M.D., Martín-Bautista, M.J.: Open data analysis for environmental scanning in security-oriented strategic analysis. In: 2016 19th International Conference on Information Fusion (FUSION), pp. 91–97. IEEE (2016)
8. Gordon, L.A., Loeb, M.P., Sohail, T.: A framework for using insurance for cyber-risk management. Commun. ACM **46**(3), 81–85 (2003)
9. Klein, A.G.: Vigilante media: unveiling anonymous and the hacktivist persona in the global press. Commun. Monogr. **82**(3), 379–401 (2015)
10. Lagazio, M., Sherif, N., Cushman, M.: A multi-level approach to understanding the impact of cyber crime on the financial sector. Comput. Secur. **45**, 58–74 (2014)
11. Nurmi, J., Niemelä, M.S.: Tor de-anonymisation techniques. In: Yan, Z., Molva, R., Mazurczyk, W., Kantola, R. (eds.) NSS 2017. LNCS, vol. 10394, pp. 657–671. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64701-2_52
12. Published by BBC: Anonymous hackers ‘cost PayPal 3.5m’ (2012). http://www.bbc.com/news/uk-20449474
13. Published by Der Spiegel: State Department Secrets Revealed, How America Views the World (2010). http://www.spiegel.de/international/world/state-department-secrets-revealed-how-america-views-the-world-a-732819.html
14. Published by Der Spiegel: Visa, MasterCard Move To Choke WikiLeaks (2010). https://www.forbes.com/sites/andygreenberg/2010/12/07/visa-mastercard-move-to-choke-wikileaks/
15. Published by The Guardian: Cyber cold war is just getting started, claims Hillary Clinton (2017). https://www.theguardian.com/us-news/2017/oct/16/cyber-cold-war-is-just-getting-started-claims-hillary-clinton
16. Richardson Jr., J.V.: The library and information economy in turkmenistan. IFLA J. **32**(2), 131–139 (2006)
17. Sheehan, N.T.: A risk-based approach to strategy execution. J. Bus. Strategy **31**(5), 25–37 (2010)
18. Solomon, R.: Electronic protests: hacktivism as a form of protest in uganda. Comput. Law Secur. Rev. **33**(5), 718–728 (2017)
19. Taylor, R.W., Fritsch, E.J., Liederbach, J.: Digital Crime and Digital Terrorism. Prentice Hall Press, New Jersey (2014)
20. The Tor Project Foundation. https://www.torproject.org/
21. UN: United Nations Office on Drugs and Crime the SOCTA Handbook Guidance on the preparation and use of serious and organized crime threat. United Nations Office on Drugs and Crime (2010)
22. Wall, D.: Crime and the Internet. Routledge, London (2003)
23. Yar, M.: Cybercrime and Society. Sage, London (2013)
24. Yüksel, İ.: Developing a multi-criteria decision making model for pestel analysis. Int. J. Bus. Manag. **7**(24), 52 (2012)