Chapter 1

Introduction

The rapid development of digital technology has transformed the world by driving our society into a digital era where the convergence of Information and Communication Technologies (ICTs) has enhanced the exchange of information between individuals and organisations. Data digitalisation facilitates this exchange and improves several processes in diverse social, economic, and political areas. Multiple platforms and communications channels connected to the Internet provide access to countless services where people interact and perform transactions that are part of their regular activities. As a result, digital technology is ubiquitous, and we depend on it every day.

Although digital innovation offers a wealth of opportunities and capabilities to enrich people's lives and promote economic growth globally, new technologies also introduce several risks that affect our society. The ease of access to information and enhanced connectivity to a global network has resulted in new crime levels. Malicious actors use or misuse technology to commit illegal activities that are enhanced by the use of electronic devices or that have arisen by using them. These crimes involving ICTs have been traditionally known as computer crime, online crime, or cybercrime.

Online crime is a global issue growing exponentially in sophistication and professionalism. It can be very damaging for victims, not necessarily just for financial reasons, as victims may also experience a loss of reputation or experience psychological damage. The Internet facilitates crime by providing offenders with

an infrastructure with millions of potential victims regardless of their location. As a consequence, an increasing number of adversaries are plotting highly elaborated schemes to commit illegal activities. Information security firms predict global online crime will cost $10.5 trillion per year by 2025, up from $3 trillion in 2015 [3].

Society's technological dependence and weaknesses in information systems present opportunities for miscreants to engage in deviant behaviour motivated by various factors, including economic rewards, political beliefs, reputation status building, and desire to inflict harm or pain. From a financial point of view, online crime is a massive business that generates enormous profits and fuels an economy where data is an illicit commodity. Such is the popularity of this type of crime that several illegal products and services are available to anyone who wants to venture into the cybercrime scene to obtain some benefit.

Online illegal activities come in different forms and happen in different layers of the Internet. This dissertation explores malicious activity on the Surface Web and the Dark Web to determine how cybercrime unfolds in different Web environments. We used honeypots, specialised crawlers and analysis tools to examine different types of online crimes. In addition, we establish comparisons to understand the modus operandi of cybercriminals in both environments, as this has never been done in our field.

This chapter will describe cybercrime, its classification and the evolution that it has undergone thanks to the advancement of technology, resulting in the use of different platforms such as the Dark Web to commit computer crimes. It will then present the rationale behind this thesis.

1.1 Cybercrime

Cybercrime is the name used most frequently in the criminological literature to describe computer crime [4]. First coined in 1982, the term *cyberspace* described a conceptually constructed virtual environment where networked computer activity occurs [5]. Cybercrime, therefore, refers to crimes committed within that "space." The term is broadly used to describe an extensive range of illicit activities using

digital electronic technologies.

Since cybercrime is an umbrella term to define distinct types of criminal activity, there have been various attempts to classify these activities. One common approach is to focus on technology's role, i.e. whether technology is used to support a crime that otherwise can be committed without it or whether technology is essential to conduct the crime. Several policing bodies have adopted this approach, such as the National Crime Agency in the UK. According to this classification, cybercrime can be split into two broad categories:

- Cyber-dependent crimes are offences that can only be committed via a computer, computer network, or other ICT forms [6]. For instance, the use of malicious software (malware) or gaining unauthorised access to someone's system or network.

- Cyber-enabled crimes are traditional crimes that can be increased in their scale or reach by using ICT, but unlike cyber-dependent crimes, they can be committed without it [6]. These offences include fraud, drug trafficking, and child pornography.

1.2 The Evolving Sophistication of Cybercrime

Cybercrime is evolving at an alarming rate, with new trends developing regularly. Cybercriminals have become more sophisticated, leveraging new technologies at breakneck speed, customising their attacks with novel approaches, and collaborating on a large scale. Sophisticated criminal networks operate all over the world, coordinating complex attacks in minutes.

At the same time, more miscreants perceive the risk of being caught or punished as too low when committing malicious activities while the pay-offs are high. Indeed, a report from the Center for Strategic and International Studies (CSIS), in partnership with McAfee [7], shows that the rapid adoption of new technologies by cybercriminals, including the rapid expansion of black markets and digital currencies, have accelerated cybercrime growth. Law enforcement agencies must therefore

understand the possibilities created by new technologies and use them as tools to combat cybercrime.

1.3 The Rise of the Dark Web

People and companies benefit from new technologies because they make them more efficient and effective; cybercriminals are not an exception. Among the reasons the CSIS-McAfee report [7] attributes to the growth of cybercrime is that malicious actors quickly adopt new technologies. However, they become more sophisticated by improving methods and techniques to achieve their goals more efficiently and conducting their operations without getting caught or blocked.

For instance, the increasing use of the Dark Web and the anonymity that this platform provides have attracted cybercriminals who can commit various computer crimes and maintain their activities hidden from law enforcement agencies. The Dark Web refers to websites hosted on networks built on top of the Internet that are not indexed by conventional search engines and only accessible by specialised software such as The Onion Router (Tor) [8].

The main feature of these networks is that they provide user privacy by obfuscating the traffic between a client and a website or online service; therefore, users can access the hosted content anonymously [9]. The Tor network offers encrypted communications through which content providers can anonymously distribute content.

Tor's features can hamper the attempts of law enforcement agencies to track illegal activities and ultimately stop criminals suggesting that criminals are using hidden platforms on the Dark Web to develop their nefarious activities [10, 11]. The same CSIS-McAfee report [7] indicates that tools like Tor have enabled the expansion of cybercrime.

1.4 The Rationale Behind this Work

As mentioned previously, cybercrime is growing rapidly, and a great deal of research has focused on malicious activity taking place on the Surface Web. The Surface Web refers to the indexed and publicly accessible part of the Internet that is searchable

using a Web search engine such as Google, Bing, and Yahoo. Users on this Web environment are subject to being tracked because their browsing histories and IP addresses are not hidden and can be identified [10].

Although the research community and law enforcement agencies make a great effort to stop illegal activities online, criminals constantly adapt and find new ways to accomplish their objectives. As miscreants use new technologies such as the adoption of the Dark Web, the challenge to understand how they operate becomes more complex. Some research has investigated various types of illegal activities on the Dark Web; however, very few studies have compared how the same crime unfolds on the Surface Web and the Dark Web. Understanding how cybercriminals operate in different layers of the Web is important for developing specific and accurate interventions aimed to prevent cybercrime.

This research aims to fill in this gap by comparing the malicious activity generated by cybercriminals on the Surface Web and the Dark Web. To this end, we define our general research question as: Does the Web environment influence cybercriminal activity? The goal is to understand the modus operandi of cybercriminals and how it varies depending on the different layers of the Internet. To address our general question, the following sub-research questions were established:

- **RQ1:** Does the Web environment and the outlet where stolen accounts are leaked influence the type and amount of activity that those accounts receive?
- **RQ2:** Does the type of network that a user utilises to connect to the Internet influence the type of advertising and the level of malvertising that the user is exposed to?
- **RQ3:** Does the Web environment where underground forums are hosted influence the type of transactions, products, and prices traded in those forums?

Similarly, the use of data collection tools to monitor some types of cybercrime and statistics for measuring malicious activity are important factors of the novelty of this work. The results of this work may contribute new knowledge to the fight against cybercrime from the Information Security perspective by designing technical

countermeasures and from the criminological angle to improve the application of their theories to cybercrime.

1.5 The Document Outline

The current document is divided as follows. Section 2 is a detailed literature review of the following sections; Section 3 is about compromised online accounts. This chapter compares the modus operandi of criminals using webmail accounts whose credentials were leaked on the Surface Web and the Dark Web. Section 4 is a chapter related to malvertising. This section offers a comparison of how malvertising is delivered between both Web environments. Section 5 is aimed to understand criminal activity in underground forums for the two Web environments. Section 6 is dedicated to the ethical issues of my research. Section 7 completes this document by presenting conclusions and final remarks.

Chapter 2

Background and Related Work

This chapter is intended to review the literature about the state of the art on the articles shaping the chapters of this document. The first section focuses on describing cybercrime from the point of view of environmental criminology and the information security approach. The following section detail focuses on the Dark Web as a new environment that facilitates cybercrime.

2.1 Understanding Cybercrime

The term ”cybercrime” encompasses a wide variety of criminal activities committed through or related to ICT systems. Due to its complexity, cybercrime presents new challenges for policing and security. There is now a substantial body of work addressing these challenges from various fields, including criminology and computer science, more specifically information security (InfoSec). While criminology research applies traditional crime theories to understand cybercrime, the Information Security literature focuses on how the technology is used to commit online malicious activities. The main issue about understanding cybercrime is that it can be conducted in different ways in a vast digital environment known as cyberspace. As a result, various academic contributions concentrate on a specific part of the problem while ignoring others.

2.1.1 Theoretical Perspective

Criminologists have studied cybercrime to understand the nature and patterns of rule-breaking activity that have emerged due to new digital changes, usually from

the offender's perspective. On the other hand, InfoSec scholars have focused on the technological aspect of the problem to shed light on how cybercriminals use ICTs to conduct their activities. It makes sense then that the issue should be treated from an interdisciplinary viewpoint. Consequently, some authors have suggested that it is vital to study cybercrime using criminology theories [12], while others have attempted to bridge criminology and InfoSec fields together to develop better mitigations aimed to fight computer crime [13].

Different positions exist in the criminology field that attempts to understand the criminal event and explain why and how people engage in criminal behaviour. Since cybercrime represents the emergence of a "new" form of crime, it is essential to investigate the modus operandi of cybercriminals from a technological perspective so that this knowledge can be studied using criminological theories. As a result, the cybercriminal event can be described using a criminological approach, and more effective strategies can be developed to combat cybercrime.

2.1.2 Environmental Criminology Theories

Criminology has proposed different approaches to understand traditional crime and deter individuals from engaging in deviant behaviour. Specifically, environmental criminology focuses on the specific places and times crime occurs. It is theorised that the environment and certain circumstances facilitate illegal activities and influence the modus operandi of criminals [14]. Thus, by manipulating the environment, crime can be prevented and controlled.

Environmental criminology theories have been proposed to explain real-world crimes situated in a physical environment or a place. In recent years there has been considerable interest in these theories as a means of understanding and tackling cybercrime more effectively. Since the concept of 'place' is essential to environmental criminology, these theories have been translated to a "cyberplace" or cyberspace, an electronic medium consisting of an interconnected network of information system infrastructures where users interact using different platforms, applications and services [13]. Among the most applied theories to study cybercrime are: routine activity theory, rational choice theory and situational crime prevention.

2.1.2.1 Routine Activity Theory

Cybercrime has been primarily studied by applying Routine Activity Theory (RAT). According to RAT, crime is likely to occur when a motivated offender and a suitable target converge in time and space in the absence of a capable guardian [15]. In essence, opportunities for crime arise when these three elements are present in the environment. The Internet amplifies criminal opportunities in cyberspace, where billions of users interact through different devices, platforms and services with little or no security implemented. The insignificant amount of time required to reach those targets in a virtual space with no geographic borders motivates offenders to engage in criminal activities [16].

Some studies have shown that RAT can be applied to understand cybercrime. However, there are differences between the physical space and cyberspace that limit the use of the theory [17, 18, 16]. These differences arise because cybercrime is a new form of crime; thus, the applicability of RAT differs between different types of cybercrime as some elements of RAT are more applicable than others. Overall, evidence suggests that users spending more time online have a higher probability of victimisation [18]. For instance, business hours give criminals more opportunities to engage in computer crimes as more potential victims are accessible [19]. Similarly, the use of banking, email and instant messaging platforms increases by 50% the risk of victimisation on Internet users [20].

Other authors have explored the relationship between some demographic characteristics and the risk of victimisation using RAT [16]. The results show that opportunities for cybercrime increase in more developed countries because victims and offenders have more access to technology. At the same time, opportunities also arise in less developed countries with there is a high level of access to technology, since social problems such as unemployment can be motivating factors for offenders to commit online malicious activities. It seems that wealthier nations with more Internet users per capita have higher rates of cybercrime according to RAT analysis.

Some studies related to malware and RAT have found that some factors that can be considered as capable guardianship reduce the likelihood of infection. For

instance, users with technical protections implemented, such as anti-virus software (considered a type of guardianship), reduced their risk of victimisation [21]. Similarly, evidence shows that users with increased computer skills are more prepared to protect themselves from cybercrime. However, the use of technical protections may not be a consequence of the level of user skills but rather of the ubiquity of this type of software in most computers.

So far, the literature has shown that RAT can be used to explain cybercrime. RAT was developed to explain traditional crime, but its elements are applicable to understanding malicious activity in cyberspace. RAT research focusing on cybercrime has yielded mixed results in terms of identifying the interaction between those elements intrinsic to the theory that increase risk victimisation[20]; however, this might be because of different variables and methods used for analysis. Therefore, it is imperative to measure different types of cybercrime to obtain valuable data that can be used to understand malicious online activities using RAT.

2.1.2.2 Rational Choice Theory

The Rational Choice Theory (RCT) was developed to understand a wide variety of traditional criminal offences. RCT proposes that offenders consider the prospective rewards and costs of committing a crime and then make rational choices or decisions based on this reasoning. The greater the benefits and the lesser the consequences, the more likely the offence will be committed [22]. These choices may not be rational all the time because judgements are subject to the offender's availability of information and cognitive abilities [23]. Therefore, offenders may not have all the necessary information before making a decision.

Regarding cybercrime, the assumption of rational choice is logical. Miscreants perceive a low level of risk since the offence can be committed from the comfort of their homes, and there is no physical contact with the victims. Similarly, there is a certain level of rationality. Due to the complexity of computer crimes, cybercriminals carefully consider the required skills and the necessary equipment to conduct a crime successfully [24]. Many studies have applied RCT to understand cybercrime [25, 26, 27]. For instance, some evidence about peoples' perceptions of a range of

computer crimes shows that they believed the likelihood of being caught was low and the punishment not severe [28], [29]. Similarly, another study found that university students self-reporting unauthorised use of computers was unrelated to the perceived probability of being caught.

Overall, cybercrime research using RCT shows that miscreants engage in illegal activities when they perceive a low chance of being detected and punished [25]. However, when they are aware of a high likelihood of detection, a deterrence effect will influence their decisions. This deterrence effect is greater when the decision is based on the risk of detection than on the consequences of punishment [29]. Evidence shows that the perceived benefits of a malicious activity outweigh the risk that cybercriminals may consider in committing the act [26]. However, as mentioned before, since they might not have all the information available to commit a crime, they may not always make the best decisions.

Research suggests that RCT is a plausible approach when analysing cybercrime. The findings demonstrate that a combination of different motivators and a thinking process explain why individuals engage in online illicit activities [27]. Understanding the process of making a decision is key to proposing countermeasures aimed to decrease the benefits and increase the costs/risks of malicious activities. This decision-making process including how cybercriminals calculate costs and benefits depends on how cybercriminals plan and conduct their operations; therefore, the measurement of different types of cybercrime provides the necessary data to carry out this type of analysis.

2.1.2.3 Situational Crime Prevention

Situational crime prevention (SCP) is an approach based on many theories, such as rational choice theory and routine activities theory. While the former theories focus on understanding the crime event, SCP aims to control and prevent crime. SCP proposes techniques to reduce opportunities for crime by increasing the perceived difficulty, increasing the risks, reducing the rewards, reducing provocations and removing excuses for offending [30, 31].

A considerable amount of literature has investigated the application of the

SCP approach to reduce cybercrime [32, 33, 34]. As a result, modified versions of the SCP techniques have been developed, and different frameworks have been proposed to prevent various computer crimes [12, 35, 36, 37, 38]. Although these studies are primarily theoretical, they conclude that SCP techniques can be beneficial in effectively reducing opportunities for cybercriminals, especially in the field of Information Security.

Research on SCP has focused on analysing environmental factors that create opportunities for different types of cybercrime. For instance, opportunity structure analysis provides comprehension of situational factors that can be modified or blocked to address several cybercrime-related opportunities. The results suggest that understanding cybercrime opportunities informs the InfoSec field to propose more effective technical interventions [32]. This is consistent with other findings in the literature showing that technical SCP measures are associated with the prevention of online crime [33].

Some authors argue that cybercrime should be tackled from a criminological perspective with a focus on the environmental factors to understand how the crime is committed [12]. The evidence shows that SCP provides insights to understand online malicious activities and tools to address the issue from a holistic perspective taking into account crime opportunity factors and the technical aspect of cybercrime. For example, target hardening techniques such as firewall systems prevent intrusions by making ICTs more difficult to penetrate [33]. As a result, SCP provides practical and applicable knowledge that can be applied to the circumstances associated with online crime, and how the environment can be modified to prevent it [38].

As stated by rational choice and routine activity theories, potential offenders act with a certain degree of rationality and make decisions based on the opportunities provided by the presence of a suitable target and the absence of a guardian. Since SCP interventions reduce these opportunities, criminals seek new approaches to perform their malicious activities. As a consequence, the problem of crime displacement and crime adaptation arises [13].

Crime displacement involves the relocation of crime or criminals from one

space (place), time, target, type of offence or modus operandi to another [39]. On the other hand, crime adaptation is when criminals discover new crime vulnerabilities and improve their modus operandi to overcome the limitations of the interventions [14]. As regards cybercrime, the ability of cybercriminals to displace or adapt their operations is relativity quick and easy in cyberspace. Thus, it is vital to understand the modus operandi of cybercriminals to refine SCP interventions to anticipate malicious activities and act accordingly.

2.1.2.4 Social Learning Theory

Although social learning theory is not considered part of environmental criminology theories, it can help explain cybercrime from an ”environmental” point of view. Social learning theory is centred around the idea that motivations and skills for criminal behaviour are learned in an environment where potential offenders associate or interact with deviant peers [40].

A certain level of knowledge and skills are required to conduct computer crimes. Consequently, individuals frequently gain insights about strategies and methods to perform these malicious activities from others [41]. Moreover, networks connecting cybercriminals are built to spread knowledge about engaging in different types of cybercrime [42]. It is now well established from a variety of research that there is a relationship between deviant peers, the acceptance of illegal behaviour and cybercrime [41, 42, 43, 44].

2.1.3 The Information Security Approach

Over the past few decades, the information security field has conducted a great deal of research on cybercriminal activity. Scholars became particularly interested in obtaining realistic measurements of the scale and extent of cybercrime by studying how technology is used or abused to commit offences. One of the main objectives is to analyse data related to computer-related malicious activity to understand the modus operandi of cybercriminals and how they use ICT to perform different types of cyberattacks.

Much of the preliminary work to measure different types of malicious activity

relied on data and findings drawn from various surveyed populations, which was remarkably inconsistent [45]. Moreover, due to the relative anonymity that computer crimes provide, a large proportion of offences are not reported [46]. Consequently, there has been a lack of official data and statistics related to cybercrime [47].

Since obtaining data has been a significant challenge, InfoSec researchers have devised innovative ways to collect primary data from ICT systems, networks and cyberspace in general. To this end, they have developed crawlers, deployed honeypots and monitored digital infrastructure to collect data from different applications and platforms such as websites, Web servers, blogs, social networks, forums and databases. Analysing this data is crucial to understanding the cybercrime ecosystem, identifying vulnerabilities and designing countermeasures for mitigation.

Several studies have begun to examine the benefits of joining environmental criminology and InfoSec research to develop better mitigations for cybercrime [13, 38, 35]. These studies demonstrate the need to understand the environmental factors of cybercrime as well as its technical components. In the following sections, we will examine some forms of cybercrime by applying the InfoSec approach to gain insight into how cybercriminals conduct their operations in the wild.

2.2 Types of Cybercrime

According to some authors, cybercrime can take different forms, including cyber-trespass (unauthorized system access), cyber-deception/theft (identity theft, online fraud, digital piracy), cyberporn/obscenity (child sexual exploitation materials), and cyber-violence (cyberstalking, cyber terrorism) [48, 47]. There is a vast amount of literature on these types of malicious activity and evidence demonstrates that cybercrime incidence is on the rise while traditional crime incidence continues to decline [49].

Much of the early research in this field was devoted to determining how cybercrime and cyberspace varied from traditional crime and terrestrial space. The results showed that cybercrime can be explained using criminological theories but the exponential growth in the development of ICTs increases the opportunities for cybercrime

[50]. Therefore, in the last few decades, there is a considerable and growing amount of literature focused on the technical aspect of cybercrime, including online fraud, malware distribution, identity theft, and cyberattacks against critical infrastructure and smart devices. These studies show how important it is to understand both the human and technical aspects of cybercrime [47].

Complex forms of cybercrime have been developed by cybercriminals to achieve their malicious goals; therefore, it is important to understand how cybercriminals conduct these malicious activities so law enforcement agencies and the research community are better prepared to tackle cybercrime. This thesis studies three forms of cybercrime focusing on malicious activities related to stolen accounts credentials, malvertising and underground forums. Each activity will be examined in the following chapters, each of which has its own literature review.

2.3 Layers of the Internet used for Cybercrime

The Internet is a vast network that connects many independent networks and computers hosting resources that enables the exchange of information and daily transactions in a rapid and decentralised fashion. Cybercriminals take advantage of these features to perform malicious activities targeting millions of potential victims accessing online resources around the world. There are distinctions in the way online resources are accessed by individuals, based on the use of either common or specialised web browsers and encryption protocols. Online resources can be accessed through three layers or sections on the Internet: the Surface Web, the Deep Web and the Dark Web. Miscreants may commit malicious activities in any of these layers.

A growing body of literature has examined malicious activities on the Surface Web. The Surface Web is made up of those Web resources that can be searched and accessed by a Web search engine such as Google, Bing or Yahoo. As law enforcement agencies increasingly monitor the Surface Web, cybercriminals become more sophisticated and continue to improve their methods and techniques to engage in deviant behaviour without getting caught [51]. For this reason, they are migrating to anonymous networks within the Deep Web.

The Deep Web refers to websites not indexed by a search engine but can be directly accessed via a Web address. Additionally, owners of these websites can block access to the content. The Deep Web hosts confidential corporate websites, databases of private companies, scientific and academic databases, medical records, etc. As part of the Deep Web is the Dark Web. The Dark Web refers to websites or Web resources on a darknet. Darknet is an encrypted network built on top of the Internet designed specifically for anonymity and is accessible through specific software and tools. Examples of a Darknet are The Onion Router (Tor), I2P, Freenet and DN42. Therefore, the Dark Web contains websites or resources whose content has been intentionally concealed [52] and are known as hidden services.

2.4 Hidden Services

Tor is the most widely used anonymity network. It was a military initiative created in the '90s by the Defense Advanced Research Projects Agency (DARPA) in the U.S. Naval Research Labs (NRL). Initially, the goal of the project was to develop a method for anonymising communications traffic to keep secret the identities of law enforcement officials on the Internet. For instance, security agents would communicate with their teams to avoid detection or browsing websites without a government IP address appearing on traffic logs [53]. Tor is typically used to anonymously access the Internet by journalists, whistleblowers, or law enforcement officials who want to maintain their privacy.

Since the research community has developed better techniques to detect and catch offenders who perform illegal activities on the Surface Web, hidden services based on the Dark Web have become more prominent over the last few years [9]. A considerable amount of literature suggests that cybercriminals are moving their operations to the Dark Web[54, 55]; however, it is not yet known whether such operations increase victimisation risks to specific attacks.

The Dark Web poses a significant challenge since the actors' identities on this platform remain largely unknown, and law enforcement agencies do not have enough resources to stop or deter illegal activities. These facts represent strong incentives

for criminals to use them. Thus, it is crucial to understand the behaviour of criminals on the Dark Web. As there is no sound information available about it so far, this study shall provide some insight by measuring the malicious activity resulting from different online crimes.

Chapter 3

Stolen Account Credentials

This chapter will explain our work related to stolen online accounts based on the paper published in the Crime Science Journal in 2018 [51], including the methodology and results presented below. We examine malicious activity on compromised online accounts on the Surface Web and the Dark Web to understand criminal behaviour exploiting those accounts.

3.1 Introduction

Individuals and companies increasingly use online services to carry out everyday transactions over the Internet for personal, commercial, and academic purposes. To have a unique identity and maintain a distinctive profile within these online services, users must create accounts that generally consist of a username and password. Such personal accounts retain a lot of sensitive information, and some of them, like webmail accounts, are primarily used to access other services. As a result, users may become victims of identity theft by cybercriminals, who steal credentials for their benefit. According to the Crime Survey for England and Wales, one in ten adults has been a victim of data theft related to their identity [56].

Cybercriminals use social engineering techniques such as phishing and spear phishing [2], malware on victims' devices [3] and also exploiting vulnerabilities in authentication databases [4, 5] to steal user credentials. After obtaining the credentials, criminals can monetise the accounts in different ways. They seek sensitive information such as credentials to other online services, financial information

and even intimate information that can be used to blackmail the victim. Similarly, they can be used to send spam or spear-phishing emails to other victims. Finally, credentials can be used as goods that are traded or shared in underground outlets.

It is a great challenge for researchers to determine what happens when an account has been compromised. Previous research focused on understanding the use of stolen accounts in the Surface Web, i.e., the portion of the Internet where websites are indexed in the search engines and it is accessible with any browser. Onaolapo et al. [6] studies the activity of cybercriminals accessing compromised Google accounts leaked through different outlets. Lazarov et al. [7] monitor criminal activity on leaked Google spreadsheets. Similarly, Bernard-Jones et al. [8] investigate the effects of language on cybercriminals navigating compromised webmail accounts.

However, at the same time, cybercriminals are continuously improving methods and techniques to engage in outlets of compromised data without getting caught or blocked. For instance, the Dark Web provides a network where content providers can anonymously publish content and cybercriminals can distribute and monetise compromised accounts. This suggests that criminals are using hidden outlets on the Dark Web to find or trade stolen account credentials [54, 11]. According to *Top10VPN.com*, the world's largest Virtual Private Network review site, someone's online identity is worth £820 to miscreants on the Dark Web as of February 2018.

Although some research has investigated various types of illegal activities on the Dark Web [54, 55], very few studies have compared how compromised accounts are used in both environments: the Surface Web and the Dark Web. As such, this chapter aims to address this gap by comparing the results of the experiment performed by Onaolapo et al. [1] on the Surface Web with the results of a similar experiment performed on the Dark Web. The new experiment follows Onaolapo's methodology which is the use of honey accounts. These accounts resemble legit email accounts from common users. For this work, Gmail accounts were leaked through outlets within online services of the Tor network. Data from both experiments were collected and analysed to provide some insights into the differences related to stolen credentials in both environments. This study addresses the following research question:

RQ1: Does the Web environment and the outlet where stolen accounts are leaked influence the type and amount of activity that those accounts receive?

To this end, we monitored during a period of one month the use of compromised Gmail accounts in the Dark Web using the infrastructure proposed by Onaolapo et al. [1]. For that purpose, we created fake Gmail accounts (called honey accounts) whose credentials were leaked in various outlets such as paste sites (online outlets where users can store and share plain text) and underground forums. The intention of the experiment is to make cybercriminals interact with these credentials. Then, all actions related to the emails in the accounts are recorded, namely when a mail is read, favourited, sent or a new draft is created. Similarly, we tracked the access to each account in order to obtain the system information and the origin of the login session. In summary, this study makes the following contributions:

- We studied the activity generated on email accounts whose credentials were leaked in different outlets of the Dark Web.
- We compare the results of this experiment with those obtained with one conducted with a similar methodology on the Surface Web [1] to gain insights into the modus operandi of cybercriminals interacting with stolen accounts in different layers of the Internet. Our results show distinctive differences between both Web environments regarding malicious activity depending on the leakage outlet.
- Using the data collected, we published a dataset containing the intrinsic characteristics of accesses to stolen accounts in a repository open to the public[1].

The results suggest that stolen accounts are more likely to receive unwanted accesses when they are leaked on the Dark Web, especially on paste sites. The analysis of the activity performed on those accounts indicates that most access events are from curious actors who may be testing the credentials but do not perform any other activity. However, some of them repeatedly log in to the same account

[1] `https://bitbucket.org/gianluca_students/surface_vs_dark_credentials_dataset`

presumably to look for new relevant information. On the other hand, the highly frequent use of unknown browsers suggests an attempt to hide the browser used in the access.

3.2 Background and related work

Online services such as email, storage, banking, and team collaboration tools provide an infrastructure in which users can communicate with one another by exchanging information. Personal credentials, which normally consist of a username and password, are used to authenticate to a service and get access to their accounts. Large amounts of sensitive data are treated within such environments every day, such as financial information, authentication details, copyrighted material, classified data, etc.

3.2.1 Credential Theft

Credential theft is a type of identity theft in which cybercriminals steal credentials from a legitimate user to access online accounts and obtain the same account privileges as the victim. Once the credentials have been stolen, accounts are abused for financial gain. For instance, accounts are used to send spam or phishing attacks, make illicit purchases, withdraw money from bank accounts or ultimately, credentials can be used as goods traded or shared in underground forums or black markets. Moreover, webmail accounts are especially valuable for searching information from other online services and can be used to take over other accounts such as banking or e-commerce.

3.2.2 Criminal Implications

Criminologists agree that individuals' online routines are creating new opportunities for identity theft. Reyns [57] applies the routine activity theory to crimes in which the victim using online services and the offender never have a physical encounter. According to the findings, individuals using online services are 50 per cent more likely to be affected by identity theft. This demonstrates the high risk of victimisation that people face when engaging in online activities.

3.2.3 Credential Acquisition

Cybercriminals use a variety of tactics to steal account credentials. Lynch [58] examines phishing, which involves criminals sending fake emails that appear to be from genuine online services and deceiving their victims to introduce personal credentials on a fake website. Similarly, spearphishing attacks use fraudulent emails that are directed to one or a specific group of users [59].

Another technique is to infect users with malware aimed to steal sensitive data [60]. This attack is particularly successful because a large number of users have little or no protection to counter this threat. Additionally, vulnerabilities in online user databases can lead to massive credential leakage [61]. The research cited above describes stealing strategies to obtain credentials but does not analyse what happens after an account is compromised.

3.2.4 Account Exploitation

Online accounts not only contain valuable information about their owners but also gain a level of reputation and trust among contacts and other online accounts over time. As a result, a growing body of literature has focused on actions undertaken by cybercriminals gaining illicit access to online accounts. Compromised accounts can be used to send spam [62], conduct phishing attacks, find confidential information or liquidate victims' financial assets [63].

Bursztein et al. [63] focus on phishing as a method of stealing credentials and show that attackers evaluate the account's value to decide how to abuse it. Similarly, they send scam attacks to the contacts in the accounts. These attacks can be more effective due to the level of trust the account may have, as contacts are more likely to be deceived and provide sensitive information or even money.

Onaolapo et al. [1] analyse the activities cybercriminals perform on compromised Gmail accounts. Their work involves leaking fake Gmail accounts on paste sites, underground forums or by using malware. The results suggest that criminals look for financial information in the accounts and try to evade Google's security mechanisms by using the location information of the account as the source of connection if this information is provided. Lazarov et al. [64] leaked Google spreadsheets

to observe the actions of criminals illegally accessing cloud-based documents using a similar approach.

3.2.5 Criminal Behaviour

Onaolapo et al. [1] provide an analysis of the interaction between cybercriminals and hijacked accounts when stolen credentials are obtained in outlets hosted on the Surface Web. Based on the observations from the accesses to the honey accounts, they identified a classification of the activity conducted by cybercriminals. There are four types of attackers according to the actions that they perform within the accounts:

- *Curious* log into the honey accounts and perform no further actions in them. They simply access the accounts to check the correctness of the credentials.
- *Gold Diggers* perform searchers on the emails contained in the account to find sensitive information that could be monetised in the underground economy.
- *Spammers* use the honey accounts to send spam messages by exploiting the trust that contacts have with the account owner.
- *Hijackers* change the account password to take complete control of it, preventing the account's original owner from having access.

3.2.6 Credential Release

Although stolen accounts can be exploited through different schemes, some criminals decide to monetise them by trading the credentials in different outlets. This is an example of how criminal activities have led to an exponentially growing digital underground economy where illicit goods and services are traded [65].

Online account credentials can be considered goods to be exchanged within this economy in several outlets such as underground forums and black markets. Holt and Lampke [66] analysed the underground markets in which criminals release or trade accounts obtained through malicious activities. In some instances, online credentials are freely released for the authors only to build a reputation within the underground community [67]. On the other hand, criminals seek financial gain and sell the stolen accounts to other criminals to monetise them [68].

Ablon et al. [69] argue that trading stolen data has become lucrative and easier to carry out than other types of illegal trade. Furthermore, a growing body of research has shown that personal and financial data can be obtained through markets for stolen data at a fraction of their true value [66]. Therefore, there is a considerable exchange rate of stolen credentials in the underground economy, which are exchanged in different outlets. The research mentioned above refers specifically to outlets located on the Surface Web.

3.2.7 Stolen Accounts in the Dark Web

A growing body of literature has studied how criminals exploit stolen credentials on the Dark Web. Large-scale data breaches are the main source of credential availability, and there is a massive demand for them, which increases their price [70]. Consequently, more miscreants are trading stolen credentials on the Dark Web through underground forums and black markets. Moreover, stolen credentials may be used to help them earn a reputation or trade for other illicit goods.

Liu et al. [71] suggest that the Dark Web is a valuable source of stolen credentials that are used for monetisation purposes through identity theft. Similarly, Biryukov et al. [72] collected and classified addresses of 39,824 hidden services in Tor to show that half of them were selling stolen accounts. These studies provide insight into the extent of stolen credential dissemination on the Dark Web. We focus on how the stolen accounts are abused by analysing criminal interaction with the accounts after credentials are obtained at Dark Web outlets.

3.3 Experimental Setup

We conducted a new experiment on the Dark Web using the honeypot infrastructure for the Surface Web experiment proposed by Onaolapo et al. [1] (see Figure 3.1). The experiment aims to imitate the way cybercriminals operate by releasing or trading stolen account credentials through some outlets on the Dark Web, specifically in some hidden services within the Tor network. The infrastructure tracks the actions performed by criminals who had the account credentials in their possession. The results of the experiment on the Dark Web are paired with the results of Onaolapo's

investigation on the Surface Web to draw comparisons. For the sake of the comparison, we followed the same methodology used in the Surface Web experiment, i.e. leaking the same number of accounts across the same type of outlets.

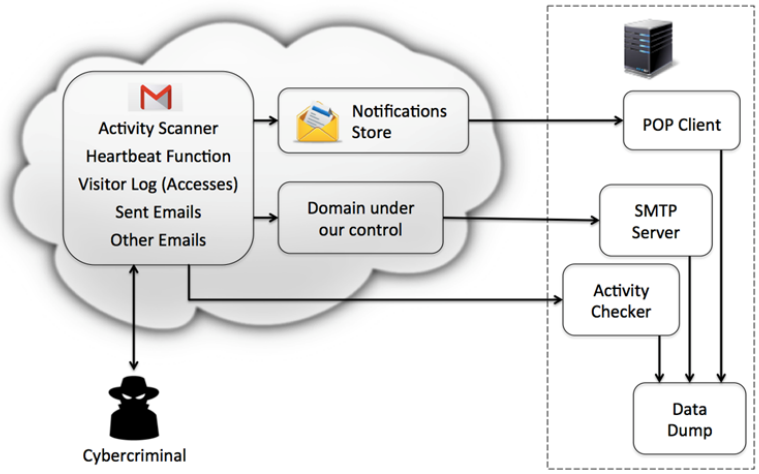


Figure 3.1: Overview of the monitoring infrastructure presented in [1]. Any login activity on the honey accounts is recorded on our side, as well as any further activity performed on them.

The first step of the experiment was the creation of Gmail accounts known as honey accounts. These accounts resemble legitimate email accounts from standard users. In the creation phase, 100 honey accounts were created manually on Gmail. The same number of accounts used in the Surface Web experiment was used in the Dark Web experiment to add consistency to the comparison. The fictitious data used to create the accounts was automatically generated using a database of random names for the accounts. The combination of names (first name and last name) was double-checked to verify that they resemble common user names.

Account creation represented a major challenge for this study because Gmail limits the number of times an account can be created from the same computer/network without the need for a phone number. This is a protection mechanism that Gmail uses to prevent abuse of the system and stop robots from registering many accounts at once. Most of the time, it was possible to create only one account from a single computer/network. Then, a phone number was required to verify the creation of more accounts. However, only one more account could be created after the verification. As there were no more phone numbers available for the experiment, the creation

of accounts took place in various places around London and Liverpool such as universities, libraries, internet cafes and more.

All the accounts were populated with email messages from the Enron dataset to simulate a real email account belonging to a regular user. Enron was an energy company declared bankrupt in 2001, and the emails dataset from the company executives was made available to the public. This corpus contains a total of 517,431 messages from 150 users [73]. Enron dataset was chosen because it simulates a real enterprise-scale communication environment. Each account received at least 200 emails. The emails were sent in batches before and after the leak to resemble an active user account that handles a lot of information. The first names, last names and the word ”Enron” were replaced in all the emails using fictitious names.

In the next phase, the accounts were instrumented with scripts to monitor and register activity from visitors. The monitoring infrastructure incorporates Google Apps Scripts hidden in a Google Sheet as a typical document within each account. Google Apps Script is a JavaScript cloud scripting language used to automate different time-based and event-based tasks across Google products [74]. Scripts were used to monitor actions over the emails by scanning the messages to determine if an email has been read, sent, marked as important (Starred) or if drafts have been created.

All these changes were registered and reported by the scripts through email notifications from each honey account. These notifications were sent to a particular email account created for this purpose. After sending a notification, the script deleted this email in the ’Sent’ folder to avoid any suspicion. Other features of the scripts included sending a heartbeat message daily to identify if the account is still functional and has not been locked by Google or if an attacker has changed the password.

We hide the scripts embedded in a Google Sheet inside each honey account. The spreadsheet was designed to look like a regular document in the owner’s account, making it unlikely for criminals to find this out and remove the script. Finally, the default `send-from` address of the honey accounts was altered so that when an email is sent from any of them, the message is directed to a controlled SMTP mail

server that was set up to receive and store these emails without forwarding them to the intended destination. The send-from address was changed using the settings menu within each Gmail account. This measure was taken to avoid abuse from cybercriminals. It is worth mentioning that Google Apps Script stops working when the account password is changed because it requires the permissions to be updated to keep running.

Similarly, other scripts extracted more information from the 'Device activity and notifications' section within each account's Gmail account management dashboard. This section uses Google fingerprinting system to extract data from the cookies generated by each login to the accounts. A cookie is a small piece of data sent to a browser by a Web server when a user requests a Web resource. Cookies are designed to be a reliable mechanism for websites to remember session information or record the user's browsing activity [75]. Cookie information includes cookie identifier, public IP address, location, login time, browser, and the device's operating system from which the login originated. Each cookie found in our dataset is considered as a unique access to one particular account. As will be explained later, leaking the accounts on the Dark Web does not imply that the accounts will be accessed through Tor. In fact, this is very unlikely because Gmail usually blocks login attempts from Tor.

As in the Surface Web experiment, the outlets chosen for the leaks are paste sites and underground forums. However, malware was not used for the Dark Web experiment because its use does not depend on the Web environment. The idea behind leaking the accounts in different outlets is to compare malicious activity among them. The experiment was performed using 100 accounts divided into groups, each group leaked on various hidden services within Tor.

The hidden paste sites chosen were `Insertor` and `Stronghold`. In terms of underground forums, the hidden services used were: `AlphaBay`, `Silk Road Forum` and `KickAss`, where there are many threads regarding illegal activities, such as data theft. The selection of sites was due to the similarity to outlets used for the Surface Web in terms of the degree of activity, as we observed many posts and messages

exchanged daily by members. The Surface Web experiment used `pastebin.com` and `pastie.org` for paste sites. For underground forums, `offensivecommunity.net`, `bestblackhatforums.eu`, `hackforums.net` and `blackhatworld.com` were selected. Furthermore, the chosen sites allow visitors to post without registration. While traffic is an important variable to consider in the experiment, we were unable to get statistics from these hidden services due to their nature. Therefore, no traffic differences could be established among the sites. We acknowledge the limitation, and we will discuss it later.

As in the Surface Web experiment, some groups of honey accounts were leaked using basic credentials: user name and password. For the remaining groups, a location was added to the leak. This is the location from which the legitimate owner is assumed to access the account. The objective was to show that accesses to compromised accounts depend on the information provided in the leakage. Hence, the locations used were near London for leaks based in the United Kingdom and Pontiac, Michigan for leaks based in the United States.

Activity on the honey accounts was recorded for about seven months for the Surface Web and one month for the Dark Web, which was the period covered for our ethics approval. However, we extracted the first month of observations from the Surface Web experiment for the comparison to be homogeneous. We chose the first month to replicate the same features in both environments as if the Surface Web experiment would have been performed for only one month to make sure not to introduce any statistical bias.

This chapter seeks to determine whether any characteristic of the accesses is associated with the environment where the credentials were exposed. The data gathered from both experiments may be helpful for researchers to understand how attackers interact with stolen webmail accounts and how this malicious activity differs in the Surface Web and the Dark Web. Therefore, we have publicly released an anonymised version of the data for academic purposes.

3.4 Data Analysis

The Surface Web experiment identified 164 unique accesses to the accounts after the leak; on the other hand, 1092 unique accesses to the Dark Web accounts were recorded in our experiment (see Table 3.1). It is important to note that even though the credentials are leaked in Dark Web outlets, they are not always accessed from the Tor network. Thus, in our analysis, the Dark Web statistics refer to accounts that have been exposed but not accessed through Tor. In fact, only 378 accesses originated from the Tor network. To perform our statistical tests, we coded the collected data into the following variables: cookie identifier, Web environment, IP address, outlet, taxonomy, login time, location, browser and the operating system of the access.

	Paste Sites	**Underground Forums**	**Total**
Surface	51.8%	48.2%	164
Dark	90.8%	9.2%	1092

Table 3.1: Unique accesses depending on the outlet. Paste sites are more likely to be used by cybercriminals in the Dark Web ($\chi^2 = 177.587$, p<.001).

We used a chi-square test [76] to determine whether a relationship exists between Web environment and outlet. The results showed a significant relationship ($\chi^2 = 177.587$, p<.001). Regarding paste sites, most accesses originate from accounts leaked on the Dark Web (90.8%) compared to the Surface Web (51.8%); for underground forums, more logins come from credentials obtained on the Surface Web (48.2%) compared to the Dark Web (9.2%). This suggests that the exposure of stolen credentials is higher in Dark Web paste sites. On the contrary, underground forums on the Dark Web are less accessible since, as we noticed, a great deal of them requires an invitation or referral to access them.

3.4.1 Taxonomy of Account Activity

Based on our observations on the honey accounts and the classification or taxonomy mentioned in previous sections, the following accesses were identified in the Surface Web: 103 *Curious*, 39 *Gold Diggers*, 2 *Spammers* and 20 *Hijackers*. On the other hand, 812 *Curious*, 227 *Gold Diggers*, 39 *Spammers* and 14 *Hijackers* were registered

on the Dark Web (see Table 3.2).

	Curious	Gold Digger	Hijacker	Spammer	Total
Surface	62.8%	23.8%	12.2%	1.2%	164
Dark	74.4%	20.8%	1.3%	3.6%	1092

Table 3.2: Unique accesses depending on the taxonomy. Hijacking is more likely to occur on the Surface Web (FET: $p<.001$)

We performed a Fisher's Exact Test (FET) [77] to observe whether there is a relationship between Web environment and taxonomy. In this case, we are not using a chi-square test to find significant differences because our contingency table has cells with expected frequencies of less than 5, which violates an assumption of this test. The test revealed a significant association between Web environment and taxonomy ($p<.001$, 99% CI) but a Cramer's V statistic showed that the strength of the association is weak ($V = 0.233$). This result is for the overall analysis and a post-hoc is performed to find individual significances. We rely on a method that yields probability values for each combination of independent category levels and uses a Bonferroni correction to control for type I error inflation [78, 79]. The test reports the percentage contribution for each cell to the overall chi-square statistic. We found a significant association between the Web environment and *Hijackers* ($p<.001$). *Hijacking* is more likely to occur on the Surface Web (12.2%) than on the Dark Web (1.3%), where this event is rare.

The variable outlet is included for further analysis. Figure 3.2 shows the distribution of accesses in each outlet and the taxonomy for each Web environment. Our FET test revealed that this association is significant ($p<.001$, 99% CI) only in paste sites (see Table 3.3). The analysis suggests that while *Gold digging* is more likely to happen on the Dark Web, *Hijacking* is more frequent on the Surface Web. This may be an indication that attackers on the Dark Web are searching for valuable information for monetisation purposes. At the same time, they try to go unnoticed without changing the password in the accounts, which in turn indicates a certain level of sophistication. Regarding the underground forums, the observed differences

are not significant.

		Curious	Gold Digger	Hijacker	Spammer	Total	Statistics (FET)
Paste Sites	**Surface**	63.5%	17.6%	18.8%	0.0%	85	$p < .001$
	Dark	74.5%	20.3%	1.3%	3.9%	992	
Forums	**Surface**	62.0%	30.4%	5.1%	2.5%	79	$p = .099$
	Dark	73.0%	26.0%	1.0%	0.0%	100	

Table 3.3: Distribution of accesses for each outlet and taxonomy class. There are significant differences between the Surface Web and the Dark Web when credentials are leaked through paste sites and used to hijack an account (FET: p<.001).

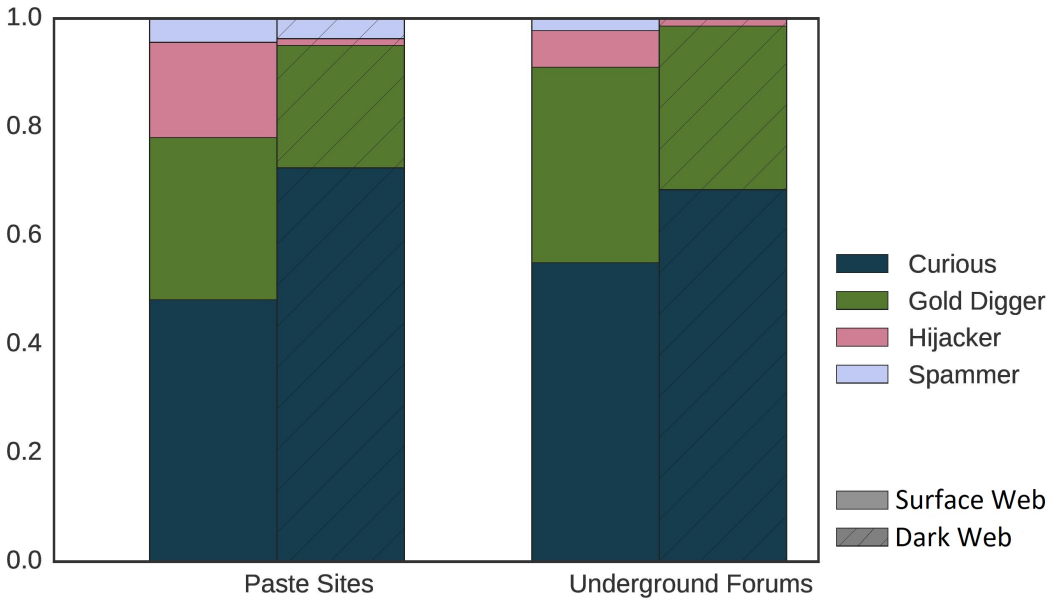


Figure 3.2: Distribution of accesses in each outlet and taxonomy class.

3.4.2 Device Configuration of Accesses

Google's system fingerprinting was used to collect information about devices accessing the honey accounts. Figure 3.3 illustrates the distribution of Web environment, operating system in each outlet where the credentials were leaked. There is a significant association (p<.001, 99% CI) between the operating system and the Web environment when credentials are obtained in paste sites (see Table 3.4). However, this association is weak (V = 0.198). Although most accesses originate from Windows, our post-hoc analysis revealed that cybercriminals are more likely to use Android devices when using credentials gathered on the Surface Web than on the

Dark Web (15.3% vs 1.1%, p<.001). This may indicate a low level of sophistication as users are probably using their own mobile devices to access the accounts. On the other hand, Linux is more likely to be used on the Dark Web (22.5% vs 7.1%, p<.001). It is reasonable to assume that Linux is used by more skilled criminals, which is consistent with the evidence that there is a higher level of sophistication on the Dark Web. In the case of underground forums, the observed differences are not significant.

		Android	Chrome OS	iOS	Linux	Mac	Unknown	Windows	Statistics (FET)
Paste Sites	**Surface**	15.3%	0.0%	0.0%	7.1%	4.7%	9.4%	63.5%	p < .000
	Dark	1.1%	0.0%	1.7%	22.5%	3.4%	8.1%	63.2%	
Forums	**Surface**	0.0%	1.3%	0.0%	8.9%	1.3%	0.0%	88.6%	p = .031
	Dark	7.0%	0.0%	2.0%	7.0%	5.0%	1.0%	78.0%	

Table 3.4: Distribution of accesses for each outlet and operating system. Most accesses originate from Windows; however, cybercriminals in paste sites are more likely to use Android devices when using credentials gathered in the Surface Web. On the other hand, Linux is more likely to be used on the Dark Web (FET: p<.001).

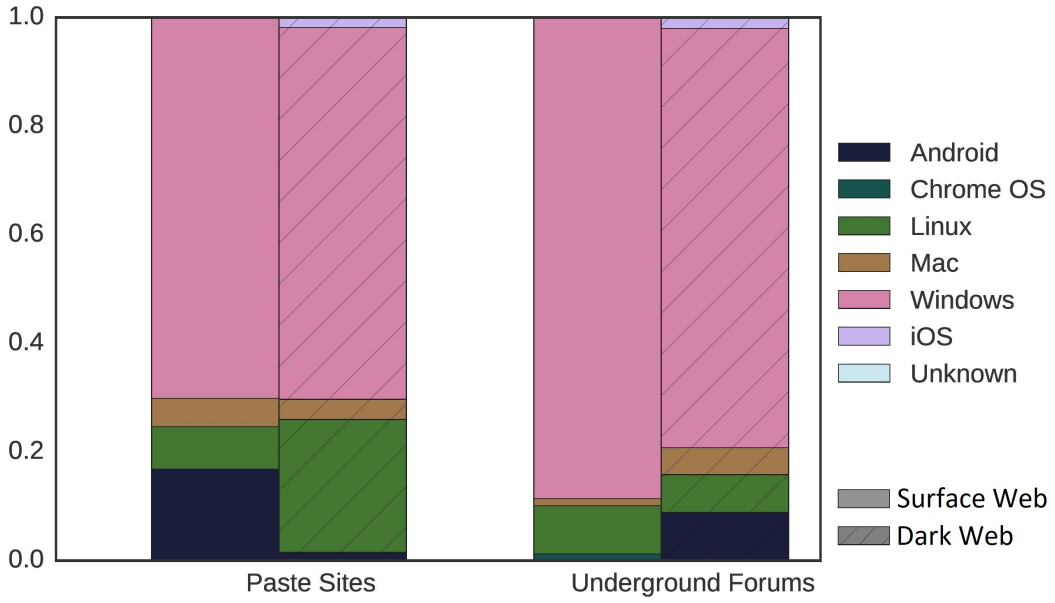


Figure 3.3: Distribution of accesses for each outlet and operating system.

The browser distribution is outlined in Figure 3.4. There is a significant association (p<.001) between Web environment and browser (see Table 3.5). The post-hoc test shows that unknown browsers are more likely to be used on the Dark Web (60%) than on the Surface Web (39.9%) for paste sites (p<.001). While this may be an

indication that criminals attempt to hide the browser user agent from the Google fingerprinting system when accessing the accounts, one could easily argue that any sophisticated attacker would use a standard user agent to avoid triggering detection mechanisms when trying to log in. The collection of further data and an analysis of the accuracy of Google's fingerprinting system would be necessary to draw strong conclusions about this aspect. Similarly, there is a significant association between Web environment and Chrome for both outlets ($p<.001$). The use of Chrome is more likely to happen on the Surface Web for paste sites and underground forums. Interestingly, in the Dark Web, we got five accesses from Mozilla Thunderbird clients. This suggests that several attackers, such as *Gold Diggers* or *Spammers*, are using the functionalities of this email application to abuse the accounts.

		Chrome	Edge	Firefox	Iceweasel	Internet Explorer	Opera	Safari	Thunder-bird	Unknown	Vivaldi	Statistics (FET)
Paste Sites	**Surface**	27.1%	0.0%	14.1%	4.7%	9.4%	2.4%	0.0%	0.0%	40.0%	2.4%	$p < .000$
	Dark	11.4%	0.3%	25.9%	0.0%	1.0%	0.8%	0.5%	0.0%	60.1%	0.0%	
Forums	**Surface**	31.6%	0.0%	19.0%	0.0%	0.0%	0.0%	0.0%	0.0%	49.4%	0.0%	$p < .000$
	Dark	12.0%	0.0%	38.0%	0.0%	0.0%	6.0%	0.0%	5.0%	39.0%	0.0%	

Table 3.5: Distribution of accesses for each outlet and browser. There is a significant association between the Web environment and the use of unknown browsers on the Dark Web (FET: $p<.001$).

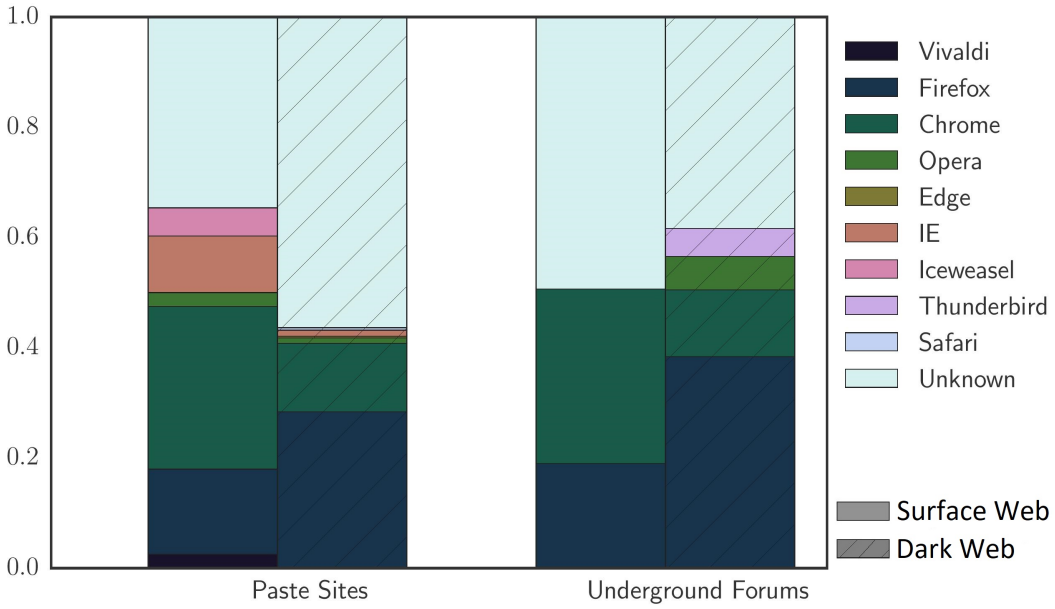


Figure 3.4: Distribution of accesses for each outlet and browser.

3.4.3 Duration of the Accesses

A cookie identifier and a timestamp of access are generated when a new login occurs in a honey account. As a result, each cookie in our dataset has timestamps of the first access and the last known access to a honey account. We used these timestamps to determine the length of access of a cookie for each unique access [1].

Figure 3.5 shows the Cumulative Distribution Function (CDF) of the lengths of accesses to the accounts on the Surface Web and the Dark Web. Most accesses were short, being less than a day, meaning that most visitors accessed the honey accounts only once and did not return. However, Dark Web accesses had a longer time between subsequent interactions with the accounts compared to the Surface Web for all taxonomies. Approximately 30% of Dark Web *Curious* logins connected to the accounts several days after the first login, and only less than 5% did it on the Surface Web. For *Gold Diggers*, the trend is the same (approximately 20% vs 5%). In the case of *Hijackers*, about 10% of accesses continued taking place during this period in both Web environments. However, this indication may not be entirely accurate because it represents the length of the access until the cookie was hijacked. *Spammers* on the Surface Web sent emails in bursts for a short period (less than a day). Conversely, spam on the Dark Web occurred for almost ten days.

3.4.4 Location of Accesses

When a user logs into a Gmail account, Google's geolocation system detects the origin of the login. In this way, we obtain information about the city and country where the login originated. However, this information cannot always be obtained as some accesses show an "unknown" location. According to information provided by Google, those accesses originated from Tor exits nodes or anonymous proxies. Thus, it is reasonable to believe that some attackers use Tor to access the accounts the same way they found the credentials.

From the data gathered in the Surface Web experiment, 148 unique accesses with the location were recorded, and 16 unique accesses originated from Tor exit nodes. Also, accesses from 29 countries were observed. In the Dark Web, 714 logins contain location information, and 378 accesses originated from Tor. Available

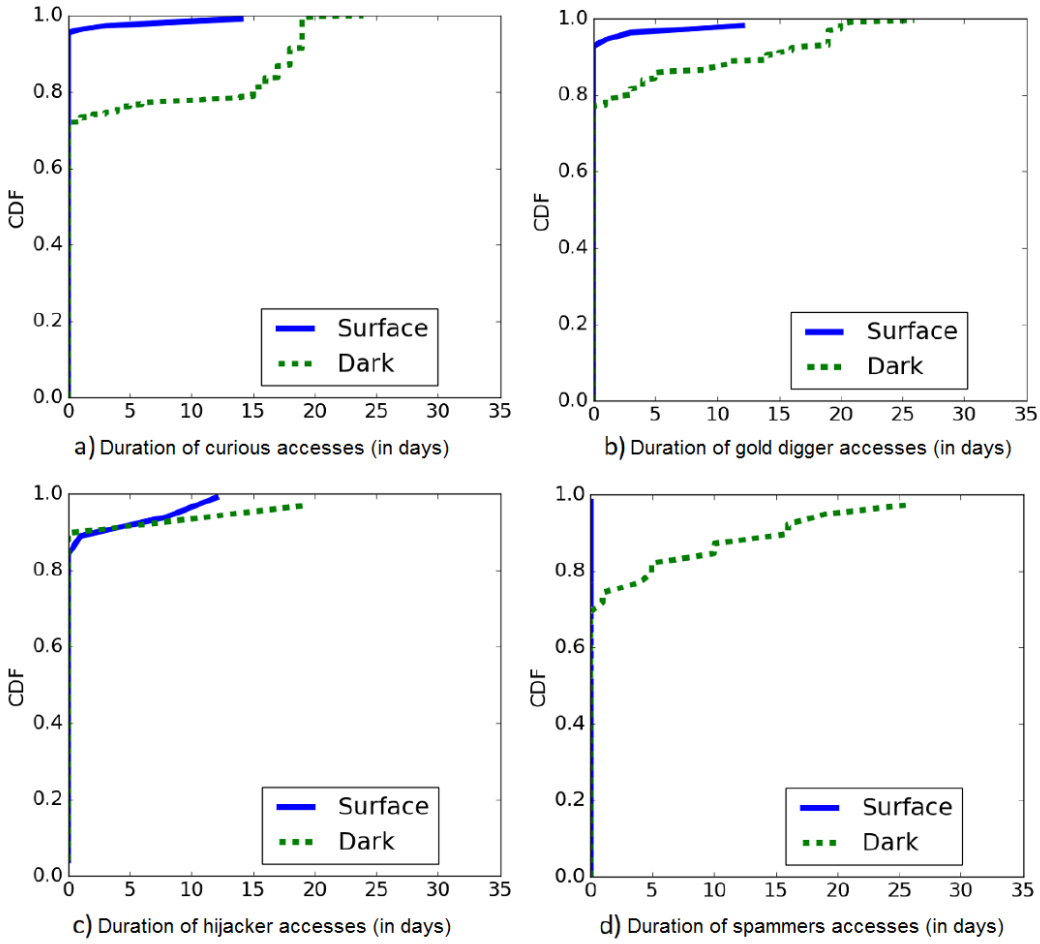


Figure 3.5: CDF of the length of unique accesses on the honey accounts for: a) Curious, b) Gold Diggers, c) Hijackers, and d) Spammers. The X-axis represents the duration of the access in days. Most accesses in all categories occurred only once.

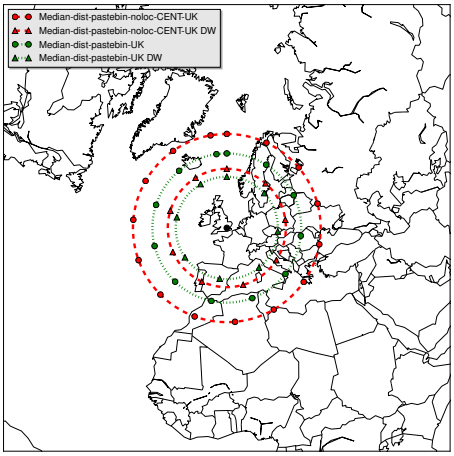
locations belong to 57 countries. Therefore, the percentage of logins coming from Tor represents 9.8% on the Surface Web and 34.6% on the Dark Web. A chi-square test ($\chi^2 = 40.929$, p<.001) revealed that these differences are significant, i.e., accesses to accounts leaked on the Dark Web are more likely to originate from exit nodes on Tor. The low percentage of Tor accesses in the Surface Web experiment may suggest that attackers acknowledge that Google uses a system to detect malicious activity that may lock the accounts if the login location is different from the usual one. Thus, they fake the login location to evade the protection systems. However, attackers on the Dark Web can set a specific exit node for the same purpose.

It has been hypothesised that criminals attempt to evade Google's anomaly detection system by trying to log in from a location that is meant to be the same as the account owner's. For the Surface Web experiment, some groups of accounts were leaked with credentials only, others included UK location information, and the remaining groups included US locations. A midpoint among locations provided in the leaks was defined for each country (London for the UK and Pontiac MI for the US). Then, the median values of the distances from the locations recorded in unique access to the midpoint are calculated for each outlet. The same approach was taken for the Dark Web experiment. As an example, Figure 3.6 shows these median values represented as circles on a UK map.

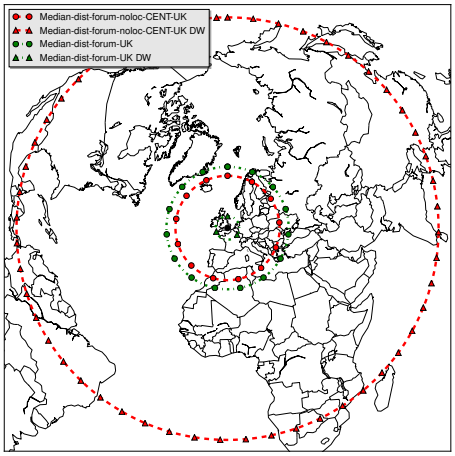
As shown in Figure 3.6a, for paste sites with advertised location information based in the UK, accesses to the accounts originate from places closer to the midpoint in the UK than accounts with leaked information containing usernames and passwords only. This phenomenon occurs both on the Surface Web and the Dark Web. Figure 3.6b represents the accesses from underground forums with location information based in the UK. The behaviour is the same for the Dark Web: leaked accounts are accessed from locations closer to the UK midpoint. However, in the Surface Web, the pattern is reversed with accesses without location information originating closer to the UK midpoint.

In the case of paste sites with US-based information (Figure 3.6c), no meaningful comparison can be made because we did not observe accesses to this category during the one month of the experiment. Finally, Figure 3.6d depicts the connections from underground forums when US information is posted. As it can be seen, the origin of the accesses with US location information is closer to the US midpoint in the Surface Web. On the other hand, accesses with location information are further from the midpoint in the Dark Web.

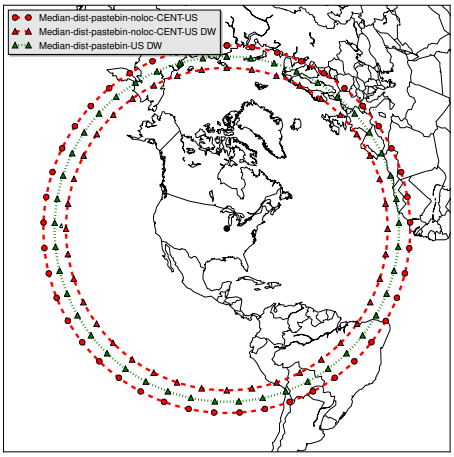
Figure 3.7 shows the Cumulative Distribution Function (CDF) of the distances from the login locations to the midpoints. As shown in Figure 3.7a, for paste sites with UK location information, accesses from the Dark Web are closest to the midpoints regardless of the information location included in the leak. Approximately



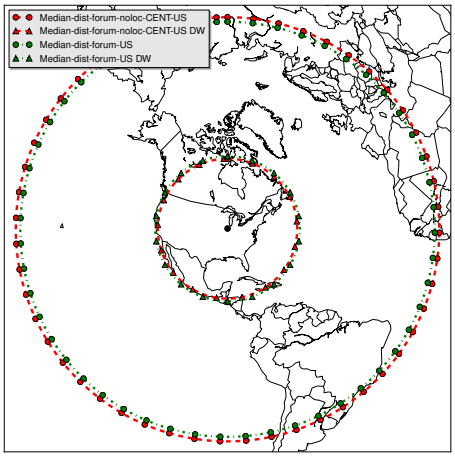
(a) Distance of login locations from the UK midpoint for paste sites.



(b) Distance of login locations from the UK midpoint for forums.



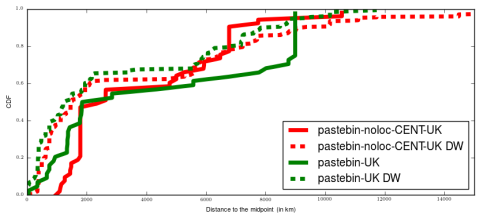
(c) Distance of login locations from the US midpoint for paste sites.



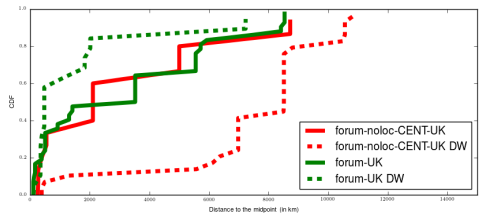
(d) Distance of login locations from the US midpoint for forums.

Figure 3.6: Distance of locations where the login originated from midpoints to locations advertised in the leak based in the UK. Red lines represent credentials leaked without the location information, and green lines represent credentials leaked with location information. (Circles represent Surface Web accesses and triangles represent Dark Web accesses). As can be seen, credentials leaked with location information are closer to the advertised midpoint but only in some outlets, and there is no pattern between the two Web environments.

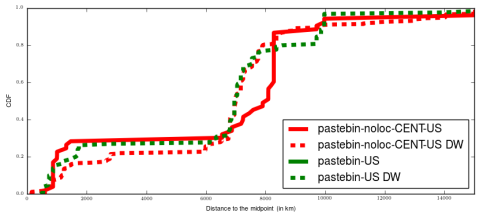
70% of all these accesses are the closest to the midpoints (2000 Km). Accesses from the Surface Web share the same trend, but they are slightly further away from the midpoints. For underground forums based in the UK (Figure 3.7b), 80% of logins from the Dark Web are close to the midpoints (2000 Km) when they have location information. Conversely, only 10% of accesses are within the same distance for accounts leaked with no location information. Regarding the Surface Web, 60% of the accesses with no location information are in the 2000 Km range, and 50% of the accesses with location information have the same distance. We cannot make location comparisons for accesses from paste sites based in the US as we did not get accesses in the first month of the experiment (3.7c). Finally, for underground forums with US information (3.7d), accesses from Dark Web are closer to the midpoint than Surface Web accesses. Approximately 50% (with location information) to 60% (no location information) of Dark Web accesses are in the 2000 Km range. In contrast, 60% of Surface accesses are further from the midpoint reaching about 8000 Km.



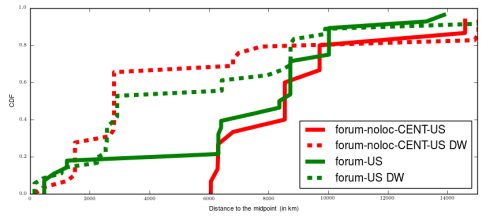
(a) Distance of login locations from the UK midpoint for paste sites.



(b) Distance of login locations from the UK midpoint for forums.



(c) Distance of login locations from the US midpoint for paste sites.



(d) Distance of login locations from the US midpoint for forums.

Figure 3.7: Cumulative Distribution Function (CDF) of the distances from the login locations to the midpoints. Red lines represent credentials leaked with no location information, and green lines represent credentials leaked with location information. (Lines represent Surface Web accesses and broken lines represent Dark Web accesses). Overall, credentials leaked on the Dark Web are closer to the advertised midpoint but only in some outlets, and there is no pattern between the two Web environments.

3.4.5 Interesting Case Studies

Interesting cases occurred during the Dark Web experiment that helped to understand the activity performed in compromised accounts. Many instances in which the honey accounts were used to create accounts on other online services such as online payment platforms (Paypal), cloud storage providers (MEGA, Dropbox), online stores (HP, Walmart), bitcoin exchangers (CEX.IO, Bitit) and more. Some orders were made in the online stores, but they were declined. Interestingly, the Google+ service was activated using one of the leaked accounts and friend requests were sent to several profiles. Some of them were accepted. Moreover, we observed spearphishing activity on some accounts. Figure 3.8 shows a handcrafted email sent on behalf of Dropbox that invited users to download a file that probably contained malware.

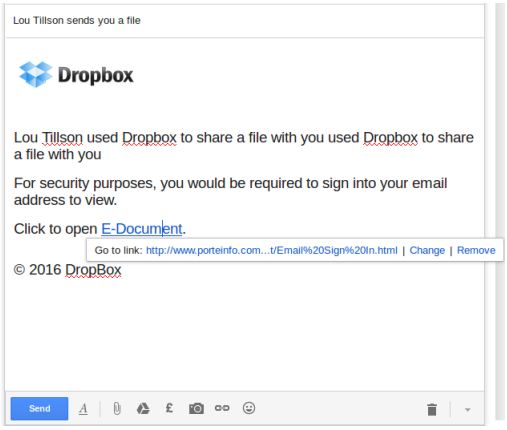


Figure 3.8: Email used to send spearphishing to the contacts within the account.

3.5 Discussion

Our findings show that accounts leaked through paste sites received more accesses in both Web environments, but the scale of access is much more extensive for paste sites on the Dark Web. While it is true that paste sites are more likely to be used to leak credentials, there is a big difference in the exposure of the leaks between the Surface Web and the Dark Web. Generally, in the Surface Web, content related to information leakage is removed from paste sites by administrators monitoring the site. On the contrary, paste sites are not monitored on the Dark Web, and leaks tend

to be published for longer. Therefore, credentials leaked in paste sites on the Dark Web are more exposed than on the Surface Web. Regarding underground forums, exposure is similar to paste sites on the Surface Web. On the contrary, credentials are less exposed in the Dark Web forums because they usually require creating an account and sometimes an invitation. One limitation of our work is that we were not able to establish whether the outlets used for our experiment are similar in terms of traffic. Therefore, the difference in the number of accesses between both Web environments may be due to the particular websites and hidden services we chose and not because of the environment itself.

In terms of the type of activity (taxonomy), there is a higher concentration of *Hijackers* on the Surface Web. *Hijacking* can be considered malicious, but its absence can mean that cybercriminals are more sophisticated and try to go unnoticed when using credentials. Thus, there is a higher level of malicious activity on the Surface Web, but miscreants tend to be more stealthy in the Dark Web. Interestingly, our data shows that there is a high concentration of *Curious* on the Dark Web. Even though no activity is performed on the honey accounts, it is reasonable to assume that more skilled attackers would not interact with the accounts to avoid detection. Unfortunately, we are not able to detect these ”sophisticated” *Curious* users. Furthermore, the high level of *Curious* activity on the Dark Web can be explained by sophisticated miscreants crawling websites searching for stolen data and using bots just to perform the login to build a credentials database for further inspection.

We showed that a variety of operating systems and browsers were used to access the honey accounts. Android is more likely to be used on the Surface Web, indicating a low level of sophistication as personal devices may be used to log into the accounts. On the other hand, the use of Linux is a sign that high-skilled attackers are accessing the Dark Web accounts. It may be the case that sophisticated attackers are using Windows bots to access the accounts, yet we are not able to measure automatic accesses with our infrastructure.

Regarding the type of browser used, accesses from unknown browsers are more

likely to happen on the Dark Web. This fact indicates that attackers try to hide their browser user agent information, suggesting some degree of sophistication. However, the use of browser extensions to change or hide the browser's user agent is common among users nowadays. Moreover, it could be argued that skilled users are prone to use known or typical user agents in an attempt to avoid being flagged as malicious users. On the Surface Web, Chrome is more likely to be used for logging in to the accounts. The use of this standard browser suggests a low level of sophistication in this environment. Our data was collected using Google's fingerprinting system; thus, the results' reliability depends on the system's accuracy. Nevertheless, the observed differences suggest that a considerable percentage of sophisticated users attempt to be stealthy on the Dark Web when credentials are obtained through paste sites. Also, the comparison shows that attackers on the Dark Web are more likely to connect several times to look for new information in the accounts.

Similarly, the evidence suggests that the duration of the accesses lasts longer on the Dark Web. The data shows that attackers on the Dark Web are more likely to connect several times to look for new information in the accounts. Furthermore, we found that accesses to accounts leaked on the Dark Web are more likely to originate from Tor node exits. There are no definite patterns in the location of the accesses for both environments. However, most accesses tend to be closer to midpoints when the accounts are leaked on the Dark Web.

The comparison shows us that although the differences in terms of the type of activity are not substantial in some cases, the Dark Web attracts individuals who seek to discover the secrets of the dark side of the Web. The high number of accesses through hidden services suggests a great interest in the information contained in the Dark Web outlets. It is reasonable to assume that this information could lead many users to use it maliciously and end up becoming cybercriminals.

We believe that security systems for account logins can be improved with the help of behavioural detection systems, which are capable of finding activity patterns that seem to be different to those commonly used in the accounts. Therefore, information about access to compromised accounts can help build algorithms that

allow early detection of malicious activity. We observed malicious activity on accounts leaked on the Dark Web suggesting an increasing use of this environment as a platform to perform illegal activities, especially as far as the trade of stolen information is concerned. For this reason, data gathered from this project may support the development of policies focused on disabling hidden outlets dedicated to those activities.

One of the important limitations of this comparison is that the Surface and the Dark Web experiments were performed in different spaces of time. Therefore, the activity level in both Web environments could have changed from one experiment to the other. Thus, the data of the experiments may not be enough to generalise our results. Our future agenda includes setting up honeypot infrastructure for both environments on other online services to establish a more accurate comparison. Another limitation was the number of Gmail accounts that we were able to create for our experiment. The creation of an account requires the registration of a phone number, and any automatic approach is flagged as spam by Gmail; therefore, we were not able to create a large number of them.

3.6 Conclusion

In this work, we compared the data from two similar experiments in which credentials of honey email accounts were leaked on the Surface Web and the Dark Web. We collected and performed a comparison based on different variables in our observations. Compromised accounts received more unauthorised accesses on the Dark Web than on the Surface Web, especially when credentials are released on paste sites due to the level of exposure of this type of outlet. We found a relationship between the Web environment and the type of activity performed in the honey accounts, as well as the configuration of the devices used to log in to the accounts. We believe that our findings can help the research community better understand the different types of malicious activity on stolen accounts. This comparison will contribute to the development of behavioural rules that can be included in detection systems aiming to protect users from attackers in different layers of the Internet.

Chapter 4

Malvertising

This chapter aims to present our work related to malvertising based on our paper published in the IEEE Workshop on Attackers and Cyber-Crime Operations 2020 [80]. We study malvertising as another type of cybercrime that can take place in the Surface Web and the Dark Web. This chapter explores user exposure to malvertising, considering the type of network they use to access the Internet.

4.1 Introduction

People's exposure to media and advertising has changed drastically since the advent of the Internet. As more online content is available, many companies seek to position their brands among millions of users worldwide through online advertising. According to the Interactive Advertising Bureau (IAB), which provides a regular update on the state of the digital advertising market, Internet advertising revenues in the United States surpassed 100 billion dollars in 2018, increasing 21.8% over the previous year [81].

Like any other type of advertising, Web-based advertising is a relationship between advertisers and publishers. While advertisers buy space on Web pages to show their ads, publishers get paid to display ads for others on their websites [82]. Due to the number of players involved, entities known as advertising networks (ad networks) manage the buyer/seller process between advertisers and publishers.

Cybercriminals are abusing the online advertising ecosystem to conduct fraudulent activities such as malicious online advertising [83], which is a more powerful

mechanism of infection compared to other dissemination strategies because of the implied trust that exists between the parties involved in the ad delivery process. Publishers and advertisers trust the ad networks to deliver only genuine ads. At the same time, users believe that the ads displayed on Web pages are legitimate, therefore they are more likely to interact with them. As a result, miscreants infect a larger number of victims in a short amount of time. Although ad networks spend significant resources to effectively mitigate malicious ads by implementing inspection and monitoring techniques, malicious advertising is ubiquitous [83, 84, 85, 86, 87].

At the same time, an increasing number of users concerned about their privacy are migrating to anonymity networks such as The Onion Router (Tor) [8] to access online content. The main feature of these networks is that they provide user privacy by obfuscating the traffic between a client and a website or online service; therefore, the user can access the hosted content anonymously [9]. Previous work showed that a growing number of websites are limiting or rejecting access to users using Tor [88]. This leads us to believe that Tor users may be treated differently by ad networks as well, relegating them to the role of second-class citizens on the Internet. As a result, low-quality ads from less popular ad networks might be delivered, putting Tor users at greater risk of being victims of malvertising. In addition, visiting websites from Tor might increase the risk of state-sponsored malvertising attacks on users with high sensitivity to surveillance or countries with strong censorship policies in an attempt to deanonymise those users [89].

There is a considerable amount of literature suggesting that attackers utilize the Tor network to commit illegal activities[54, 55]; however, it is not yet known whether such network increases victimization risks to certain attacks. In this chapter, we take the initial steps to explore user exposure to malvertising considering the type of network they use to access the Web. Therefore, the following research question is formulated:

- **RQ2:** Does the type of network that a user utilises to connect to the Internet influence the type of advertising and the level of malvertising that the user is exposed to?

To this end, we collected online ads by crawling 20,000 websites from 6 different IPs using a regular connection to the Web (*regular network*) and the *Tor network*. Then, we establish a comparison between both Web environments to understand malicious activity related to advertising. Similar to our work about compromised online accounts, the purpose of this chapter is to understand how malicious Web advertising is performed in different layers of the Web. In doing so, we hope to shed some light on how cybercriminals abuse the advertising ecosystem.

Our results show that ad networks deal with ad requests the same way regardless of the type of access and do not discriminate against users coming from Tor. Additionally, we observed that even though the level of maliciousness is similar in the regular network and the Tor network, ad servers when accessed from Tor are more likely to deliver ads from low ranked landing pages. This may suggest that ad networks are redirecting Tor users to less reputable ad networks associated with malicious websites.

4.2 Background and related work

Online advertising has rapidly evolved since it appeared in the 1990s as static images at the top of websites. Nowadays, advertisements (ads) can be created in multiple formats and targeted to different audiences who need a specific product or service. Online advertising is delivered using several channels such as display advertising, email advertising, search engine marketing, social media marketing and mobile advertising.

4.2.1 Online Advertising

Ad networks are centralised platforms that buy ad space provided by a group of publishers and sell it to advertisers. Internally, an auction process determines how ads are allocated among the publishers depending on the most profitable advertisers. To this end, the ad network manages ad traffic in the form of ad requests and ensures maximum revenue for all its members by displaying the most suitable ads on a website to the right visitors. Similarly, ad exchanges enable buying and selling ad traffic from multiple ad networks.

When the auction process is successful in the ad exchange, an ad is shown on the publisher's webpage, then the advertiser pays the ad network, and a percentage is paid to the publisher. Publishers are paid every time a visitor watches an advertisement on their site or every time a user clicks on the ad. In this case, the ad network sends a redirected URL to the browser, which is the landing page hosted by the advertiser. Moreover, when a user clicks the ad and performs further actions such as purchasing an item or filling a form, the publisher gets more revenue [83].

Previous studies have focused on measuring the online ad ecosystem to understand its features. Liu et al. [90] implemented a browser-based tool that provides detailed measurements of the prevalence of different ad targeting strategies. Likewise, Barford et al. [91] developed a Web crawler to collect display ads to determine whether the delivered ads depend on the user profile (cookies and browser profile). We have used some strategies presented in these papers to develop our tool to collect ads.

4.2.2 Malvertising

The complexity of the ad ecosystem allows cybercriminals to conduct malicious activities. Attackers operate in different ways to abuse the online advertising ecosystem. Click fraud, for instance, is an activity in which miscreants, who are members of the ad network, generate fraudulent clicks to make more money than they deserve. Several studies have shown that this malicious ad traffic is generated from botnets aimed to visit fraudulent websites created by the fraudsters [92, 85]. For instance, Stone-Gross et al. [82] study click fraud from the ad network perspective and describe the ad exchange process in detail.

In our work, we focus on malicious advertising known as malvertising. It aims to infect users with malware or redirect them to malicious websites under the control of criminals [84]. To this end, they inject ads into the ad network or set up a shady ad network to deliver malicious ads. Malicious ads take advantage of browser vulnerabilities to infect the victim's machine, lure users into downloading and installing malicious software or redirect users to websites they have not planned to visit [83].

Previous research suggests that malvertising operations are low-cost and more effective than other malware dissemination techniques such as spamming and social networking [93]. Li et al. [85] investigated the topology of malvertising and proposed MadTracer, a machine learning detection tool that identifies prominent features from malicious advertising nodes and their related content delivery paths. Furthermore, similar research involving crawling and monitoring has been found in some prior academic publications [94, 95]. Different from our work, they relied on static analysis.

Zarras et al. [83] studied malvertising using different open tools to collect ads and detect malicious behaviour. Using a similar approach, we developed an ad collection tool from scratch. Some detection techniques using statistical models have been proposed as well. For instance, Huang et al. [93] applied the Bayesian game model to inspect Web-based malvertising. We differ from these studies as we focus on measuring the malvertising ecosystem from the end-user perspective, considering the type of network used for the access.

4.2.3 Criminal Implications

Malvertising is a form of malware distribution that is usually designed to gain unauthorised access to ICT systems and steal information. It can be classified as a type of ”cyber-trespass” and ”cyber-theft”. Bossler et al. [96] argue that malware infection is analogous to a burglary in that it infects and compromises computer systems in the same way that thieves do when they break into a house. While burglars use concealed points of entry and weak points in the security to access the building, online attackers utilise shady ad networks to distribute the malicious software anonymously and reach more vulnerable victims.

An increasing number of studies have examined the relationship between RAT and malware infection. For instance, Holt et al. [97] showed that the risk of infection increases when engaging in deviant online activities such as viewing pornography or accessing and downloading copyrighted media. Furthermore, antimalware software proved to be a powerful protection factor for victimisation against unauthorised access [98].

4.2.4 Malvertising on the Dark Web

Anonymity networks serve an important purpose on the Internet, ensuring privacy for users accessing Web resources. Therefore, an increasing number of privacy-conscious users are using Tor to access online content. However, a rising number of websites (publishers or advertisers) or Web resources discriminate against Tor users, offering them a degraded service. In addition, visiting sites from Tor might increase the risk of state-sponsored malvertising attacks on users with high sensitivity to surveillance or countries with strong censorship policies in an attempt to deanonymise those users [89].

Previous work showed that a growing number of websites are limiting or rejecting access to users using Tor [88]. This leads us to believe that Tor users may be treated differently by ad networks, relegating them to the role of second-class citizens on the Internet. As a result, low-quality ads from less popular ad networks might be delivered, putting Tor users at greater risk of being victims of malvertising. Khattak et al. [88] demonstrate the existence of differential treatment of Tor users by analysing website responses to Tor requests. Our work aims to observe whether ad networks are relegating Tor user requests to less reputable ad networks delivering malicious ads.

4.3 Methodology

Our research focuses on the ad delivery process from the end user's perspective. This section presents the methodology we use to generate and analyse a large corpus of advertisements. To do this, we crawl websites using the regular network connection and the Tor network to extract the displayed advertisements served by the ad networks. We then measure and analyse similarities and differences in both access environments.

4.3.1 Data Collection

Collecting data about display ads is a task that requires the ability to identify and analyse the various elements of a webpage. One crucial step is to identify ad elements and record the redirections triggered by the ad request [90]. When a user visits a

publisher's website, third party tags embedded in the HTML code collect information from different ad networks to serve the ad. Typically, third party tags are enclosed within iframes and execute JavaScript at the time of page load to render the ad dynamically. It allows the inclusion of external objects such as images, videos, and other HTML documents.

When the third party tag is invoked, a request is made to an ad server that deals with the arbitration process. The ad request is redirected through the different ad networks to find the most suitable ad to be displayed on the publisher's website. Ad request redirections are typically implemented through JavaScript, and we are unable to record the HTTP chain of redirections taking place internally by the ad servers. However, most of the time, the code retrieved by the execution of the third-party tag contains the last chain of the HTTP redirections, consisting of a URL that includes the ad server serving the ad and the advertiser's landing page.

Our approach consists in extracting all the advertising URL chains retrieved by the third party tags when the ad is rendered in the browser. To this end, we developed a Web crawler based on Selenium (a browser automation framework used to test Web applications). The main feature of Selenium is using a WebDriver capable of driving a real Web browser natively as a user would. We use Mozilla Firefox for our experiments. This approach allows us to visit websites automatically and retrieve all the content, including HTML and JavaScript code rendered, especially for advertisements. Figure 4.1 shows the crawling infrastructure used for this study.

The top 20000 websites from the Alexa top 1M list are visited to observe patterns on the most popular Internet websites. Our crawler extracts the site's source code, including the code dynamically generated by JavaScript in the main HTML document and within the iFrames. Depending on the DOM structure of the page, ads are often embedded in nested iFrames spanning multiple levels [90]. Since most of the source code related to ads is rendered in the inner iFrames, it is important to extract the data from these elements. Therefore, the code within iFrames is recursively extracted up to six levels and stored on disk. Appendix A.1 shows the function used by the Web crawler for this purpose.

After storing the code generated for each website, we identify all the advertising URL chains using EasyList, a database of regular expressions used to detect ads [91]. Although all the URLs obtained are ad-related, we discard *Single URLs* (e.g. `http://domain.com`) as they do not represent ad-delivery paths. We focus our analysis on *Compound URLs* which are two or more single URLs linked by specific attributes such as a query string. For example: `http://domain.com/path/?query1=test&query2=http://domain2.com`. We test the single URLs of the compound URL against the EasyList again to infer ad membership. If positive, the URL belongs to an ad server (ad URLs). Otherwise, the URL refers to the landing page of the advertiser (landing URLs). Appendix A.2 show part of the code implemented to perform this operations. In our observations, we have noticed that most landing pages come after specific patterns in the URL chain such as `adurl=`, `redirecturl=`, etc.

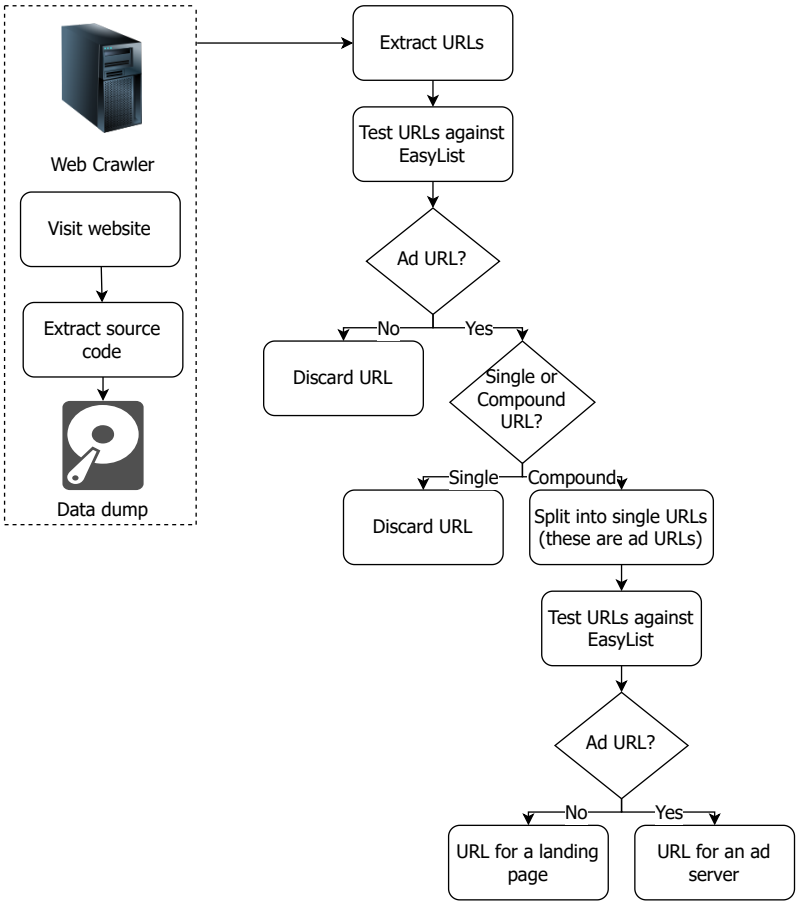


Figure 4.1: Overview of the online advertising crawling infrastructure.

It is important to note that not all the advertising URL chains included in the advertisements contain the landing page. Sometimes the URL chain consists only of URLs related to the ad servers. We still consider these URLs for our analysis as they represent the ad servers involved in the ad delivery process. As we observed manually in these cases, when the ad is clicked, further redirections are triggered within the ad network to open the landing page in a different tab or a new window in the browser. Since we focus our study on display ads, we only collect data from automatic redirections without clicking on the ads.

This work aims to measure malvertising from the user's perspective accessing the Internet using Tor and a regular network to compare the data we collect. Thus, we deploy our crawler in both access environments using three physical machines on three different IP addresses. Each physical device hosts two virtual machines (VMs). The purpose of the first VM is to crawl the Web using a regular connection and has direct access to the Internet using the IP of the physical machine. The second VM, intended to crawl the Web using Tor, is configured to use a transparent proxy that establishes a connection to the Internet using the Tor network. In total, we have three crawlers for the regular network and three for Tor, each visiting the top 20000 websites of the Alexa Top one million list.

To add consistency to our experiment, i.e. to ensure that the differences we find in our analysis are due to the access environment and not because of other factors, we consider the following controls:

- We visit the same set of websites.
- We synchronise our crawlers to visit each website simultaneously and in the same order for each of the 6 VMs.
- We use the same location: United Kingdom. Our physical machines are based in the UK for crawling the regular network. For Tor VMs, we configure them to use a single UK exit node.
- We remove websites not present in all 6 IPs to build a homogeneous dataset for our analysis.

4.3.2 Detection Systems

This chapter aims to detect malicious and fraudulent activities that exploit online ads and how they differ between users accessing websites using their regular network connection and the Tor network. Therefore, we analyse the components of the advertising URL chain (ad URLs and landing URLs) to detect malicious activities (e.g., delivering malicious content, illicitly redirecting users to malicious websites, etc.). If any component of the URL chain is identified as malicious, we flag the ad as malicious. Correspondingly, we call any chain containing a malicious component a malvertising chain. Note that not all the ad URLs or landing URLs of the chain are always malicious. For example, a malicious component may redirect a user to a legitimate website.

There are several online detection systems or blacklists to which URLs or files can be submitted to identify known malicious activity. We use antivirus scans provided by Virus Total and URL/domain blacklist services provided by Google Safe Browsing to classify ads as malicious. These online tools are free to use and provide APIs that allow us to automatise our classification process. We submit all the ad URLs and landing URLs collected by our crawler to our detection systems. We describe each detection system in the following paragraphs.

4.3.2.1 VirusTotal

VirusTotal is an online tool used to analyse files, hashes or URLs to detect malware. It inspects items with over 70 antivirus scanners, URL/domain blacklisting services and other tools to extract different traits from the submitted content [99]. VirusTotal aggregates data from various antivirus engines, website scanners, file and URL analysis tools, and user contributions. VirusTotal also includes several characterisation tools used for different purposes such as heuristic engines, known-bad signatures, metadata extraction, identification of malicious signals, etc.

We submit our URLs to VirusTotal using a Python script (see Appendix A.3) that uses the HTTP-based API provided by the company. We obtain a result showing whether a given antivirus solution detected a submitted URL as malicious for each URL. Not all the features included in the VirusTotal service are open to the public.

Hence, we established collaboration with VirusTotal to access all the tools provided, especially the metadata extraction, which extracts the category of the URLs submitted and is part of our analysis.

4.3.2.2 Google Safe Browsing

In a blacklist-based detection system, any domain/URL hosting malicious content is added to a blacklist for future reference [83]. Google Safe Browsing is a blacklist service provided by Google that allows client applications to check URLs against lists of Web resources related to malware and phishing and are constantly updated by Google. [100] This detection system aggregates information about maliciousness from various sources, including data crawled by Google's search engine robots and client-side checks.

From the client's point of view, Google Safe Browsing warns users about websites or links that may lead to potentially harmful Web resources. Safe Browsing has also provided an API that can be used to submit URLs using an automated script. Similarly to VirusTotal, we submit our URLs and receive as a response the type of threat related to the resource if the detection result is positive.

4.4 Results

This section analyses several aspects of malvertising in the regular network and the Tor network. We deployed our crawler on 20,000 websites from 3 different IPs for each type of access (60,000 in total), and we were able to extract data from 53,946 (90%) sites accessed from the regular network and 45,470 (76%) from Tor. When using the regular network, 9% of websites presented '404 Not Found' errors, 0.06% required a captcha, and 0.04% showed an 'Access Denied' message. From Tor, 22% resulted in '404 Not Found' errors, 1.7% required a captcha, and 0.3% showed an 'Access Denied' message. The difference in accessibility has to do with the fact many websites block Tor users [88]. After cleaning our dataset, 13,036 websites were available for our analysis for each IP representing 65% of the 20000 Alexa list. We aggregated the data from each type of access to perform the analysis.

4.4.1 Measurement Dataset

From the top 20000 Alexa websites crawled for each network, 10% more websites were accessible from the regular network than in the Tor network. We extracted 11,060 unique advertising URL chains from advertisements in the regular network and 10,041 unique chains from advertisements in Tor. As mentioned previously, a chain may contain two or more URLs, including ad URLs and a landing URL.

Table 4.1 shows the distribution of URLs found for both access networks. We include the domains to which those URLs belong, and we call them *ad domains* for ad URLs and *landing domains* for landing URLs. It is important to note that several URLs may belong to the same domain, but the behaviour for each one is different depending on the redirections performed. As can be seen, we obtained more advertising URL chains in the regular network than Tor. However, the number of URLs and especially domains are similar, which may be an indication that the same ad servers and advertisers participate in the arbitration process in both environments.

	Regular	**Tor**
Advertising chains	11060	10041
Ad URLs	2191	2016
Landing URLs	5639	5447
Ad domains	450	430
Landing domains	4877	4710

Table 4.1: Distribution of the advertising URL chains in the regular network and the Tor network. Chains may contain two or more URLs (ad URLs or landing URLs), and each URL belongs to a domain.

Table 4.2 reports the top ten ad servers for each network. As expected, Doubleclick from Google leads the list with about 35% of all the ad requests in both cases. While the top ten ad servers cover approximately 65% of all the ad servers, the remaining servers (440 in the regular network and 420 in the Tor network) represent about 35% of all the ad traffic generated in our experiments.

4.4.2 Maliciousness

To identify malicious ads shown to users visiting the websites from the regular network and the Tor network, we submitted our collected URLs using the VirusTotal

Regular		Tor	
Ad network	**Percentage**	**Ad network**	**Percentage**
doubleclick.net	35.60%	doubleclick.net	34.27%
pubmatic.com	6.77%	pubmatic.com	8.99%
adnxs.com	6.10%	adnxs.com	4.64%
mathtag.com	4.21%	tumblr.com	3.62%
openx.net	3.42%	openx.net	3.28%
tumblr.com	3.17%	mathtag.com	3.02%
smartadserver.com	2.09%	smartadserver.com	2.29%
casalemedia.com	1.79%	casalemedia.com	2.07%
weborama.fr	1.38%	de17a.com	1.76%
aolcdn.com	1.32%	weborama.fr	1.66%

Table 4.2: Top 10 ad servers for each Web network. Doubleclick is the most common ad server representing one-third of all them. 65% of all the ad traffic comes from the top 10 ad servers, and the remaining, which are more than 400 ad servers, account for about 35% of the total.

API and Google Safe Browsing API. If any of the detection systems flag any ad URL or landing URL as malicious, we assume it comes from a malvertising request. Therefore, we consider its ad server or landing page as malicious. It is common for VirusTotal vendors to disagree with each other [101] causing false positives. In order to minimise this risk, we sample some URLs to define a threshold empirically. Therefore, we label URLs as malicious only if at least three vendors flag them as positive. In addition, we understand that the results of engines may update over time, or they need some time to include new information about a URL that has not been previously scanned. Therefore, we performed subsequent scans for our URLs to obtain the most updated results.

As shown in Table 4.3, less than 0.5% of the ad URLs involved in the ad delivery process are malicious for each network. Likewise, about 0.3% of the landing URLs where the ads are pointing are infected. In terms of domains, malicious ad domains account for about 1% of all observed ad entities, and approximately 0.3% of landing domains are flagged as malicious. In general, the level of maliciousness is similar when accessed from the regular network compared to the Tor network. It is important to note that only one ad URL and only one landing URL are detected as malicious for regular access by Google Safe Browsing. For Tor, a single ad URL was flagged

as malicious.

	Regular	% of Total	Tor	% of Total
Ad URLs	10	0.46%	8	0.40%
Landing URLs	18	0.32%	19	0.35%
Ad domains	5	1.11%	5	1,16%
Landing domains	15	0.31%	17	0.36%

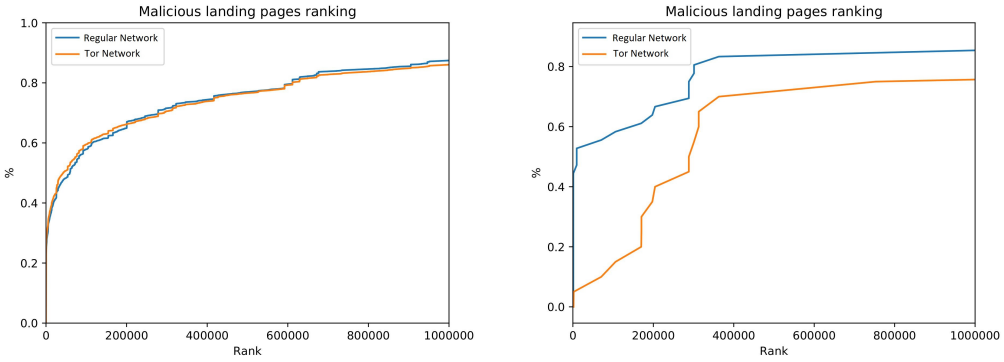
Table 4.3: Level of maliciousness in the regular network and the Tor network. In general, the level of maliciousness is similar in the regular network compared to the Tor.

We further analyse the landing pages resulting from the ad-delivery process to understand the nature of the advertisers involved in the ad ecosystem in the regular network and Tor. Firstly, we focus on the domain rank for each type of access to understand the level of popularity of the landing pages. Domain ranks are available in the Top 1M Alexa list. If a rank is above the one million, we extract it from the Alexa Web server directly. Usually, malicious pages are less popular because they are not indexed to popular search engines. The landing domains observed are in the range between 1 and 2,000,000. Approximately 7% of them were 'unranked' in both networks, i.e. Alexa did not have enough traffic data to assign a rank. Presumably, these domains were used for a short period as part of a malicious campaign or are very unpopular. Among the unranked domains, about 0.1% are malicious.

Figure 4.2a shows the CDF of the ranking of the landing domains observed for each type of access. It shows that the ranks of all the domains involved in the ad ecosystem are similar in the regular network and Tor. About 60% of the landing domains are in the top 100K list and 80% in the top 600K. The similarity in both networks may be an indication that the ad networks deliver ads from high-profile advertisers and do not discriminate against traffic coming from the Tor network.

In terms of maliciousness, as shown in Figure 4.2b, the rank of malicious landing domains is lower in Tor than in the regular network. About 5% of the infected domains in Tor are in the top 10K compared to 50% in the regular network. In other words, most malicious landing pages delivered from ads accessed from Tor are low ranked websites. We used a Kolmogorov-Smirnov test [102] to determine whether a significant difference exists between malicious domain ranking distributions. The

results showed a significant difference ($p<.001$) between the two types of access. This suggests that ad traffic from Tor is more likely to be redirected to shady networks which deal with less popular landing pages of dubious reputation.



(a) General ranking **(b)** Malicious ranking

Figure 4.2: CDFs of the ranking of landing domains. For all domains found, more than half (60%) of the landing domains are in the top 100K list and 80% in the top 600K. For malicious domains, the ranking in the regular network is higher than the ranking in the Tor network.

Furthermore, we identified the unique ad domains for each network to verify if certain ad networks only deliver ad traffic in one network or another. We found 35 unique ad domains in the regular network compared to 15 in the Tor network. Each domain accounts for 0.01% or less of all our observations, except one in the regular network representing 0.06%. Of these unique ad domains in the regular network, only three are malicious. No malicious unique ad domains were found in Tor.

4.4.3 Domain Categorisation

We analysed the categories of the landing domains taken from our advertisements to understand the type of websites targeted in the ad-delivery process. We used the VirusTotal API to extract the categories for each landing domain. We obtained 527 unique categories in the regular network and 511 for the Tor network. Table 4.4 shows the top 10 categories of landing pages found in our ads. The category of the landing pages is similar for both networks suggesting that ad networks do not target websites of different categories if the ad traffic originates from Tor.

Figure 4.3 illustrates the abovementioned results by showing the top 20 cate-

Regular		Tor	
Category	**%**	**Category**	**%**
business	21.99%	business	21.78%
shopping	7.80%	shopping	7.99%
uncategorized	7.18%	uncategorized	7.21%
information technology	6.18%	information technology	5.92%
news and media	4.65%	news and media	5.49%
education	4.41%	financial services	4.43%
financial services	4.39%	education	4.14%
travel	3.43%	travel	3.19%
marketing	2.89%	marketing	3.04%
advertisements	2.07%	advertisements	1.99%

Table 4.4: Top 10 categories of the landing pages observed. The landing page categories are similar in the regular and Tor networks, suggesting that ad networks do not target websites of different categories if the ad traffic originates from Tor.

gories in our set of advertisements for each access network. The trend is similar in the two networks. Most websites are related to business activities. The list also includes uncategorised domains. These are new or expired domains (as we checked manually), not associated with malicious activities. We found only three infected uncategorised domains. It is important to note that the 'advertisements' category refers to websites offering advertising services but no ad networks or ad servers.

Finally, we determine which categories are unique for each access network to understand if ad networks are more likely to serve ads pointing to landing pages with particular contents in one network or another. There are 46 unique categories for ads collected in the regular network and 50 categories that only appear in the Tor network. After a closer inspection, we found that these categories represent less than 1% of all the categories. In terms of maliciousness, some domains are categorised as dangerous by the categorisation engines. For each Web environment, we retrieved three unique domains with a malicious category.

4.5 Discussion

Our study shows that the ad-delivery process is similar when users visit websites from the regular and Tor networks in terms of ad traffic. There are minor differences in the number of advertising URL chains, ad servers and landing pages in both

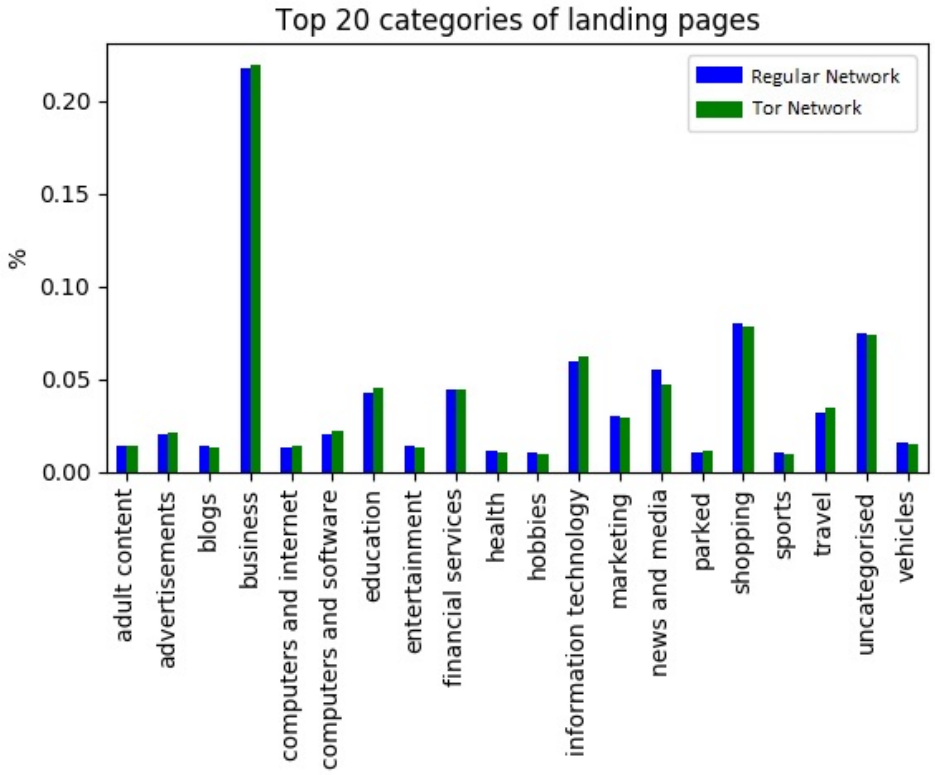


Figure 4.3: Top 20 categories of the landing pages.

networks, but these are insignificant. This may suggest that the volume of traffic related to the ad arbitration process in the regular network resembles the traffic observed in Tor. While it is known that Web traffic coming from Tor is blocked by some publishers [88], ad traffic flows in both networks without restrictions, with the same ad networks and landing pages participating in the process most of the time.

Doubleclick serves one-third of the advertisements in both networks, and the top 10 ad networks serve two-thirds of them. The list is similar for the two networks, with only one different ad server on each side. Thus, the same ad networks and advertisers are present across the different publishers regardless of the network. This supports our previous evidence that ad traffic is homogeneous in both networks and originates from the same ad networks. Considering that our advertising URL chains result from all the redirections after the ad exchange process, we may argue that ad requests coming from Tor clients are treated in the same way as clients requested from standard Internet connections.

The level of malicious activity is similar when browsing websites from the

regular network and the Tor network, as the percentage of infected domains is similar for ad servers and landing pages. None of the top 10 ad servers, representing two-thirds of all the ad networks, are flagged as malicious. We find that approximately 0.4% of the ad domains and 0.3% of the landing domain are malicious in both access networks. This suggests that cybercriminals perform malicious advertising activities at the same level on the Surface Web and the Dark Web, presumably because their target criteria are not based on the Web environment but on other factors.

Regarding the domains of landing pages, the general ranking of the landing pages served through the advertisements in the regular network corresponds to the general ranking in the Tor network. Most landing pages are in the top 100K. Therefore, advertisers involved in the ad-delivery process have a high level of popularity. Their ads are served based on the arbitration process performed by the ad network without discrimination traffic from the Tor network. However, focusing our analysis on the ranking based on malicious landing pages, there is a significant difference between both networks. The ranking for malicious landing pages in the Tor network is lower than the ranking on the regular network. This may suggest that ad traffic from Tor users is redirected to less reputable ad networks associated with lower-ranking advertisers, and based on our previous evidence, these advertisers likely host malicious content.

The categorisation of landing domains is similar between the regular and the Tor network. The categories of the landing websites are the same in both networks, with a slight variation in the ranking. Furthermore, unique categories for landing pages found in each network are an insignificant portion of all the observations. This suggests that ad networks are not serving ads based on landing page categories for each network. The same pattern holds for unique ad servers suggesting that ad networks participate in the ad-delivery process regardless of the network.

The comparison shows that there are no substantial differences in the ad-delivery infrastructure between the regular network and Tor, with almost the same ad networks participating in the process with few exceptions. Similarly, the categorisation of landing pages follows the same trend in both Web environments. Although ma-

licious activity is similar in both environments, malicious landing pages belong to less popular domains probably associated with shady ad networks. Data collected from our observations may be helpful for ad network administrators to understand malicious advertising activity in different layers of the Web. Therefore, if malicious campaigns are set in any network, they can be differentiated, and the respective countermeasures applied.

There are some important limitations in this work. Although we send synchronised requests to load websites, our crawlers cannot finish loading the content of the websites simultaneously, especially for Tor connections. The high latency produced by Tor while bouncing the traffic through several nodes causes a delay between the two networks. Therefore, our results may be influenced by the time difference in each crawling, as advertisements may change in time. Furthermore, we have sampled 20000 popular websites, but malvertising might be more prevalent in low-ranked websites. Therefore, we plan to crawl different subsets from the top 1M Alexa list to obtain more representative results as part of future work. Then, we will crawl the same subsets multiple times and average the results to remove some biases such as the latency. At the same time, multiple crawls will allow us to build user profiles and explore whether malvertising campaigns target users due to their browsing habits depending on the type of access.

4.6 Conclusion

In this chapter, we have explored the advertising ecosystem in the regular and Tor networks to understand the ad-delivery infrastructure and specifically compare malicious activity on display advertisements. We crawled websites in both networks to extract advertisements and studied various aspects related to malvertising. We found no significant differences in the ad-delivery process and that the level of maliciousness is similar in the regular and the Tor network. However, some ad networks deliver malicious advertisements related to low ranked landing pages. We believe that this work can help develop behavioural security systems aimed to detect and prevent malicious advertising campaigns in different layers of the Internet.

Chapter 5

Underground Forums

In this chapter, we analyse criminal activity in underground forums based on our paper presented at the APWG Symposium on Electronic Crime Research 2021 [103]. We explore trading patterns of several cybercrime-related forums hosted on the Surface Web and the Dark Web to understand the role of underground forums in the cybercrime ecosystem.

5.1 Introduction

Underground forums are platforms that facilitate communication among individuals engaging in illegal activities. These forums are used to exchange knowledge and ideas about illicit activities and trade items and services from an illegitimate origin or purpose. Similar to traditional Web forums, underground users establish relationships to collaborate and interact with potential trading partners. As a result, underground forums promote innovation in the cybercrime ecosystem, and individuals with different levels of expertise are attracted to these communities to obtain resources to perform cyber attacks. According to security firm McAfee, underground cybercrime profits in China have likely already exceeded US$15.1 billion [104]. This deviant behaviour is encouraged by an underground economy where the tools and communication vehicles for miscreants are becoming inexpensive, readily available, and online attacks are monetised.

The underground economy continues to expand through a large number of underground forums dealing with various types of online criminal activities, including

fraud, abusive monetisation techniques, money laundering, malware distribution, trading illegal physical goods and more [105]. Users advertise products and services in threads across forums, turning them into marketplaces where vendors and buyers trade assets for a price, using different digital currencies and payment providers. Potential buyers either reply to threads in the forum, use a private message service, or use an escrow service to complete the transaction [106]. These forum marketplaces offer many items and services such as stolen personal information, hacking services, malicious software and tools, bullet-proof hosting, currency exchange and even drugs and weapons.

Due to their essential role in the cybercrime ecosystem, significant research has been conducted on the Surface Web to understand better the structure and interactions on these forums [106, 107, 108, 109]. Underground forums on the Surface Web are indexed in the search engines and are accessible with any browser. Security practitioners and researchers constantly monitor forums to detect significant data breaches, zero-day exploits and vulnerabilities affecting information systems. Forum marketplaces, for instance, provide valuable information about the state of the underground economy related to emerging threats and attacks that are being traded as goods or services. Analysing these monetisation structures at a large scale is essential to understand the motivations and interests of forum members, which in turn provides insights into the pathways to crime.

Furthermore, cybercriminals are becoming more sophisticated and continue to improve their strategies to engage in underground communities without getting caught [51]. For this reason, miscreants are migrating to anonymous underground forums within the Dark Web to conduct illegal transactions [54]. This suggests that criminals are using hidden outlets on the Dark Web to connect to other individuals and trade illicit items or services [11].

Although a great deal of research has focused on criminal activity in underground forums of the Dark Web from different perspectives [11, 54, 110, 111, 112], very few studies have compared how the underground economy unfolds in both environments: the Surface Web and the Dark Web [51, 80, 113]. As such, this study

seeks to address this gap in the literature by conducting a quantitative exploratory analysis of the trading patterns of several cybercrime-related forums hosted on the Surface Web and the Dark Web and then compare the results to shed light on many elements related to commercial activities in both environments. Our goal is to broadly characterise trading activity in each environment and unveil patterns in products and services offered, prices, payment methods and currencies accepted. This knowledge can be useful to understand the difference in the modus operandi of cybercriminals operating in underground forums. This chapter addresses the following research question:

- **RQ3:** Does the Web environment where underground forums are hosted influence the type of transactions, products, and prices traded in those forums?

To this end, we crawled several popular underground forums in both Web environments, all dealing with the same type of criminal activities. Our forums deal with topics related to hacking, black-hat activities, financial fraud and stolen data. We extracted all threads (related to trading or not) posted during a year in the same range of dates for each forum. Then, we applied a tool based on natural language processing and machine learning to automatically identify the type of post, i.e. whether items are being sold (offered), bought (requested), or any currency exchanged. In addition, we identify products and prices. Finally, we aggregated our data by environment and analysed forum trading activity in both Web environments. In summary, this paper makes the following contributions:

- We studied commercial activity related to trading goods and services in underground forums from the Surface Web and the Dark Web.

- We compared the results of this analysis to understand the underground economy behind forum communities in different portions of the Web.

- We provide a comprehensive overview of underground forums by identifying the type of potential transactions and several product types with different prices.

- Our work contributes to the literature the knowledge to understand the modus operadi of criminals in underground forums, which is important to detect and stop illegal activities.

5.2 Background and related work

The evolution of technology and the Internet has enabled a broad spectrum of criminal offences that generate revenue for actors conducting these activities. Therefore, over the last decade, criminological research into cybercrime has expanded, with a particular focus on crimes related to economic motivation [114].

5.2.1 Criminal Implications

Underground forums open up cybercrime opportunities for potential offenders motivated to obtain easy money. According to the criminological theory of social learning, individuals gain insights into criminal activities while interacting with others as a natural behaviour in closed communities [40]. Consequently, they acquire the required skills and knowledge to engage in deviant behaviour. A great deal of information and techniques used by miscreants has led to transforming traditional forums into underground communities where cybercrime grows exponentially.

5.2.2 The Underground Economy

Criminality in underground communities is considerably promoted by a thriving underground economy where information about the availability of goods and services is shared by individuals seeking to monetise their illegal activities [115]. There are underground forums for a wide spectrum of illicit activities. People who do not have the necessary skills but are motivated seek the tools to engage in deviant behaviour on these forums.

Underground forums make available countless products and services. These illegal items are posted in threads by sellers who advertise what they have to offer, including price, payment method, contact information and specific rules regarding the transaction process. In turn, potential buyers contact the seller to ask questions and discuss the terms of the sale. After an agreement is reached, buyers make the

payment, and the goods are delivered [116]. Correspondingly, potential buyers make posts to advertise any product or service they are requesting. The myriad products and services offered include stolen accounts credentials, credit card data, malicious software, botnets, hacking and cashout services, currency exchange, etc. In terms of payments, several payment methods, currencies, cryptocurrencies and money transfer systems are accepted, or an escrow system is used. The success of these transactions relies on trust and informal social control through forum moderators and reputation indicators [117].

5.2.3 Forum Interactions

Since forums are a form of an online social network, prior research has analysed underground forums to understand interactions between criminals, the type of assets being traded and reputational factors [107, 118]. Pastrana et al. [115] focused on understanding criminal pathways and characterising key actors related to illegal activities using a social network approach. McAlaney et al. [119] studied discussions within online forums to better understand how individuals may be influenced in hacking behaviours and beliefs. As for analysing the purpose of posts, Caines et al. [120] examined the function and intent of posts from a corpus of several underground forums. Other authors studied private interactions between forum members by identifying whether a thread is likely to generate private messages aimed to complete transactions [109, 121]. Our work also considers the intrinsic characteristics of underground forums; however, we specifically focus on the aspects related to the underground economy, such as types of potential transactions, products and prices.

5.2.4 Forum Content

Other researchers have investigated specific types of content on underground forums. For instance, Samtani et al. [108] identified the characteristics and functions of hacker assets (tools) used in cyberattacks and obtained in underground forums. Similarly, Fang et al. [122] studied threads related to data breaches. Haslebacher et al. [106] focused on carding forums trading stolen financial data and analysed

products, prices, seller prolificacy, seller specialisation, and seller reputation. We differ from these studies as we focus on measuring posts related to the type of potential transactions, a variety of products and prices but not only from forums associated with a specific topic (such as carding) but with a broader spectrum in terms of the topics and activities performed on them.

Analysing data from underground forums is a demanding task that involves reviewing a large number of posts related to different domains and specific jargon. Portnoff et al. [2] developed a series of tools to extract high-level information from forum unstructured data using natural language processing and machine learning to overcome this challenge. We leverage these tools to extract data from underground forums at a large scale to answer our research question.

5.2.5 Underground Forums in the Dark Web

Cybercriminals are becoming stealthier and improving their strategies to avoid being monitored and detected in underground forums. They are joining underground communities on the Dark Web to share crime-related content and conduct illegal transactions for this particular purpose [54].

A growing body of research has examined underground forums on the Dark Web. Similar to the Surface Web, several studies focused on social network interactions. Nunes et al. [112] analysed forum discussions on the Dark Web to identify vulnerable platforms, vendors and products (e.g. hardware or software) that are at risk of exploiting by hackers. Likewise, Pete et al. [111] explored post discussions from six Dark Web forums to understand networks structures and structural patterns within these communities.

Other papers studied the characteristics of hidden services acting as specialised marketplaces more focused on trading items than on engaging in thread discussions [9]. Dolliver et al. [54] conducted a quantitative analysis of drug vendor characteristics on two marketplaces to determine differences among them. Hardy et al. [123] investigated the effect of seller reputation on the prices of goods and services in the Silk Road marketplace. Our work examines trading activity on underground forums that might include a marketplace section, not pure marketplaces.

Although there is a vast amount of forums located in both Web environments, their dynamics, interactions and the goods or services they trade might change depending on the Web environment. Since underground forums drive a large underground economy, it is important to understand the differences between individuals trading in these outlets. As there is not enough well-founded information to date, this research is intended to provide insights by measuring trading activity on underground forums on the Surface Web and the Dark web and compare the results to understand criminal activity in different layers of the Web.

5.3 Methodology

In this section, we present the methodology we used for our data collection. First, we developed a Web crawler to extract posts spanning a one-year period from a selected group of underground forums on the Surface Web and the Dark Web. Then, we instrumented our crawler with a modified version of the automated analysis tools proposed by Portnoff et al. [2] to retrieve the information we used for our comparison.

5.3.1 Forum Selection Criteria

To add consistency to our comparison, we considered selection criteria based on two conditions. Our forums should deal with similar topics and activities, and the time coverage should be comparable, i.e. data should be available for the same range of dates. Therefore, we selected two sets of forums (one for the Surface Web and another for the Dark Web) matching our criteria. In the following subsections, we describe the application of our criteria for the selection of our forums.

5.3.2 Forum Search

We searched for popular and prominent underground forums on the Surface Web and the Dark Web to build our forum dataset. We used the CrimeBB dataset provided by the Cambridge Cybercrime Centre (CCC) for a comprehensive list of forums. CrimeBB is a collection of posts from Surface and Dark Web forums scraped using the CrimeBot tool developed for researchers from CCC [124]. At the time of writing, CrimeBB hosted data from 18 Surface Web forums and 5 Dark Web forums.

However, not all of them met the selection criteria because the time coverage was not comparable. Thus, we searched for additional forums on Google, onion directories and Tor search engines based on our criteria.

Our search resulted in an extensive number of forums for both Web environments. Most of them had been shut down, especially hidden services. We first considered forums of the same nature to match our first selection criteria from those still active. Then, we chose forums where the same range of dates was available for extraction. Our selection included forums found in the CrimeBB dataset. However, we extracted a more recent version of the data based on our criteria.

As a result, we selected eight forums (Table 5.1): four from the Surface Web and four from the Dark Web covering activity over a period of 12 months, specifically from January 2020 to January 2021. Our forums focus on hacking, black-hat activities, malware distribution, financial fraud, stolen data and illegal monetisation techniques.

Hack Forums. According to the Alexa traffic ranking, Hack Forums is the number one website in the ”Hacking” category. It covers a wide variety of topics relating to hacking (tools and training), exploits, malware, stolen data and black-hat activities in general.

RaidForums. RaidForums focuses on ”raiding”, which is the practice of engaging in online collective activities such as DDoSiNg, doxxing, spamming or trolling. It is an exclusive database sharing and marketplace forum where members trade stolen information. RaidForums is a trendy forum hosting discussions on topics including raiding, hacking, leaks and tutorials.

Cracked.to. Cracked.to focuses on cracking tools and techniques aimed to gain unauthorised access to information systems. It includes database and stolen data trading, illegal monetisation techniques, hacking, etc.

Nulled. Nulled is another cracking community focusing on a wide variety of topics, including stolen personal information, databases, social engineering, malware, etc.

CryptBB. CryptBB is a cybercriminal forum on the Dark Web that focuses on hacking, carding and fraud. Initially, it was exclusive for experienced hackers who

needed to pass an application process. In 2019 a public section was open, resulting in an increase in users on the site to roughly 10,000 as of August 2020.

DarkWeb Forums (DWF). DWF focuses on several black-hat activities such as hacking, cracking, carding, illegal monetisation techniques, databases and accounts.

Deutschland im Deep Web (DiDW). DiDW is a German-speaking forum covering topics related to hacking, financial fraud, weapons and drugs. This is the only forum in our dataset dealing with drug trafficking, but we filtered that data out. DiDW was seized and shut down in 2017 but went online again with different onion addresses. We consulted a German speaker to confirm the accuracy of our analysis.

HM Forum. HM Forum is hosted on the Dark Web and specialises in fraud, carding, stolen financial data and illicit monetisation techniques. It also covers topics about hacking, stolen accounts and leaks.

	Forum	Threads (Trading)	Users
Surface	Hack Forums	29,634 (18.69%)	4,971,816
	RaidRorums	8,251 (15.39%)	558,729
	Cracked.to	10,234 (11.13%)	2,128,049
	Nulled	6,805 (12.78%)	4,039,030
Dark	CryptBB	1,617 (23.87%)	≈10,000
	DarkWeb Forums	1,308 (27.68%)	8,892
	Deutschland im Deep Web	5,024 (28.26%)	10,468
	HM Forum	2,740 (24.29%)	11,578

Table 5.1: Number of threads extracted from each forum for a one-year period. *Trading* represents the proportion of posts identified as trading posts. *Users* account for the total number of members since the forum was created.

5.3.3 Post Extraction

Online forums consist of tree structures composed of boards, threads and posts. Boards divide forums into categories for relevant discussions. Under the boards, members create threads by writing an initial post in which other users contribute by posting replies. Our approach is to extract the initial post within a thread. We excluded information about the author because we are not interested in community members. Similarly, we did not include replies to the initial post. According to Portnoff et al. [2], relevant information about trading is found in the initial post of

a thread, and further replies do not improve the performance of the analysis tools. Some forums have dedicated marketplaces to trade goods and services; however, we observed that commerce threads are posted in different sections across the boards. Therefore, we extracted all the threads for the given range of dates.

To this end, we developed a Web crawler based on Selenium (a browser automation framework used to test Web applications). The main feature of Selenium is the use of a WebDriver which has the ability to drive a real Web browser natively as a user would. We designed our crawler following some of the strategies proposed by CrimeBoT [124] such as completeness, stealthiness, efficiency, and downloading only textual content. We visited the list of forums matching our selection criteria and scraped the posts according to our approach. Appendix A.4 shows the source code used to parse data from the first post of each thread in the boards included in the forums.

5.3.4 Trading Information Extraction

Based on the tools for automated forum analysis described in Portnoff et al. [2], we developed a modified version that integrates our crawler with the modules that extract the data required for our analysis. We extracted three elements from each post appearing in a forum. **Type of transaction** refers to the type of post, i.e. the nature of the trading; a user may initiate a post to *buy* or *sell* a product or service, *exchange currency* or a topic related to anything other than trading. It is important to note that we observe the type of posts, not actual transactions. It is not known if the transaction is successfully executed or not. We also extracted the **product** and **price** involved in the post. If a post is tagged as currency exchange, we extracted the currencies exchanged.

5.4 Results

This section describes our analysis of the trading activity in underground forums on the Surface Web and the Dark Web. We collected and aggregated data from 4 forums on each Web environment (8 in total) for a period of one year. We extracted 54,924 posts from forums on the Surface Web and 10,689 posts from forums on

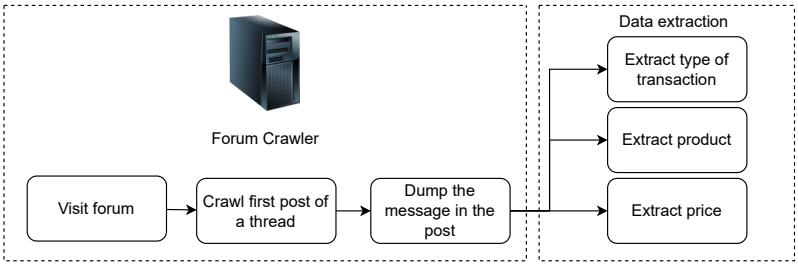


Figure 5.1: Overview of the infrastructure used to crawl underground forums. It included the tools proposed by Portnoff [2] to extract type of transactions, price and products from a post.

the Dark Web (see Table 5.2). We used a chi-square test (χ^2) [76] to determine differences in trading activity between both Web environments. Trading threads are more likely to take place on the Dark Web (26.67%) than on the Surface Web (16.05%) ($\chi^2 = 689.83$, p<.001). Moreover, our values are consistent with previous research [2] stating that less than 40% of forum posts are related to commerce.

	Trading	**%**	**Non-trading**	**%**	**Total**
Surface	8,818	16.05%	46,106	83.95%	54,924
Dark	2,851	26.67%	7,838	73.33%	10,689

Table 5.2: Number of trading and non-trading threads. Trading threads are more likely to take place in the Dark Web than in the Surface Web.

5.4.1 Type of Transaction

Table 5.3 shows the type of transaction posted in forums on the Surface Web and the Dark Web. A chi-square test revealed that the type of transaction significantly differs between both environments ($\chi^2 = 1050.01$, p<.001). Overall, selling posts (19.24%) are most likely to take place on the Dark Web, while posts related to buying or requesting products (6.85%) and exchanging currency (2.17%) are more likely to occur on the Surface Web. Even though, in general, all kinds of activity (trading and non-trading) are higher on the Surface Web, selling posts are more prevalent on the Dark Web. This indicates that hidden underground forums are suitable venues for miscreants to monetise their illicit activities.

	Sell	Buy	Curr	Other	Total
Surface	7.04%	6.85%	2.17%	83.94%	54,926
Dark	19.24%	5.84%	1.60%	73.32%	10,690

Table 5.3: Type of transactions. Selling posts are most likely to be found in Dark Web forums while buying and exchanging currency posts are more likely to occur on the Surface Web.

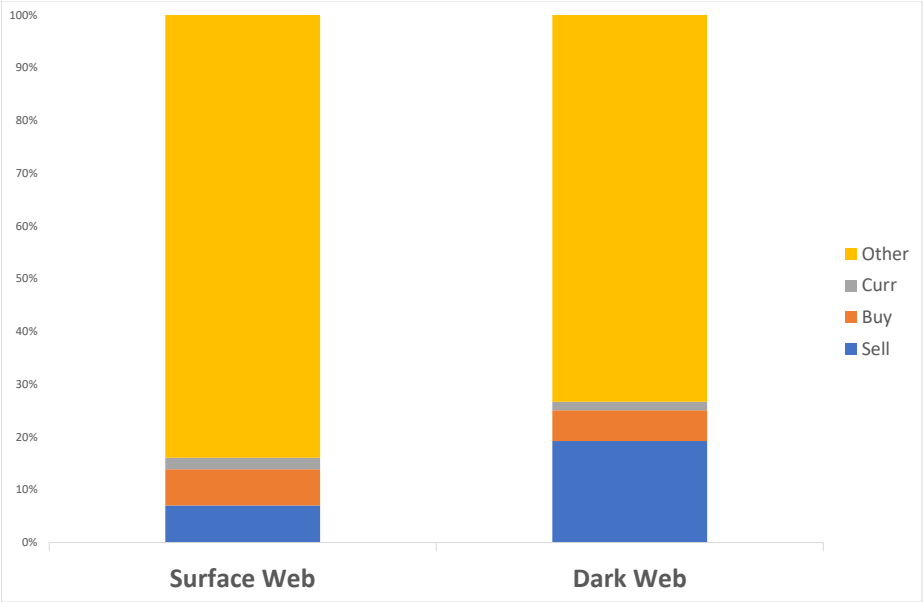


Figure 5.2: Type of transactions.

5.4.2 Products

We identified 3,865 products offered to sell and 3,762 requested to buy in Surface Web forums. For the Dark Web, 2,057 selling posts and 624 buying posts. As stated by Portnoff et al. [2], the product extraction tool only captures one product per post even if there are more. Similar to their work, we sampled 100 posts from the Surface Web and 100 from the Dark Web to confirm that multi-product posts rarely occur. In our samples, only one post offered more than one product for each environment. Also, where possible, we grouped posts selling/buying the same product with a different name. For example, we grouped posts about ”RAT” (Remote Access Trojans) and ”backdoor” because they are the same.

5.4.2.1 Selling Activity

Table 5.4 shows the top 10 selling products for each Web environment. On the Surface Web, most of the goods offered are associated with stolen personal information. This is in agreement with the findings of previous research [51] showing that stolen data exposure is higher in Surface Web forums than on the Dark Web. More specifically, online accounts (22.02%) and databases (17.71%) are the products that are for sale the most on the Surface Web. A wide selection of account types is offered, such as email accounts, media streaming services, social media, online stores, payment systems, and more. Databases or dumps offered can be email addresses used for spamming, email credentials from free providers or private companies, personal information and credentials from different websites (shops, dating sites, government systems, financial institutions), etc. Also, a high percentage (12.01%) of "make money fast" schemes are offered. These are paid packages including software, tutorials and resources aimed to conduct affiliate marketing activities. Although the activity itself is not illegal as claimed by vendors, the packages offered might not work as reported by some users. There is also activity related to malware (botnets: 9.12%) and DDoS services (7.24%).

Regarding the Dark Web, most of the selling activity revolves around malware. Approximately 50% of the posts offer malicious software, including RAT (23.52%), botnets (14.01%) and crypters (10.25%). RATs are popular because they give complete access to a victim's device, allowing an attacker to manipulate files, and turn on or off the camera or even the device. A crypter is a tool used to encrypt and obfuscate malware, making it undetectable to antivirus programs. Another well-known activity is money laundering, advertised as "transfers" (18.11%), where criminals sell money they obtained illegally by transferring it to the buyer's bank account or any payment system account such as PayPal. There is also selling activity related to financial stolen data such as credit cards credentials (6.17%) and fullz. Fullz are packets of information about individuals. They contain Personal Identifiable Information (PII) such as full name, phone number, address, social security number, date of birth, driver's license, which is used by criminals to commit fraud and identity

theft.

In essence, the Surface Web has a higher amount of activity surrounding stolen data. In contrast, the Dark Web has a more considerable amount of malware and financial fraud activity. Interestingly, we observed several posts offering products related to COVID-19 disease at high prices. We grouped these products under the same tag, and we found that 1.76% of all selling posts on the Dark Web are related to COVID-19 goods. These products include antibody tests, N95 masks and drugs that allegedly could be used to treat the disease such as Chloroquine, Lopinavir and Ritonavir.

Sell			
Surface		**Dark**	
accounts	22.02%	rat	23.52%
database	17.71%	transfers	18.11%
money	12.01%	botnet	14.01%
botnet	9.12%	crypter	10.25%
ddos	7.24%	credit cards	6.17%
verification	5.31%	ddos	4.47%
traffic	3.09%	fullz	2.33%
proxy	1.97%	accounts	2.02%
rdp	1.28%	COVID-19	1.76%
stock	1.08%	wallets	1.21%

Table 5.4: Top 10 most products for sale in forums for each Web environment. Surface Web forums have a higher amount of activity related to stolen data, whereas Dark Web activity is more focused on malware and financial fraud.

5.4.2.2 Buying Activity

Table 5.5 reports on the top 10 products required for buying for each Web environment. Similar to selling posts, buying activity on the Surface Web mainly pertains to stolen data, especially user accounts, representing around a third of all buying posts (31.61%). Correspondingly, combinations of users and passwords known as ”combos” (7.01%) and databases (3.72%) are popular data-related items requested in Surface Web forums. Furthermore, there is a high demand for DDoS services (8.83%) requested to damage specific companies. Posts purchasing hacking services (”hack”) are prevalent in the Surface Web (4.24%). These are requests to gain

unauthorised access to accounts, devices, websites or servers. Usually, a specific individual or company is the target, but no context is provided in the request.

Hacking services are the most requested services in Dark Web forums. Unauthorised access (18.13%) is requested to download files or execute commands on a website or server. Code for shells (14.03%) and RATs (12.09%) are popular buying items. These requests vary depending on the goal of the adversaries and target system. Likewise, /emphattacks (6.61%), where a hacking method or technique is requested, are among the most popular products to purchase. A high level of buying activity involves malware installs (7.15%) that are services aimed to deploy malicious software on devices or network infrastructures and set up Web hosts that are used to host malware, botnets, phishing sites, spam tools, etc.

While in Dark Web forums, buying activity is mainly based on hacking services, items related to stolen data are the most requested on the Surface Web. There are also hacking services requests on the Surface Web; however, these requests do not provide any information about the target. They usually focus on requesting access to a single personal account. Conversely, requests for hacking services on the Dark Web are more sophisticated and provide context about the system such as platforms, vendors, potential exploits, IP addresses, etc.

Buy			
Surface		**Dark**	
accounts	31.61%	access	18.13%
ddos	8.83%	shells	14.03%
combos	7.01%	rat	12.09%
hack	4.24%	cryptominer	9.54%
databases	3.72%	install	7.15%
e-whoring	1.96%	attack	6.61%
keylogger	1.64%	id	2.11%
review	1.39%	giftcard	1.47%
keys	1.12%	ddos	1.32%
giftcard	1.07%	server	1.24%

Table 5.5: Top 10 most products required for buying in forums for each Web environment. While buying activity is mainly based on stolen data in the Surface Web forums, hacking services are the Dark Web's most requested items.

5.4.3 Price

In our dataset, 48.72% of the selling posts in the Surface Web forums and 52.44% of the posts on the Dark Web mention pricing information. The rest of the posts require those interested to contact the vendor by private message. In some cases, there is an escrow or contract system available by the forum administrator to manage transactions. Posts with pricing information show the value of the product, usually in USD dollars and the payment method. Less than 1% of the posts do not include the payment method. On the contrary, buying posts never include pricing information.

Overall, prices are higher in Surface Web forums than on the Dark Web (see Table 5.6). Since our prices distributions are not normal, we performed a non-parametric Mann-Whitney-U test [125] to determine statistically significant differences between price products. For instance, PayPal accounts cost more on the Surface Web compared to the Dark Web ($Mann-Whitney-U = 5314$, $z = 11.86$, p<.001). On average, a PayPal account with a $4,500 balance costs approximately $220 on the Dark Web, whereas the Surface Web costs $260. Similarly, the average price for a single email account ($0.70 vs .$0.50) is higher on the Surface Web than on the Dark Web ($Mann-Whitney-U = 6804$, $z = 5.86$, p ¡ 0.01). User databases price is also higher ($15 vs. $12) on the Surface Web compared to the Dark Web ($Mann-Whitney-U = 4276$, $z = 8.15$, p<.001). We believe that stolen data (especially financial information) is more expensive in Surface Web forums because is ready to be monetised and has higher demand among non-skilled users looking to get a quick benefit such as money from a transfer or a free media streaming account.

Malware is generally cheaper in Dark Web forums compared to the Surface Web. For example, crypters and RATs are among the cheapest items that can be found on the Dark Web costing $1 each at the lowest. The average price for crypters is higher ($27 vs. $3) on the Surface Web compared to the Dark Web (Mann-Whitney-U=2042, z=10.53, p<.001). The same happens for RATs ($Mann-Whitney-U = 1974$, $z = 6.34$, p<.001), costing $39 and $5 on the Surface Web and the Dark Web respectively. As we observed in several posts, malware tools offered in the Surface Web are proprietary software developed by, according to the authors, companies

providing support and updates. Therefore, more expensive. On the other hand, malware on the Dark Web is sold by individuals who sometimes also include the source code in the deal. Then, the average price for DDoS services is higher on the Surface Web compared to the Dark Web. These services can be hired daily, weekly or monthly with an average price of $202 on the Surface Web and $175 on the Dark Web per day.

Some exceptions are prices for fullz and Web hosting. Fullz are significantly higher on the Dark Web than on the Surface Web ($Mann-Whitney-U = 3211$, $z = 13.02$, $p<.001$) with an average cost of $23 on the Dark Web and $17 on the Surface Web. According to the sellers, these are high-quality fresh fullz; therefore more expensive. Similarly, Web hosting is, on average, more costly on the Dark Web than on the Surface Web ($5 vs $12 per month), although the difference is not significant. Probably the reason for a higher price on the Dark Web is that this is a specialised bulletproof Web hosting aimed to host botnets, malware, spam and phishing sites, etc.

In terms of payment methods, the most used payment methods in Surface Web forums are Bitcoin, PayPal, Monero, Ehtereum and Perfect Money. Regarding the Dark Web, Bitcoin, Monero, Litecoin, Bitcoin Cash and Dash. This suggests that the underground economy on the Dark Web is based on cryptocurrency more than on the Surface Web.

5.4.4 Currency Exchange

Figure 5.3 shows the number of posts related to currency exchange in both Web environments. Rows represent the currency or payment system offered and the columns the one sought. Each cell reports on the number of posts for each offered/desired combination. We identified 23 payment mechanisms on the Surface Web and 14 on the Dark Web.

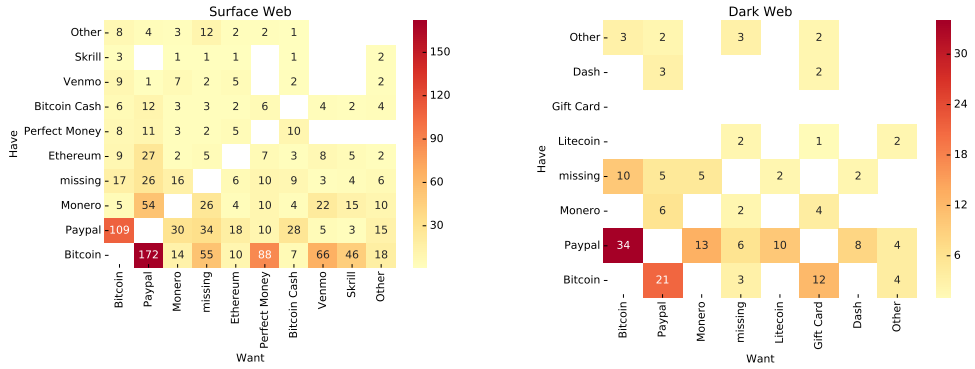
We extracted 1,193 exchange posts from the Surface Web (see Table 5.3a). In this environment, the most exchanged currencies and payment systems are Bitcoin, PayPal and Perfect Money. The most common operation is offering cryptocurrency in exchange for money in a payment system where exchangers profit by charging a

Product	Surface	Dark
PayPal accounts ($45,000 balance)	$260	$220
Bank accounts	$48	$22
Email accounts	$0.70	$0.50
Social media accounts	$90	$62
Social media followers	$7	$9
Netflix accounts	$3	$1
User databases	$15	$12
Fullz	$17	$23
DDoS (per day)	$202	$175
RDP Servers	$23	$17
Crypter	$27	$3
Remote Access Trojan (RAT)	$39	$5
Web hosting (per month)	$5	$12
Fake passport	$1850	$2100
ATM skimmers	$800	$800

Table 5.6: Prices for products. Overall, prices are higher in Surface Web forums than on the Dark Web.

fee for the transaction. Moreover, we believe that some criminals seek to cash out illegally obtained cryptocurrency for money in a payment system. In general, the most popular exchange is Bitcoin for PayPal. Our experiments confirm previous findings from Portnoff et al. [2] about this pattern.

Regarding Dark Web forums, we extracted 171 posts related to currency exchange, as shown in Table 5.3b. PayPal, Bitcoin and Monero are the most popular currencies on exchange. Contrary to the Surface Web, the most common operation is offering money in a payment system in exchange for cryptocurrency. We presume the reason for the demand is that miscreants seek to obtain cryptocurrencies anonymously, which otherwise must be obtained from official exchanges that request personal data; therefore, they can perform their illegal activities unidentified. Another reason might be cybercriminals selling money from stolen payment systems in exchange for cryptocurrency as a form of money laundering mechanism. Generally, the most popular exchange is PayPal for Bitcoin.



(a) Number of posts on the Surface Web. The most exchanged currencies and payment systems are Bitcoin, PayPal and Perfect Money.

(b) Number of posts on the Dark Web. PayPal, Bitcoin and Monero are the most popular currencies on exchange.

Figure 5.3: Number of currency exchange posts for each pair of currency or payment system (offered and desired) in each Web environment. *missing* stands for transactions not mentioned in the post or not extracted. "Other" represents the sum of other currencies or payment methods with less than two transactions.

5.5 Discussion

Our results show that there is a higher level of trading activity in underground forums within the Dark Web compared to the Surface Web. While posts related to selling goods and services are higher on the Dark Web, requesting products to buy and currency exchange posts are higher on the Surface Web. Most of the products for sale on the Surface Web are related to stolen data, whereas the Dark Web is focused on malware assets. Similarly, items related to stolen data are the most requested for buying on the Surface Web. On the other hand, the most requested products on the Dark Web are hacking services. In terms of prices, these are higher in Surface Web forums than on the Dark Web. Moreover, currency exchange posts in the Surface Web deal with a great variety of currencies, including government-issued currency and cryptocurrency. Conversely, cryptocurrency is the main mean of exchange on the Dark Web.

Overall, there is more activity in terms of threads and users on the Surface Web (commercial and non-commercial). Usually, Surface Web forums have a longer lifespan; therefore, more users create threads over time. In contrast, forums on the Dark Web are online for shorter periods because they are continually shut down by law enforcement agencies or taken over by competitors [126]. Furthermore,

Surface Web forums are more accessible for any user regardless of their technical skills. In contrast, visiting forums on the Dark Web requires some knowledge and specialised software to access anonymity networks and a level of skill to search these hidden services. Thus, we believe there is a higher concentration of skilled users in Dark Web forums. Selling posts are more prevalent on the Dark Web; this suggests that Dark Web forums are specialised marketplaces where skilled users offer products to a knowledgeable audience that already knows that the products they need are available and do not need to request them. Although posts related to buying products are higher on the Surface Web, we found a demand for hacking services in both environments. The main difference is that requests on the Dark Web are more sophisticated, with users dealing with specific types of tools, methods and techniques used to perform cyberattacks. In terms of currency exchange, since financial activity on the Dark Web is conducted via the tools of cryptocurrencies (especially Bitcoin) to maintain anonymity, there is more activity on the Surface Web as users deal with a variety of currencies and payment systems.

Our comparison shows that there is a higher level of sophistication among Dark Web users and the activities they perform in these underground forums. We believe that underground forums on the Dark Web are mainly run and catered by users with high computer-related knowledge who plan and perform criminal operations focused on making big profits. On the other hand, Surface Web underground forums cater for non-skilled users who are motivated by curiosity and looking for personal gain (like getting an account with money or a free media streaming account). While users in Surface Web forums are more likely to search, request and buy readily available products such as stolen accounts or databases for monetisation purposes, Dark Web users seek to monetise their skills by offering products or services for specific goals or targets. Therefore, the underground economy is increasingly growing facilitated by underground forums which are key drivers in the adoption of technologies with great potential such as encrypted messaging and cryptocurrency that challenge our notions of privacy and money. Moreover, although the COVID-19 has negatively affected economies worldwide, it seems to have opened up new opportunities for

online black markets.

We believe the results of our work are key to improving threat intelligence strategies regarding underground forums; therefore law enforcement agencies, security practitioners and the research community can make better decisions about how to develop procedures and policies to prevent, stop or deter illegal activities conducted in underground forums depending on the layer of the Web where shady commercial activity is conducted. Threat intelligence analysis is important because it provides information about the techniques, tools and assets of adversaries. By improving this analysis, researchers will be able to detect threats more efficiently depending on what portion of the internet they originate from. Trading activities related to malware, large data leaks, or zero-day exploits may be detected promptly; thus minimising the damage caused to the individuals or companies involved. For instance, security practitioners may trace malware variants to underground forums on the Dark Web, which is where these assets are mostly sold. Therefore, interventions could be performed to stop the authors of these variants from selling their malware (e.g. shutting down a forum). Similarly, using patterns found in this work, practitioners may improve their monitoring techniques for Surface Web underground forums. This may be helpful especially when prominent information leaks occur because, as we observed, these are the outlets to find stolen information.

Our work clearly has some limitations. First, some forums, especially on the Dark Web, require to pay a membership or an invitation to access premium areas; therefore our data may not represent all the content posted in those forums. One forum on the Surface Web and two on the Dark Web had areas with restricted access. Second, since our crawling was performed from a user's perspective, we did not include private messages that might provide information about finished or closed transactions in the forums. Third, we only analysed messages related to potential transactions, not actual trades. Fourth, the collection period is not enough to analyse trends and changes in trading activity, for example, those related to the COVID-19 pandemic that may have changed patterns in the underground economy. Lastly, some trading posts could be aimed to scam users. As a result, our analysis is not

based on actual trading data and might not present a complete picture of the trading activity in the forums. However, since we collected a great deal of data, we were able to perform a thoughtful analysis of trading patterns of offered/requested items in underground forums.

As part of future work, we plan to include more forums in each Web environment and collect data for a longer period of time to obtain more representative results. Our work focuses on trading posts such as types, products and prices. Similarly, other studies have examined only social networks interactions including user behaviour and private interactions. Therefore, future research should consider the analysis of all the variables of the forum ecosystem to understand the role of underground forums in the evolution of cybercrime. Researchers should also analyse scam threads to identify patterns and proportions of real trading in underground forums. These findings would be useful to improve threat intelligence activities aimed to counter illegal activities in these underground communities.

5.6 Conclusion

Underground forums have become increasingly central to exchanging information about malicious activities on the Internet, including performing illegal commercial activities. At the same time, since more cybercriminals are migrating to the Dark Web to hide their operations, it is vital to understand how illicit goods and services are exchanged in different layers of the Web. In this chapter, we collected data related to trading activities from forums on the Surface Web and the Dark Web covering a period of 12 months. Our findings suggest that the Dark Web has a higher degree of trading activity compared to the Surface Web. Most of this activity is related to products offered for sale and shows a higher level of sophistication. This work is an initial step in a larger research agenda to understand the modus operandi of cybercriminals in different portions of the Web.

Chapter 6

Discussion

This chapter will discuss different aspects of the stolen accounts study, malvertising study, and underground forums study. Although each study has its associated discussion sections, this chapter aims to show the shared characteristics between these contributions and relevant findings. Moreover, we explain the implications of this research in terms of the Information Security and Environmental Criminology fields. In addition, future research to expand on these findings will be suggested.

6.1 A Summary of the Problem

The Dark Web presents new opportunities for criminals to commit malicious activities as these platforms provide anonymity so they perceive the risk of being caught as very low. This thesis sets out to understand the modus operandi of cybercriminals on the Surface Web and the Dark Web by presenting measurement studies for three different types of cybercrime. Primary data related to different types of criminal activity was collected and analysed for comparison. Chapter 3 studies how criminals interacted with compromised accounts whose credentials were leaked in various outlets in both Web environments. Chapter 4 examines malvertising and how it is delivered to users accessing the Internet through a regular connection or Tor. Chapter 5 analyses underground forums to observe trading activity related to illegal goods and services between the Surface Web and the Dark Web. These three studies are important to understand cybercrime from different sections of the Internet.

6.2 Knowledge Gained

The purpose of this study is to determine if online criminal activity is different depending on the Web environment where it takes place. Our analysis suggests that there is a higher level of malicious activity on the Dark Web compared to the Surface Web for the three types of cybercrime we study. This indicates that more criminals are taking advantage of anonymous networks like Tor to avoid monitoring and detection, making them more difficult to catch by law enforcement agencies. Consequently, criminals are more sophisticated on the Dark Web, but not only in terms of stealthiness, because they conduct their activities in a more elaborate and knowledgeable way. Although not all the activity found can be defined as sophisticated in both environments, it is more prevalent on the Dark Web.

Several instances of this sophistication were found on the Dark Web that shares certain characteristics among the crimes analysed. For example, stolen accounts on the Dark Web are traded in underground forums where more skilled criminals have access. Therefore, access to these accounts is performed in a more stealthy way to avoid leaving a trace. Although several reports from security companies [127, 128, 129, 130] indicate that the Dark Web forums are the outlets where stolen accounts are traded the most, our data shows that this activity is more common on the Surface Web. However, we argue that stolen information in Dark Web forums is of higher quality and obtained from recent security breaches that occur less frequently.

The findings of this study indirectly contradict the work of Chavez [34] based on the Surface Web, who argues that hackers are not highly skilled individuals, but simply wily opportunists. Most of the time, miscreants looking to venture into the cybercriminal world seek help from skilled hackers to perform an attack, and these hackers are usually found on the Dark Web. This is what our findings showed. We discovered that criminals on the Dark Web pay close attention to their actions, which shows a certain level of expertise. For instance, they access a stolen account from a location meant to be the same as the account owner's if that information is provided in the leak. Or, they offer more sophisticated tools and services in underground forums on the Dark Web.

The above suggests that the Dark Web enhances criminal behaviour by hosting malicious activity from better-skilled criminals who deploy their operations on a Web environment that provides anonymity. Consequently, more criminals are willing to join this environment conducive to cybercrime, thus increasing the risk of victimisation for individuals and businesses in cyberspace. Although many of these individuals will first join in out of curiosity and without the required skills, after a period of knowledge sharing and learning from other criminals, they may improve their tactics. However, this does not mean that crime on the Surface Web will decrease, but rather the way these activities are carried out is different.

6.3 Criminological Analysis

The findings of this study show that there is a difference between malicious activities conducted on the Surface Web compared to the Dark Web. Since cybercriminals perceive a higher level of anonymity on the Dark Web, they conduct more sophisticated operations by executing a broader variety of illegal activities that include enhanced techniques and offering more sophisticated products or services. This is consistent with the literature about criminological theories as they see anonymity as an opportunity to offend in a platform where guardianship is more difficult to implement. Therefore, miscreants take the opportunity to run more complex operations. Although they may be less sophisticated, malicious activities are still present on the Surface Web. However, criminals make rational decisions about how to perform their operations based on the Web environment.

Although a great deal of research has been done to understand online crime using criminological theories, the difference in how these offences are performed by miscreants migrating to anonymity networks remains largely fragmentary. Criminology currently lacks data available to further explore the cybercriminal environment. Therefore, the findings of this thesis contribute to this area with the necessary knowledge to understand the opportunities and motivations behind online malicious activities, specifically for criminals moving to the Dark Web in search of anonymity. In this way, criminological theories can be translated into cyberspace. As a result,

the area of Information Security can implement more reliable preventive technical measures based on the outcomes of the criminological analysis.

6.4 The Role of Criminological Research

Understanding the state of this innovative and growing criminal ecosystem and how cybercriminals operate in different portions of the Internet provide helpful information about potential or current attacks targeting vulnerable individuals or organisations. This research offers new insights for researchers, the industry and law enforcement agencies about the modus operandi of cybercriminals which can be studied using environmental criminology theories to draw conclusions on how the Web environment creates criminal opportunity in cyberspace [47]. As a result, more effective defence mechanisms can be built to prevent and control cybercrime.

Theory is the core of all criminological research, and as has been shown in several studies, it can be applied to complex forms of crime [47]. By using the theories of environmental criminology, cybercrime can be treated more objectively. While RCT and RAT can be beneficial in understanding the criminal event, SCP helps prevent it [14]. The idea is to identify the circumstantial factors that lead criminals to commit malicious activities in each Web environment and apply the necessary measures to modify those factors and thus prevent cybercrime.

According to RCT, the first thing to consider is that criminals make rational decisions based on the factors surrounding the malicious activities they want to conduct [22]. One of these factors is the risk of detection. Cybercriminals seek to keep this risk low by operating from underground infrastructures o platforms such as Tor. In the online crimes analysed in this work, it has been observed that criminals use Tor to increase their privacy and remain anonymous from law enforcement agencies. For instance, while some criminals on the Surface Web access stolen accounts from mobile devices that can be easily traced, the Tor browser is used most of the time on the Dark Web.

Following RAT [15] and its emphasis on identifying the ”crime triangle”: a motivated offender, a suitable target and a lack of a capable guardian, it can be

inferred that privacy-enhancing networks attract more motivated offenders willing to engage in deviant conduct. At the same time, these networks make the application of enforcement policies complex, including activities such as monitoring and surveillance; consequently, a lack of guardianship can be perceived in this environment, which is why criminals carry out their malicious activities to find suitable targets in cyberspace. Below are several instances of RAT in the cybercrimes analysed:

Stolen account credentials. Cybercriminals are increasingly turning to underground forums in the Dark Web to buy and sell stolen accounts. Therefore, there is motivation to steal private data and trade it in underground forums on the Dark Web. Also, by providing the community with valuable stolen information, offenders gain respect and reputation from other members, which is another motivating factor in obtaining account credentials.

Malvertising. Malware distribution through advertisements (malvertising) is more attractive to criminals because it is a more efficient dissemination mechanism. The implicit trust that exists between publishers, advertisers and ad networks makes these malicious ads reach more (suitable) end-users exposed to these ads. At the same time, by distributing these ads through ad networks located on the Dark Web as hidden services, it is more difficult to track their authors.

Underground forums. As mentioned for stolen accounts, criminals are motivated by the fact that there is a higher level of trading activity in underground forums of the Dark Web. Likewise, as the results of this work have shown, different types of malware used to perform cybercriminal operations (unauthorised access to ICTs, spam delivery, DDoS services, money laundering) are among the most frequently offered assets in the Dark Web forums. These are criminal business schemes known as 'crimeware-as-a-service' (CaaS) [131]. Therefore, there is a motivation for highly skilled cybercriminals to monetise these services in underground forums and similarly others with limited skills who seek to contract these services. Another motivation is the expectation of exchanging knowledge and creating associations to find and exploit more criminal opportunities. In terms of guardianship, underground forums are more challenging to monitor. As a result, a lack of control and surveillance is

perceived.

6.5 Applying Situational Crime Prevention

The information security field has developed technical cybercrime prevention measures, which are a form of SCP [30], to prevent cybercrime. A few examples are malware detection software, intrusion detection/prevention systems and firewalls. However, technical measures alone cannot prevent cybercrime. Understanding the crime event, especially the circumstantial factors that create criminal opportunities in cyberspace, is critical to developing a deeper theoretical understanding of cybercrime. As a consequence, SCP frameworks and methodologies can be designed to prevent cybercrime. Due to the different forms that cybercrime can take, the application of SCP techniques should be adapted to a variety of offences; therefore, the information about the modus operandi of cybercriminals for the three types of cybercrime studied in this thesis may be used to develop more efficient frameworks.

The development of SCP frameworks should include policy and technology-based solutions aimed at preventing specific types of cybercrime. The basis of an SCP framework should start with identifying threats and vulnerabilities against different types of ICTs that will be mitigated with the techniques proposed in the framework. This identification is achieved by studying how criminals operate in the wild in a manner similar to what has been done in this investigation. Then, SCP techniques are applied to all identified threats to alter the Web environment in such a way that malicious online activities become more difficult to execute, riskier, less rewarding and less excusable for cybercriminals.

6.5.1 Specific Countermeasures

Using the knowledge gained from this thesis, several countermeasures using SCP techniques can be implemented to benefit the cybersecurity ecosystem. Some of the countermeasures proposed may already be implemented by some online services but it is important to point them out because they are not broadly applied on the Internet.

6.5.1.1 Stolen Accounts

Online resources based on personal accounts should implement strict password and account management policies and practices, especially for popular online services. For example, enabling multi-factor authentication may reduce the opportunities for cybercriminals to take control of an account. In addition, login attempts from foreign computers may be recognised by implementing device fingerprinting to identify the user's device. Even if the online services block login attempts from unrecognised devices or locations, administrators could send warnings to the user when a login occurs from a Tor node. Similarly, security awareness campaigns help the public to understand and mitigate risks. On the user side, enabling malicious software protection on their devices is important to avoid data theft.

6.5.1.2 Malvertising

Those involved in the online advertising ecosystem such as publishers and ad networks should develop policies to detect and monitor malicious ad networks and malicious ads. For instance, data analytics including machine learning can be implemented to detect malicious ads and track irregular ad redirections. Also, static analysis of online advertisements and malicious executables used in malvertising can detect malicious activity. Therefore, malicious campaigns can be stopped as soon as they are released. These analysis techniques should identify maliciousness originating from users connecting from Tor browsers to keep an updated knowledge about the risk of using anonymity networks. At the same time, browser developers should implement real-time analysis of JavaScript execution to detect malicious ads. As a result, malicious ads are not delivered to end-users. Same as for other types of cybercrime, enabling malicious software protection on their devices is important to avoid malware infection.

6.5.1.3 Underground Forums

Since underground forums are usually run by cybercriminals, the challenge to deploy countermeasures is greater in this environment. In this case, law enforcement agencies are the bodies in charge to disrupt or shutting down underground forums

aided by the research community. Therefore, it is important to analyse forums to gain deep insights into the ecosystem. To do so, data analytics including machine learning techniques should be implemented to identify threats in underground forums. As a consequence, popular assets including goods and services traded in underground forums can be analysed in order to make them unavailable or unusable. Thus, they lose their value and cannot be shared in the forums. This is especially important when it comes to forums where zero-day vulnerabilities can be found. Similarly, key actors can be identified by searching for administrators or prominent contributors. As a result, they can be traced to other forums and even to a physical environment to finally get caught.

6.6 Future Work

Although future work suggestions have been outlined for the cybercrimes studied in this thesis, future research also needs to focus on understanding cybercrime in a more detailed way from a ”cybercriminology” perspective. The idea is to use the information security measurements' knowledge and match it with environmental criminology theories. To this end, considerably more work will need to be done to determine the exact mechanisms that influence criminal decisions at different layers of the Internet and the specific elements of the ”cybercrime triangle” that converge in the digital realm. In addition, more studies are needed to combine technology-based solutions from the perspective of InfoSec and research in environmental criminology to propose and test practical frameworks that can be applied to different types of online crime. Then, further validation of the effectiveness of these frameworks will be necessary to continually improve their application on malicious activity that takes place on both the Surface Web and the Dark Web.

6.7 Ethical Discussion

All the academic research is subject to ethical considerations regarding the validity of the results presented by researchers. This section will discuss some ethical concerns that may arise from the research projects presented in this thesis. This research is focused on measuring malicious activity; therefore, it is essential to observe actual

incidents on the Internet which may involve various ethical issues related to privacy. I will go through the different projects to explain the ethical consideration faced during the research.

6.7.1 Stolen account credentials

The chapter about stolen accounts compares data from two different projects conducted by my first supervisor Dr Gianluca Stringhini. The first project collects data from Gmail honey accounts released on the Surface Web, and the second project, which is part of my research, collects data from accounts released on the Dark Web. Both projects received ethical approval from UCL. After the data collection stage, the comparison itself does not pose ethical issues.

The experiments were developed considering several ethical factors in order not to affect actual Gmail users. First, the default `send-from` address of the honey accounts was altered so that when an email is sent from any of them, it was sent to a controlled SMTP mail server that was set up to receive and store these emails without forwarding them to the intended destination. The send-from address was changed using the settings menu within each Gmail account. This measure was taken to avoid abuse from cybercriminals. Similarly, we collaborated with Google to ensure that accounts are suspended when they are hijacked or in case of problems beyond our control.

It is important to note that the only personal information that our experiments would collect is the cybercriminal's IP address and their user agent, which identifies the Web browser in use. Although this information identifies the cybercriminal's device, it is not enough to identify them. At most, it can give us information on the country or the area where the cybercriminal's device is located. Although we do not collect any sensitive information, we encrypted our data and will securely delete the data after the end of the project.

6.7.2 Malvertising

In a nutshell, the malvertising chapter uses a Web crawler that gathers the HTML code or the JavaScript code generated by public websites on the Surface Web and the

Dark Web for further comparison. Although this information is open to the public, we perform our experiments with various ethical considerations in place. We crawl our list of websites at a reasonable rate (less than one visit per day for each website for each IP) with a legitimate user agent string. Furthermore, we only save the data we need for our analysis, i.e. links related to advertisements.

After collecting the advertising links, we submit them to detection systems such as VirusTotal and Google Safe Browsing to check for maliciousness. In this stage, we also take into account some ethical considerations. For example, we use the API provided by our detection systems which establishes a daily request quota to prevent abuse. The information provided by our detection systems is in the public domain as their databases rely on the output of many different antivirus engines, website scanners, file and URL analysis tools, and user contributions.

Similar to our previous project, the data comparison does not present ethical issues as we are not analysing any personal information. Dr Stringhini's previous works and the research conducted by the computer security community show that the ethical issues associated with this type of research are minimal.

6.7.3 Underground forums

Several ethical issues should be considered when collecting and analysing data from underground forums. Pastrana et al. [124] argue that while data collection is about understanding underground forums as ICT systems, data analysis is about understanding people's behaviour in forums. Therefore, it is important to recognise this distinction when considering ethical issues in forum research.

In terms of data collection, we crawl underground forum websites and extract HTML code from posts covering a period of one year. It is important to note that the volume of data extracted is not significant. We only extracted the first post for every thread in each forum. For this step, we take into account the following ethical considerations. First, since we extracted a small portion of the forums, we only needed a single visit to collect all the data required for our analysis. Therefore, we do not generate a large amount of traffic on the crawled websites, and no Tor relay was abused to visit the Dark Web forums. For the Surface Web forums, no CAPTCHA

was required to access the websites. On the other hand, a CAPTCHA was needed to enter the Dark Web forums. We introduced them manually for each site. In any case, some authors [55, 132] argue that bypassing CAPTCHAs is ethical because session cookies can be provided to the crawler as a functionality offered by site administrators. Second, all the posts extracted are publicly available and login is not required to access them. As the same authors point out [55, 132], the use of crawlers for data collection does not compromise a website or any user using it.

Regarding the data analysis, no identifying information is extracted from the publications to avoid harming individuals. For instance, usernames are anonymised using a cryptographic hash just for reference; only the content of posts is analysed. These posts do not have contact information because subsequent communication is done through the private message system of the forums, to which we do not have access. Similarly, an escrow system is used to complete the transaction, but it requires creating an account. In addition, we analyse trading activity by aggregating all the posts in each Web environment; thus, we do not identify any individual member in the forums. Moreover, no additional private information was identified in the posts. We analysed a sample of 100 posts in each Web environment and found that none of the posts contained confidential information such as names, email addresses, or cryptocurrency addresses. Finally, our crawler only collects text data to avoid downloading any malware or any illegal content in the form of photos or videos.

Chapter 7

Conclusions

This thesis has presented an analysis of three types of cybercrime conducted in two different Web environments: the Surface Web and the Dark Web. This analysis has been used to explore the research question: Does the Web environment influence cybercriminal activity? The findings of this work can be beneficial to the research community in the fight against the constant evolution of cybercrime. This chapter summarises the research, describes the effectiveness of our methods for each study and concludes on the research question.

This research aimed to understand the modus operandi of cybercriminals at different layers of the Internet. Several tools such as crawlers, honeypots and parsers were used to measure malicious activity on stolen accounts, online advertising and underground forums on the Surface Web and the Dark Web. Then, based on a quantitative analysis of the data collected for each type of cybercrime, it can be concluded that the Web environment is an important factor considered by cybercriminals when they conduct illicit activities on the Internet.

The existing literature has investigated malicious activity related to different types of cybercrime on the Surface Web and the Dark Web separately. This thesis is the first work attempting to discover how the same type of cybercrime develops in both Web environments. Essentially, this work attempts to answer the research question by strategically comparing malicious activity in both Web environments. The differences between the two environments are highlighted, and the common characteristics of the cybercrimes analysed are outlined to understand criminal

behaviour in different layers of cyberspace.

7.1 Stolen account credentials study

The first study seeks to discover what happens to stolen email accounts whose credentials have been leaked on paste sites and underground forums on the Surface Web and the Dark Web. The methodology used for the measurement involves comparing a set of predefined behaviours extracted from the literature, which are based on the interaction of criminals with stolen accounts. This analysis concluded that cybercriminals are more cautious and try not to leave traces when logging into accounts using credentials obtained from the Dark Web. Additionally, interesting case studies occurred during the experimentation period that showed the willingness of criminals to commit criminal acts with the accounts, especially on the Dark Web. While the limitations related to time and account availability presented in Chapter 3 limit the results' generalisability, this approach provides new insights into the ecosystem of stolen online accounts from different sections of the digital realm.

7.2 Malvertising study

In this study, the online advertising ecosystem is explored from the end-user perspective. A growing number of users seeking to protect their privacy have chosen to use the Tor network for their online activities instead of a regular connection to the Internet. This analysis aims to determine whether the type of access that the user employs to browse the Internet affects the distribution of advertisements by advertising networks, specifically malicious advertisements. To this end, some new methodological approaches were developed. The first is related to the automatic collection of online ads. In this study, existing methods for crawling websites and identifying ads were improved. These techniques were shown to extract the ad network and landing page from a displayed ad on most websites. This research clearly illustrates that the ad delivery process, including malicious ads, is similar for the two connection types. However, the fact that ads delivered to Tor users are more likely to redirect traffic to low ranked landing pages, it also raises the question of whether cybercriminals are setting up fraudulent ad networks targeting Tor network

users.

7.3 Underground forums study

The third and final study compares trading activity between underground forums located on the Surface Web and the Dark Web. Trading activity was examined in terms of different variables related to transactions in forum posts. After extracting posts from various underground forums in both Web environments, an improved version of a tool that identifies trading information in posts was used to obtain the required variables. This tool builds upon existing work that uses natural processing language to extract high-level information from forum posts. The methodology used proved effective in extracting transaction types and prices, but further research is needed to identify products more accurately. The results indicate a higher level of trading activity in underground forums on the Dark Web, most of which refer to more sophisticated hacking services. Although these results shed light on the products and services traded in underground forums in both environments, forums where an invitation is needed to join the community should be included to gain a deeper insight into trading activity in forums where high-skilled cybercriminals operate.

7.4 Final remarks

The studies presented in this research were used to determine whether the Web environment in which cybercriminals conduct malicious activities influences their behaviour. This thesis found evidence to support the argument that there is a higher level of criminal activity on the Dark Web and that criminals are more stealthy in this portion of cyberspace. The results suggest that cybercrime on the Dark Web is more sophisticated compared to the Surface Web.

Despite the limitations noted in previous sections, this work has made positive contributions to the academic literature with detailed measurements that provide information about cybercrime, especially how it differs between the Surface Web and the Dark Web. These contributions can also benefit security professionals seeking to understand better online crime in different layers of the Web and improve techniques for cybercrime mitigation.

Since cybercrime research should be the key information source for law enforcement agencies, security professionals, policymakers and the research community in general, this work can be a starting point to explore more types of cybercrime and understand how they differ depending on the Web environment. Therefore, cybercrime can be approached from several angles, thus preventing the growth of these malicious activities in any section of cyberspace.

This thesis argues that environmental criminology theories can be used to identify specific criminal events in cyberspace so that new practical frameworks can be proposed to combat cybercrime by combining crime prevention techniques with technological measures and information security policies. As a consequence, more innovative approaches can be developed to prevent and control the different types of cybercrimes in the wild.