Understanding and Characterizing the Cybercriminal Ecosystem Enabling Attack Innovation at Scale

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van
de rector magnificus prof. dr. S.K. Lenaerts,
voor een commissie aangewezen door het College
voor Promoties, in het openbaar te verdedigen op
donderdag 7 maart 2024 om 11:00 uur

door

Michele Campobasso

geboren te Castellaneta, Italië.

Dit proefschrift is goedgekeurd door de promotoren en de samenstelling van de promotiecommissie is als volgt:

Voorzitter:	prof. dr. E.R. van den Heuvel
Promotor:	prof. dr. S. Etalle
Copromotor:	dr. L. Allodi
Leden:	dr. N. Zannone
	prof. dr. R. Leukfeldt (Netherlands Institute for the Study of Crime and Law Enforcement – NSCR)
	prof. dr. A. Hutchings (University of Cambridge)

Het onderzoek of ontwerp dat in dit proefschrift wordt beschreven is uitgevoerd in overeenstemming met de TU/e Gedragscode Wetenschapsbeoefening.

The work in this dissertation is partially supported by the ITEA3 programme through the DEFRAUDIfy project funded by Rijksdienst voor Ondernemend Nederland (RVO), Grant No. ITEA191010

The work in this dissertation is partially supported by the INTERSCT project, Grant No. NWA.1162.18.301, funded by Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)

Printed by: ADC Nederland - 's-Hertogenbosch

Cover design: Michele Campobasso

Original cover art: IMGAEP, Reddit contributor. Generated with Disco Diffusion 4.1.
Link to resource: `https://i.redd.it/sk7vyzzcmmc81.png`

A catalogue record is available from the Eindhoven University of Technology Library.
ISBN 978-90-386-5971-8

An electronic version of this dissertation is available at `http://repository.tue.nl` and `http://michelecampobasso.github.io/`

It is an important and popular fact
that things are not always what they seem.

Douglas Adams, The Hitchhiker's Guide to the Galaxy

Summary

The digitalization of our society has improved the quality of human life at large, but also created new opportunities for malicious (cyber) actors to exploit weaknesses of digital systems and to find new venues to perpetrate fraud. It is well known that a sizable portion of cybercriminals meet in hundreds of online communities (often in the so-called "dark" or "deep" web) to engage with like-minded individuals, share knowledge, and trade offensive tools and services. Monitoring and extracting information from these communities is a challenging task, yet crucial to obtain valuable threat intelligence, which can be used to better prepare against new threats and deploy tailored defenses to thwart attacks originating from the underground. Private sector and law enforcement agencies alike would then be able to better protect final users and societal institutions from the risks associated with criminal activities.

However, as of today it is unclear whether all hacking communities (generally in the form of forums) equally contribute to supporting cybercriminal activities and technological innovation. Many underground markets seem to only trade low-tech products like old password leaks and obsolete malware that is immediately detected from most existing defenses. Still, cybercrime and its economic and societal impact increase year by year. Considering this, it is critical to investigate the properties of marketplaces in relation to the traded products, to identify what are the characteristics that distinguish those where the trade of effective cybercriminal technology happens from those where it does not. Lacking this ability may lead to biased threat intelligence or overestimating the risks associated with petty crime or scammer activities. However, extracting forum data from underground criminal communities (particularly prominent ones) can be a non-trivial task, due to the scale of the phenomenon, their diversity in terms of access, interaction, language, and the countermeasures they put in place against internet crawlers. Hence, we formulate the following research question:

How can we identify which cybercriminal marketplaces can support the trade of innovative offensive products and services, and how can we effectively monitor their activity to evaluate the threat they pose?

We first conduct a preliminary study on the segregation of underground marketplaces in relation to the maturity of the offensive tools and services they provision. We aggregate information from industry reports and scientific literature, and we corroborate these findings by manually exploring 13 underground markets. Our study finds indication that more segregated markets tend to offer more mature offensive tools, and they tend to be better protected against unwanted access and monitoring activity.

Following this, we must address the problem of circumventing network monitoring performed by markets to detect and thwart automated data extraction. Therefore, we investigate crawler detection mechanisms to devise a method and implement a prototypical crawler

(THREAT/crawl) for stealthy automated data collection for segregated underground markets. We design the crawler to offer a simplified supervised procedure to learn how to crawl a forum and what data to collect, and we test it against 7 live forums. We publicly release the code and documentation of the crawler.

Making use of the techniques and insights developed in the first stage of this work, we then investigate the presence of prominent threats in the underground. During our investigation across 30+ underground forums, we discover Genesis Market, an (at the time) emerging market proposing a novel criminal service for Internet user impersonation at scale. After obtaining invitation codes to access the market from affiliate communities and users, we infiltrate it with multiple identities and use those to investigate its operations and derive a model of the threat it generates, that we called "Impersonation-as-a-Service" (IMPaaS). Further, we develop a specialized crawler based on THREAT/crawl to scrape the whole market's offer, enabling us to assess the scale of the threat globally and perform statistical evaluation of products pricing in relation to their characteristics. We devise a rigorous statistical methodology based on dimensionality reduction and multi-factor analysis to account for the intrinsic limitations of data collection in this domain, to estimate market revenues, scale, and attacker (i.e., market customers) preferences, and evaluate the overall posed threat. We conclude that Genesis is a mature marketplace, and IMPaaS is a (now) established threat at scale that could be used as a convenient alternative for initial access to mount targeted attacks. The extracted datasets are available to interested researchers.

Finally, we condense findings and insights from our research to investigate the characteristics of cybercriminal marketplaces trading innovative threats like IMPaaS. We identify issues typical of "markets for lemons" and derive mitigating mechanisms employed by markets by manually investigating 20+ cybercriminal marketplaces and the relevant literature. We cast the obtained dimensions into a preliminary framework based on the Business Model Canvas to evaluate what business aspects affect the trade of innovative products. Our findings show that "functioning marketplaces" on average tend to be more segregated, scrutiny their sellers, and are concerned with offering a fair and competitive marketplace.

In conclusion, this thesis shows that the underground ecosystem is diverse, and it is possible to identify and stealthily monitor the fraction that convincingly solves trade problems and drives innovation, thus obtaining more refined threat intelligence, while better understanding the criminal decision-making process. Ultimately, we argue that identifying factors that support innovation in the cybercriminal landscape has the potential to provide insights on the criminal's decision making, hence allowing defenders to, potentially, be prepared for the 'next big attack' before it arrives.

Contents

1

Introduction

1.1. Scope and motivation

1.1.1. Online underground forums as criminal marketplaces

Both traditional and digital crime need meeting places with specific characteristics to thrive. These places represent venues where criminals get to know other co-offenders, find work opportunities, and trade stolen and illegal goods; these places are sometimes called *offender convergence setting* [239, 93]. In the digital context, IRC servers were initially used as an offender convergence setting [274, 98, 249], but these started to lose popularity in favor of forums in the early 2000s [16]. With the advent of new technologies, faster Internet, powerful and affordable personal computers, and the growing concerns posed by unencrypted communication channels for unlawful activities, forums have become the places where cybercriminals meet [16, 274, 170]. To protect the anonymity of their members and administrators, some of these forum marketplaces are hosted in the dark web, accessible via Tor network[1] [52, 107, 238, 262]. Forums allow their members to engage in technical discussions on anonymity, vulnerability research, fraud and, most importantly, trade products. Many of these forums feature more or less specialized marketplace sections where participants can post advertisements to sell illicit products and services [274]. Some marketplaces may be completely dedicated to specific frauds; for example, carding forums are some common specialized underground markets [275]. A carding forum often features sections focused on several aspects of carding, offering its members a place to discuss how to create cloned credit cards starting from 'dumps' (card details stolen from a card's magnetic stripe with the use of devices called 'skimmers'[2]), how to spend cards' balance online (e.g., how to buy and safely pick up goods from e-commerce with stolen credit cards), and how to cash out offline (e.g.,

[1] In a nutshell, Tor network is an overlay network that relays traffic through different nodes to greatly improve the privacy of its users [77]. Apart from the nefarious usages that criminals do of Tor, The Tor Project's goal is to offer anonymity to its users, improve their privacy, and offer protected communications to political activists and unrestricted access to Internet whereas censorship is an obstacle.

[2] A 'skimmer' is a device installed on top of an ATM or POS that replicates the appearance of the original device. They are used to read the PIN of a card and the data contained in the magnetic strip to subsequently be available to the attacker.

with the use of money mules) [117]. Some markets feature the trade of additional products that could be indirectly related to that specific fraud; in the case of carding again, it is possible to spot the presence of sections trading counterfeit documents, useful in identity theft attacks, and a plethora of proxy services to preserve the anonymity of the attackers [211, 260]. Aside from carding marketplaces, there are more generic marketplaces that cover a broader range of products and services that can be used to support other illicit activities.

Zooming out from the specific, according to industrial reports and literature [252, 262, 36, 3, 127, 131], underground cybercriminal marketplaces mostly offer three categories of items: stolen information, products, and services. Among the *stolen information* and apart from stolen credit card information, the underground features leaked or stolen credentials, or the more complete 'logs' (collections of stolen credentials, cookies, together with some information about a victimized system, originating from a single browser); in addition, markets offer collections of personally identifiable information (PII) including name and surname, address, phone number, and more sensitive information like social security numbers, document numbers, and their digital copies. *Products* can be either physical or digital; among physical products, aside from those that are carding specific (e.g., blank EMV cards[3] and the associated devices to write them, or cloned credit cards ready for cash out), markets trade counterfeit goods like documents and cash. Digital products instead include but are not limited to malware, remote administration tools (RATs), exploit kits, and phishing kits. *Services* include an extended range of options for different purposes; they could be services that support the creation and maintenance of illicit websites (Bulletproof-Hosting-as-a-Service) or improve the anonymity of attackers (e.g., proxies, VPNs, SSH tunnels, ...), licensed copies of malware (Malware-as-a-Service), services improving a malware's stealthiness ('crypter', 'packer', and 'obfuscator' services), or full-stack solutions to perform ransomware attacks (Ransomware-as-a-Service). Other services connected to malware include payload delivery with the use of spamming, called 'traffing', and pay-per-install. Attackers could also consider buying access to compromised systems from initial access brokers (IABs), thus outsourcing the need for technical knowledge necessary to carry out an attack and obtain a stable foothold within an organization; others could be more interested in disrupting the availability of a system with the use of booter services (DDoS-as-a-Service). Finally, some services help criminals in cashing out the revenues of their wrongdoings, with cryptocurrency mixers and money mules.

Considering the multitude of products and services traded in these communities and the role they play as criminal meeting places, studying them could offer a valuable opportunity to analyze the mechanics of trade and the provision of illicit products in the context of the underground cybercriminal economy.

1.1.2. Online underground forums as social opportunity structures

In the digital space, *virtual offender convergence settings* offer cybercriminals the possibility to meet with like-minded individuals. These convergence settings come often in the shape of cybercriminal forums. These forums create a range of opportunities: attackers can expand the boundaries of their social networks, moving from their offline acquaintances to

[3]Empty cards with magnetic strip and chip that can be used to 'write' a copy of a stolen card.

meet transnational partners, customers, and suppliers, with the benefit of preserving their anonymity and physical safety [164, 239]; in addition, these communities offer them the opportunity to discover and engage in new forms of fraud [69] while offering access to an unparalleled supply of different products and services that build the value chain to create new business opportunities [36]. This is well framed by the concept of *social opportunity structure*, which suggests that people do crime in response to temptations and opportunities for a (financial) gain [59] and that the social interactions within such communities are instrumental to reach these goals [258, 164]. In addition, virtual offender convergence settings have the ability to bridge gaps in criminal organizations, thanks to the multiple roles that members can cover [258]. For example, the late 2010s witnessed the rise of ransomware gangs; the massive Conti ransomware group leak in February 2022 sheds light on the structure and management of a complex organization: managers, human resources, finance, and developers worked together on different parts of the operation. Interestingly, this leak offered insights into the recruitment process of the group: many forums hosted job opportunities offered by the human resources personnel, but the interviews happened off-site via safer, private communication channels [220] (as similarly reported by [239]).

1.1.3. Research scope

Before identifying the general direction of this thesis, we define its scope based on some observations. First, the cybercriminal ecosystem is vast, and literature sometimes offers conflicting views on the role and characteristics of underground communities. Although defining a cybercriminal community as a marketplace may seem belittling, the context in which we operate may have a substantial overlap between the two concepts; when considering cybercriminal communities as social opportunity structures, actors are considered as rational players guided by a cost-benefit analysis that favors the engagement with illicit activities in perspective of large financial gains, despite the involved risks. Therefore, from our observational standpoint (as in, researchers infiltrating underground communities), the visible 'symptom' of these interactions is the participation in trade activities. Clearly, communities do not solely offer a marketplace section, but they host sections where community members can have fruitful conversations. However, to the best of our knowledge, the criteria for identifying what makes this possible or what created the perception that a virtual convergence setting is adequate to meet criminal partners are unclear. In addition to that, investigating the offenders' choices with qualitative methods may pose some challenges; first, we do not know which participants make up a representative sample of the offenders' population within a community; second, to draw conclusions on the ecosystem as a whole, it is necessary to conduct investigations on a large scale. Therefore, investigating in this direction with the use of interviews and surveys while ignoring the aforementioned challenges would likely lead to only partial, or potentially biased, results. Furthermore, discussion within communities appears to gravitate around trade, and follow-up conversations may happen in private or off-site; to obtain full insights into these dynamics, we would require the analysis of dumps and leaks of said communities. However, this could introduce biases in the interpretation of the results; data leaks are relatively rare events and they are often incomplete, while dumps of a community are often the result of law enforcement operations. In either case, the information may be partial and refer to a past state of the examined communities, leading to inconclusive results.

Despite the relevance that qualitative approaches have to the overall comprehension of the cybercriminal ecosystem, at this stage, studying the subject with quantitative methods could greatly help us to mitigate the previously identified challenges, thus gaining insights on its dynamics while developing tools that can support the analysis of active communities.

1.1.4. Research gaps

To formulate our research questions, we first identify some gaps present in the literature.

Assessing product quality in underground markets

From the literature, it emerges that effectively trading valuable products in the cybercriminal context is not a trivial task. The seminal work from Herley and Florêncio [122] paves the road for the discussion on what seems to be one of the biggest underlying problems of the underground economy: *quality uncertainty*. The authors argue that, under many circumstances, a buyer cannot assess the quality of a product before buying it, because of the information asymmetry between buyers and sellers [11], ultimately causing adverse selection (i.e., the incapacity of buyers to get the product that fits their needs). According to the theory of quality uncertainty described by Akerlof [11], sellers are in a privileged position, owning complete information about their products, and they could offer buyers incomplete or plainly wrong information to appear more competitive against other sellers. In the cybercriminal context, not offering extended information about an offered product could make sense, as the seller would avoid the risk of spoiling it (e.g., giving details on how a weaponized 0day exploit works could give away too much information, allowing other people to reproduce it) [182]. However, this creates the opportunity for sellers to use their position to willingly scam victims. For example, it is relatively trivial (with simple internet searches) to find some famous marketplaces mostly known for being specialized in drugs which feature sections advertising hacking tools. When examining these advertisements, the majority of them comes with a generic description of the product, with little to no evidence about the product's functionalities, and they are offered at a surprisingly low price tag (Figures 1.1a, 1.1b). To aggravate uncertainty, digital goods like stolen credit card information can be sold multiple times. The buyer has no warranties that the seller will offer this information only once; in that case, even if the offered bank account truly has the reported amount, multiple buyers of the same credentials will compete for cash out its balance, reducing the expected value.

In literature, some studies look into the features that black markets implement to mitigate quality uncertainty; for example, Yip et al. look into the facilitating factors for trade limited to the context of carding forums [274] and report that marketplaces implemented via forums, as opposed to those over IRC, support the identity establishment for market participants (despite the anonymous context) thanks to the existence of historical records of interactions among members and hypothesize that member screening and rippers punishment have a positive effect on trade in a carding forum. On the other hand, Dupont et al. [81] examine a data leak originated from the infamous Darkode cybercriminal forum, which performed member screening; they find out that, despite the existence of this procedure, trust among peers remained elusive and interactions were often fraught with suspicion and accusations, even when hackers who were considered successful were involved in the trade. In another

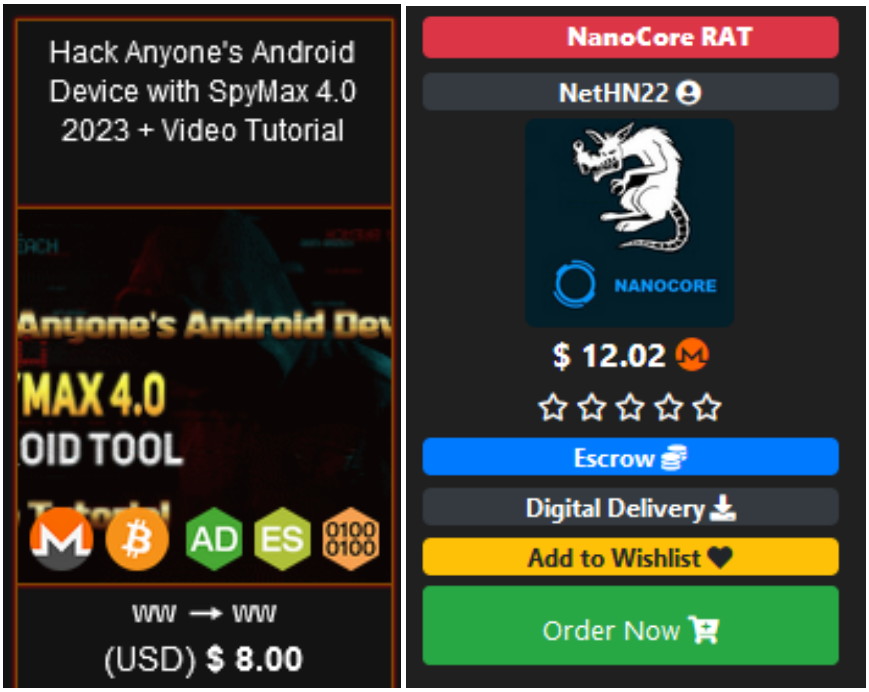


Figure 1.1: Malware of unclear quality offered on two different underground marketplaces.

work, Dupont et al. examine the role of reputation systems in underground communities and find out that only a small percentage of market participants use them, causing the impact of reputation as a signal of trust to be marginal [82]. Holt provides a qualitative analysis of trust-establishing features in underground marketplaces, and identifies the mechanisms of product verification, reputation, and guarantor systems as beneficial to the overall mitigation of quality uncertainty [127]. In light of this, it emerges that prior work has successfully identified the problems affecting trade in distrusted environments and some of the associated mitigations are in place, but we still lack a clear understanding of how these mechanisms are implemented in the wild, and what is their overall composition in underground markets to be able to positively influence the success of a marketplace.

Marketplaces may not equally matter

As suggested from [11, 122], in the presence of information asymmetry, dishonest sellers thrive by selling products of lower quality than advertised. This has two negative effects on the market: first, it lowers the expected average product quality, making buyers less inclined to buy a product considering the high risk of being scammed, and increasing the overall costs they face to acquire a good product (i.e., if one has to buy multiple products to find a functional product, the final price will be inflated by a *scammer tax*) [274, 122]; and second, it forces honest sellers to lower their prices to compete with other sellers, reducing their profits to eventually push them out from the market, and ultimately causing

the market to fail [11]. Given the existence of these problems, it is unclear whether the activity of all cybercriminal marketplaces should be considered as a threat equally. Lusthaus suggests that the research does not equally scrutinize all underground communities [171]; the author identifies different tiers of marketplaces based on their access model (from free access to high-profile, closed communities) and argues that those that are more segregated are underrepresented in literature. Furthermore, the author suggests that marketplaces belonging to different layers may include different fractions of the underground economy, thus excluding a portion of these may hinder our comprehension of cybercrime as a whole and the planning of intervention [171]. On a similar wavelength, in literature multiple instances of studies aim at identifying the prominent actors within a community [207, 70, 127, 128], potentially supporting law enforcement intervention; however, similar studies on the role of markets as a whole are scarce, and the selection of marketplaces as subject for studies is often based on their popularity [36, 81], or other indicators like years of activity and number of participants [207, 12, 117]. At first glance, markets fraught with information asymmetry appear to build the majority of 'hacking forums', and their inclusion as potential sources of threats may be (at least partially) the cause for some unrealistic assumptions on the damages caused by cybercriminal activity. For example, some scholars indicate that industry largely overestimates the costs and negative externalities caused by cybercrime [20, 246]; furthermore, it seems that there is no general agreement between threat intelligence providers about the sources of threat: some studies highlight how the Indicators of Compromise (IoC) produced by threat intelligence companies show marginal to no overlap [43] and, as a result, defenders, on average, rely on 7.7 TI providers to ensure coverage in their threat feeds [216]. With that considered, it remains unclear whether all markets 'matter' the same in the overall threatscape. We speculate that there should be a portion of markets mature enough to have addressed trade issues by any means, hence becoming able to pose an actual threat; furthermore, threat actors may have preferences that optimize different needs aside from those of an economic nature. Therefore, we believe that it is necessary to investigate *which* are the *measurable* characteristics of markets in relation to their foundational problems, and explore which play a positive role in the decision-making of attackers when selecting their trading venue of choice. By doing this, we aim at understanding *what* features cybercriminals prefer, thus allowing us to better characterize the venues where the mass of illicit activities happen and to produce useful information for law enforcement operations. [171]

Automation of market infiltration and data extraction

To investigate how markets operate and to identify their differences, it is necessary to obtain data from these communities. Data could be obtained from data leaks, law enforcement operations, or data collections. There is a multitude of studies based on data from the first two scenarios [81, 36, 188]. However, these datasets have some limitations; data leaks are relatively rare, data tends to be outdated, or it allows only to perform a post-mortem inspection of the dynamics of a market [209]. Nonetheless, despite their rarity, this data offers researchers a privileged view of the operations of the marketplace, showing private messages and interactions otherwise invisible from the outside. Alternatively, data extraction is the only way to obtain fresh data from a live market. In literature, we witness several studies that rely on data collected from live markets [12, 33, 52, 92].

However, gaining access to part of these marketplaces, especially high-profile ones, is a challenge in itself. These markets screen members with the use of applications, asking for references across affiliated communities and proof of intentions, either with a 'curriculum' or by committing economically with the payment of fees [33, 188, 171]. In most cases, investigators need to create believable identities online in a lengthy process and use them across different communities, interacting with their members to gain reputation [171]. In addition to these challenges, once access is granted, data extraction is a non-trivial process. In fact, there is evidence of marketplaces featuring network monitoring capabilities to detect and ban accounts manifesting signs of crawler activity [69, 209]. Other markets feature 'traps' that allow the identification of whether the session is automated, deliberately make complex DOMs, or require interaction with the aim of disrupting crawler operations [79, 156, 209]. Thus monitoring the activity and extracting data from these markets using automation remains a complicated task that hinders researchers' and law enforcement's investigations.

As an effect, research that relies on crawled data from live underground communities requires significant efforts by scholars to develop crawlers to extract information from them [209], although there are some exceptions where these datasets are made available to other scholars [213]. However, the crawling infrastructures used for these studies are rarely available to other researchers, and often these are *ad-hoc* solutions to extract data from a single market, making them potentially unsuitable for further research. The intrinsic difficulties of monitoring underground communities, together with the costs associated with the development of such tools hampers with the research [209]. Hence, to fill this gap, exploring novel methods to remain 'under the radar' is fundamental to extract fresh data from a wide range of illicit communities of different degrees of sophistication. Finally, providing a set of guidelines and possibly an open-source tool to the scientific community to conduct such investigations could greatly impact the quality of future research in this field.

1.2. Research questions

Main Research Question (MRQ)
How can we identify which cybercriminal marketplaces can support the trade of innovative offensive products and services, and how can we effectively monitor their activity to evaluate the threat they pose?

To begin with addressing our main research question, we primarily need to define the scope within which cybercriminals operate. Thus, we formulate the following research question:

RQ1: *How can we preliminarily characterize the space of underground marketplaces supporting the provisioning of offensive cyber capabilities?*

To address this research question, we investigate the relationship between marketplaces and their role as providers of offensive capabilities for the Access-as-a-Service threat model. To achieve this, we collect information from the relevant literature and industry reports, and then we conduct a preliminary exploration of underground communities.

During this preliminary investigation, we learn that some underground marketplaces employ bot detection and perform network monitoring to thwart automated data extraction. To successfully tap data from underground forums, a crawler must operate in disguise. To achieve this, we define the following research question:

RQ2: *How can we stealthily extract information on market activity from underground forum markets while circumventing crawler detection mechanisms, and scaling up monitoring to multiple forums?*

To tackle this issue, we first study prior work on stealthy crawlers and data extraction from cybercriminal forums, including crawler detection and detection evasion techniques. We then design a prototype that embeds these techniques and models human behavior to achieve stealth, and we benchmark its performances against human subjects. Finally, we extend this prototype to provide its users with a guided procedure that allows them to instantiate a crawler tailored to a specific forum and to repurpose it with ease for multiple targets.

To validate our assumptions about the existence of marketplaces fostering innovation and trade of cybercriminal products, we need to identify and evaluate potential candidates. To conduct this research, we need to employ the developed crawlers to extract data to analyze. Therefore, we ask:

RQ3: *How can these monitoring capabilities and market characteristics be used to identify and evaluate high-relevance cyber-threats?*

To answer this research question, we employ our crawlers to study an (at the time) emerging criminal marketplace for user impersonation at scale. With the collected data, we derive the underlying, novel threat model it operates on and look into the economic aspects of the platform. Following this, we perform additional data collection to measure the market activity to estimate its customers' preferences and use this data to inform and expand a cyber risk model.

As discussed before, we postulate the existence of differences in marketplaces, instrumental to the presence of trade of successful cybercriminal products. To investigate these, we introduce the following research question:

RQ4: *What characteristics differentiate underground forum markets capable of supporting high-relevance cyber-threats from those that cannot, and how can this difference be evaluated through market observation?*

To investigate these differences, we propose to look at the foundational problems affecting trade and map them to an economic framework. We analyze the implemented strategies by marketplaces to mitigate these issues and use these to describe said marketplaces. Finally, we compare these marketplaces and look at the recurrent characteristics of those featuring criminals who been indicted or convicted in the past, and we generalize our findings.

1.3. Thesis outline and contributions

This thesis is divided into three parts.

1. In Part I, we provide a preliminary overview of the underground markets' provisioning of offensive cyber capabilities, and we investigate the problems and solutions associated with the data extraction from underground markets. More specifically:

   1.1. In Chapter 2, we lay the foundations for this thesis, and provide a preliminary characterization of underground markets to answer to **RQ1**;

   1.2. In Chapter 3, we devise and evaluate a general methodology to stealthily monitor underground communities, supporting monitoring in adversarial environments, as per **RQ2**;

   1.3. In Chapter 4, we operationalize and extend the contribution from Chapter 3; we present a method and tool supporting researchers to scale up stealthy data collection from underground communities, as per **RQ2**.

2. In Part II, we employ the findings and tools from Part I to investigate a prominent cybercriminal service provider driving attacks at scale, and we model the threat and the risk it poses. More specifically:

   2.1. In Chapter 5, we look at an underground market that advertises itself as an innovative criminal service. To answer to **RQ3**, we infiltrate the criminal platform and observe its operations to derive the underlying threat model and the scale at which it operates; we quantitatively assess the product's characteristics, the associated pricing model, and its maturity as a global threat;

   2.2. In Chapter 6, we further investigate the criminal platform examined in Chapter 5 to further examine the scope of **RQ3**. We perform an analysis of the platform's threat levels, understand attacker preferences, measure the market's economic activity to estimate market revenues, and infer victimization rates.

3. In Part III, we extrapolate our findings on mature and innovative markets by building a general framework capturing the key characteristics of underground markets to distinguish those that can convincingly drive real-world threats from those that cannot. More specifically:

   3.1. In Chapter 7, we investigate the key aspects that support criminal activities in cybercriminal forums. To answer to **RQ4**, we propose a preliminary framework describing the structure and functionalities of underground forums; we obtained these features from the observation of underground markets and the analysis of relevant literature on economy and cybercrime. We then examine the characteristics of these forums in relation to the presence of notorious cybercriminals, and we use this information as a proxy signaling the threat level posed by the examined marketplaces;

   3.2. In Chapter 8, we discuss our findings from Chapter 7 and offer some perspectives for future research. In particular, we look at the increasing adoption of instant

messaging apps as new cybercriminal marketplaces. We argue that investigating their characteristics could inform us about the problem(s) they address, the market target they aim for, and could allow us to conjecture about their future and associated threat levels.

Finally, Chapter 9 concludes the thesis.

1.4. Publications

Our research lead to the following contributions (in inverse order of publication):

1. **Campobasso, M.**, and Allodi L., *Know Your Cybercriminal: Evaluating Attacker Preferences by Measuring Profile Sales on an Active, Leading Criminal Market for User Impersonation at Scale*, 32nd USENIX Security Symposium (USENIX Security 2023)
2. **Campobasso, M.**, Rădulescu, R., Brons, S., and Allodi, L., *You Can Tell a Cybercriminal by the Company they Keep: A Framework to Infer the Relevance of Underground Communities to the Threat Landscape*, 22nd Workshop on the Economics of Information Security (WEIS 2023)
3. **Campobasso, M.**, and Allodi, L., *`THREAT/crawl`: a Trainable, Highly-Reusable, and Extensible Automated Method and Tool to Crawl Criminal Underground Forums*, 17th Symposium on Electronic Crime Research (APWG eCrime 2022)
4. DeSombre, W., **Campobasso, M.**, Allodi, L, Shires, J., Work, JD, Morgus, R., O'Neill, P. H., and Herr, T., *A primer on the proliferation of offensive cyber capabilities*, Atlantic Council 2021, In-Depth Research & Reports
5. **Campobasso, M.**, and Allodi, L., *Impersonation-as-a-Service: Characterizing the Emerging Criminal Infrastructure for User Impersonation at Scale*, 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS 2020)
6. **Campobasso, M.**, Burda, P., and Allodi, L., *`CARONTE`: a Crawler for Adversarial Resources Over Non-Trusted, high-profile Environments*, 2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW) - 1st Workshop on Attackers and Cyber-Crime Operations

Publications to which I contributed, but that do not appear in the corpus of this thesis:

7. Rosso, M., **Campobasso, M.**, Gankhuyag, G. and Allodi, L., *`SAIBERSOC`: A Methodology and Tool for Experimenting with Security Operation Centers*, Digital Threats: Research and Practice (DTRAP). Volume 3, Issue 2, Article No.: 14, pp 1–29
8. Rosso, M., **Campobasso, M.**, Gankhuyag, G., and Allodi, L., *`SAIBERSOC`: Synthetic Attack Injection to Benchmark and Evaluate the peRformance of Security Operation Centers*, Annual Computer Security Applications Conference (ACSAC 2020)

I

The underground ecosystem and how to measure it

2

Background: a preliminary characterization of offensive capabilities provisioning from underground markets

This chapter is based on [Campobasso5]:
DeSombre, W., **Campobasso, M.**, Allodi, L, Shires, J., Work, JD, Morgus, R., O'Neill, P. H., and Herr, T.
A primer on the proliferation of offensive cyber capabilities
Atlantic Council 2021, In-Depth Research & Reports

In this chapter, we provide the background for this thesis. We look at the life-cycle of cyber operations under the Access-as-a-Service (AaaS) threat model to conduct an exploratory analysis of the underground markets landscape. In particular, we preliminarily investigate how and to what extent criminal marketplaces support the proliferation of offensive cyber capabilities, and we look at their role in the execution of offensive cyber operations. To achieve this, we break down offensive cyber operations into five pillars. The pillars are the following: (1) *vulnerability research and exploit development*, (2) *malware payload generation*, (3) *technical command and control*, (4) *operational management*, and (5) *training and support*. These pillars are inspired by the tactics indicated in the MITRE's Enterprise ATT&CK Framework [64], which we adapted to frame with greater granularity the categories of offensive capabilities available from marketplaces. The sets of offensive capabilities (products or services) indicated by each pillar are needed to support (to a different degree) a stage of an offensive cyber operation. Not all offensive cyber operations rely on tools and services from all the identified pillars. In this chapter, we focus on offensive operations under the AaaS threat model, as they represent an example helping our investigation that includes cyber capabilities from all five pillars.

To illustrate the differences among markets, we rely on a review of the relevant literature, industry news, media outlets, and governmental reports. Said markets vary in segregation and maturity, thus affecting the quality of their offer and making their provisioning suitable for different purposes and threat actors. From our preliminary analysis, we note that the level of segregation (as opposed to the 'openness' of markets to new members) is positively correlated to fluctuations in the quality and sophistication of the offered products.

These variations suggest that differences across markets are relevant and that looking only at markets with similar characteristics does not imply that the observation is representative of the whole cybercriminal landscape. Therefore, in this chapter, we delve into the ecosystem provisioning offensive capabilities to threat actors, and we discuss our findings in the context of Access-as-a-Service with a focus on the underground markets space.

2.1. Introduction

The proliferation of offensive cyber capabilities (OCC) has created the opportunity for a number of state and non-state actors to engage in offensive cyber operations (OCO). Over time, the barrier to entry in this domain has become more of a gradual rise than a steep cliff, and this slope is expected to only flatten increasingly over time [232]. As states and non-state actors gain access to more and better OCC, and the in-domain incentives to use them increase [95], the instability of cyberspace grows. In fact, said actors use OCC for different purposes: some seek economic gain, others conduct offensive cyber operations to disrupt critical infrastructure [106] for military purposes or to cause political unrest, while others aim at maintaining national security by fighting terrorism[1]. To pursue such goals, an option for adversaries is to rely on OCC that operate under the Access-as-a-Service (AaaS) threat model. AaaS is a threat model that enables attackers to directly obtain access to a targeted system (e.g., a network, a personal device, ...). Depending on the maturity of the supplier and clearance the customer (i.e., the threat actor, the attacker) has, these OCC greatly vary in sophistication and maturity, ranging from general purpose components to complete, operationalized off-the-shelf espionage solutions defying state-of-the-art technology.

Considering the wide spectrum of services supporting AaaS, we investigate their maturity in relation to the characteristics of their providers, ranging from unsophisticated underground marketplaces to state actors. To investigate this matter, we rely on the relevant scientific literature, industry news, media outlets, and governmental reports.

2.1.1. Offensive cyber capabilities: Seeing the whole chain

Offensive cyber operations are possible with the use of a diverse set of offensive cyber capabilities. Malware and exploits represent only some of the pieces that compose the jigsaw of an offensive operation. Emblematic is the case of Stuxnet, a malware[2] attributed to Israel

[1] Although this definition is prone to interpretations used to justify the haunting of human rights activists and journalists critical to the government [175]

[2] For the sake of precision, Stuxnet is a worm, a malicious software that does not require any interaction from the user, like opening a file, and has self-propagation capabilities

and the United States, used in one of the first documented state-backed offensive cyber operations targeting critical infrastructure. Stuxnet was designed to infect the PLCs controlling the centrifuges used to enrich uranium in the nuclear program of the Iranian government. The malware had the ultimate goal of destroying the centrifuges by briefly operating them at a faster speed than their nominal values, without raising any alert. The development of Stuxnet was especially well executed and required extensive collaboration between the involved parties. On the technical side, its developers weaponized Stuxnet with five different 0day exploits[3], tailored to infect the target systems. Post-mortem analysis of the malware shows that Stuxnet included command and control capabilities to push new updates over time. This exemplifies that, to accurately frame OCC, it is necessary to frame them as a chain of capabilities that contributes to the realization of an offensive cyber operation.

Therefore, we propose a taxonomy of offensive cyber capabilities based on the tactics from MITRE's Enterprise ATT&CK Framework to characterize their function within the context of an offensive cyber operation. The five pillars are (1) Vulnerability Research and Exploit Development, (2) Malware Payload Development, (3) Technical Command and Control, (4) Operational Management, and (5) Training and Support. Table 2.1 summarizes these pillars.

2.1.2. Semi- and self-regulated markets for OCC proliferation

Providers and developers of OCC can be roughly separated into self-regulated and semi-regulated spaces. Both spaces provide access to different technologies, such as malware, supporting infrastructure, and services. The two environments operate respectively in lack or presence of legal jurisdiction, and they are capable of offering capabilities of different maturity and sophistication. The self-regulated space is typically represented by underground marketplaces, which operate autonomously and offer illicit products and services to their customers, whereas the semi-regulated space features governments and private sector companies offering cutting-edge espionage services to vetted customers.

The self-regulated environment is characterized by the presence of marketplaces with different levels of segregation. The majority of these markets are freely accessible, allowing wannabe hackers and unskilled actors to join; pull-in marketplaces represent another conspicuous fraction of the overall picture, which verify and screen their participants; a minority consists of segregated marketplaces, criminal venues where only highly-skilled actors have access [151, 16]. These differences in the access model help marketplaces to identify their market segment; committed criminals with verifiable credentials and experience obtain clearance to more elite and segregated marketplaces, thus contributing to the introduction of innovation and mature attack capabilities [16]. In addition to this distinction, marketplaces in the self-regulated space largely differ in terms of the offensive cyber capabilities they offer. For example, `0day.today` is a marketplace operating in the clearweb selling exploits targeting multiple software and operating systems, although there is anecdotal evidence showing customers attempting to buy a high-priced 0day becoming victims of scam [265]. On the other hand, marketplaces like `exploit.in` and `darkc0de` that operate(d) in the underground are regarded as exclusive and well-regulated [83, 177, 153], capable of facilitating

[3]A zero-day (or 0day) is a vulnerability that is currently unknown to the software vendor and the organization whose system the vulnerability affects, and for which a patch does not exist.

Table 2.1: The Five Pillars of Offensive Cyber Capability Proliferation

	Definition	Government Examples	Criminal Examples	Industry Examples	AaaS Examples
Vulnerability Research and Exploit Development	Discovered vulnerabilities, or disclosure programs that facilitate the proliferation of discovered vulnerabilities and written exploits	Chinese intelligence community vulnerability research and exploitation, specifically within the MSS and its associated CNNVD	Exploit kits sold on underground forums	Bug bounty programs, vulnerability disclosures, Zerodium	NSO Group's use of a WhatsApp 0day
Malware Payload Development	Any malware or tool written or used by attackers to conduct offensive cyber operations, or any forum that encourages or conducts exchange of malware	Custom malware developed by state teams that is reverse engineered and published by malware analysts	Commercial malware market	Red-team tools developed and sold through commercial offerings and companies; posting malware for research on GitHub	NSO Group's Pegasus Spyware
Technical Command and Control	Technologies aimed at supporting offensive cyber operations, e.g., bulletproof hosting, domain name registration, server side command-and-control software, VPN services, or delivery accounts involved with the initial creation of an offensive cyber operation	IPs and domains attributed to state operations by threat intelligence reports	Bulletproof hosting, and other pre-built command-and-control infrastructure	Test servers built to send phishing tests against one's own companies, infrastructure used for penetration testing purposes	Infrastructure used by Appin Security for Operation Hangover
Operational Management	Operations management, strategic organization of resources and teams, initial targeting decisions, and other functions that are required to effectively manage an organization that conducts cyber operations	Chain of command within an organization of government intelligence agencies	Criminal outsourcing, ransomware affiliate programs	Delegation of duties within a red-team exercise; escalation policies during an incident	Good Harbor Consulting's organizational management of UAE DREAD cyber capabilities
Training and Support	Training or education provided on the offensive cyber operation process, expanding the number of trained professionals and creating connections between them that facilitate the growth of OCC	NSA's National Cryptologic School or other government-sponsored cyber training program	Fraud tutorials, phishing kits, customer support, provided within forums	Kali Linux tutorials on YouTube, cyber security certifications, conference training and talks	DarkMatter training provided to UAE cyber operators

trade between committed threat actors and expelling untrustworthy members from their platforms. These differences appear to stem from the capability of a marketplace to provide a regulated environment within which trade happens; being able to establish a clear set of rules and enforce them builds trust among participants, positively affecting the average quality of traded products.

In the semi-regulated space, private sector and governmental agencies benefit from ample access to funds, which allow them to produce high-quality and fully-fledged versions of products similar to those offered by criminals, or to produce them in-house. When a governmental agency is not able to access the needed know-how, the private sector can supply them with the necessary capabilities, ranging from the mere espionage tool, possibly accompanied by personnel training, to fully-fledged AaaS operations. Among these companies, a notorious example is the Israeli 'NSO Group', a company that 'creates technology that helps government agencies prevent and investigate terrorism and crime to save thousands of lives around the globe'[4] [195].

The next sections develop a more detailed picture of the markets in which these transactions take place and describe the five pillars of this chain of OCC.

2.2. The five pillars of Offensive Cyber Capability Proliferation

This section will look in greater detail at each pillar and into the availability of the associated capabilities with regard to both the self- and semi-regulated spaces. For each pillar, we present a short table summarizing their availability across the two spaces. We indicate with a dash (-) no capabilities for that pillar from that marketplace, ○ to indicate that actors have only basic capabilities on that pillar (e.g., obtained by operating automated frameworks), ◐ when actors can repurpose and modify existing technologies in that dimension (e.g., to obfuscate known malware/exploit code), and ● to indicate actors that can generate novel methods or efforts in that dimension (e.g., 0day exploits).

2.2.1. Pillar one: Vulnerability Research and Exploit Development

	Self-regulated space (Black markets)			Semi-regulated space	
	Free	Pull-in	Segregated	Private - AaaS	Government
Vulnerability Research and Exploit Development	-	○	◐	●	●
	-	◐	●	●	●

To operate, digital systems run operating systems and a multitude of software. These components can be affected by vulnerabilities, flaws that attackers exploit to obtain unauthorized

[4]This claim is largely disputed due to the use of their products to conduct attacks against human rights activists and journalists in various countries [175].

access, extract sensitive data, and compromise their correct operations [234]. To achieve this, attackers write exploits, code that leverages said vulnerabilities to achieve their malicious intents [90].

Security researchers put their efforts into finding and responsibly disclosing vulnerabilities to software vendors, allowing the latter to produce security updates to fix their software. However, threat actors can pursue the same path to find vulnerabilities to use to their own advantage. An undisclosed vulnerability for which an exploit exists takes the name of 0day. A 0day is a particularly powerful type of exploit, as there is no fix available at the time of exploitation and the affected systems are unprotected against it [206]. As such, 0days represent an unmatched offensive capability against a vulnerable system, but exploits for known vulnerabilities remain usable even after a patch has been released (known as ndays). Albeit less effective, the latter composes the bulk of all exploits used in the wild [140].

Self-regulated space. Underground markets often group and retail multiple exploits in exploit kits, both for rental or on license, with prices ranging from a few hundred to a few thousand dollars, depending on the number of exploits included and their reliability [160]. However, this seems to be the case only in portions of the self-regulated space; novel and effective exploits are almost non-existent in free-access underground marketplaces [202]. These markets have been reported to have a conspicuous fraction of scammers among their members, whose activity causes quality uncertainty and hinders the trade of effective technological products [122, 274]. In rare cases, we witness some discussion on vulnerability research, but it generally revolves around the repackaging of old technology [15]. In reaction to that, pull-in marketplaces scrutinize new members, causing these markets to be populated by more knowledgeable and trustworthy threat actors [274, 76], which are able to offer some effective attack technology in some cases, especially in more segregated sections of these markets, thanks to the presence of trust-enabling mechanisms that support trade [16]. In some cases, pull-in markets have proved able to supply new malware payload generation techniques and to make progress in the management of more complex command-and-control architectures [Campobasso7]. Segregated marketplaces instead present mature and reliable trust mechanisms, often corroborated by the presence of members that have bonds offline [172, 166], thus creating deeper trust among peers and potentially favoring a more fertile environment for discussion and trade.

Semi-regulated space. Following Edward Snowden's 2013 disclosure on the NSA's mass surveillance operations [112] and declarations on the NSA 0day stockpiling [29], there has been growing evidence that these tactics do not belong only the cybercriminal world, but are largely employed from multiple countries worldwide to conduct cyberwarfare (North Korea [253], China [53], Iran [24], the United Arab Emirates (UAE) [180], South Korea [149], the United States [212], and multiple other countries [180]), *de facto* indicating the existence a shift in the conflict to the intangible cyberspace. Similarly, companies like Candiru and NSO Group providing espionage and counter-terrorism services to governments rely on vulnerability research to identify unknown vulnerabilities and craft 0day exploits used in their espionage software [250, 54].

2.2.2. Pillar two: Malware Payload Development

	Self-regulated space (Black markets)			Semi-regulated space	
	Free	Pull-in	Segregated	Private - AaaS	Government
Malware Payload Development	○	◐	◐	●	●

Malware is usually the most discussed component of an offensive operation. Malware is quite diverse; it includes but is not limited to, payloads generated by penetration testing toolkits (e.g., Metasploit's MSFVenom and Cobalt Strike) [230], disruptive ransomware, and infostealer malware licensed as commercial software, and state of the art espionage tools using 0days. However, malware becomes obsolete quickly: exploited vulnerabilities get patched, and antivirus software detects them. This led to the emergence of practices and tools that help improve the lifecycle of malware. Aside from the many creative techniques used during the delivery of a malicious payload, attackers develop and use tools to obfuscate malware in the eyes of an analyst or antivirus software. This is achieved by mingling with the code executed by the program without altering its functionality.

Self-regulated space. Malware traded in the self-regulated space largely varies in quality. Free-access marketplaces feature sometimes some malware, but it might be obsolete or mainstream; sometimes, with the use of an obfuscation layer, this malware could be used for some scenarios [40]. In pull-in marketplaces, it is not uncommon to find advertisements of vendors providing more professional information about the malware's capabilities, how effectively it evades detection, and other technical details on its features, alongside some screenshots [15, 109]. These markets sometimes feature some malware on a license basis (Malware-as-a-Service) [197, 76]. From our observation, more segregated markets are quite on the same line. The major differences mainly stem from the increased amount of details in the description of the offer and the potential presence of malware advertised for 'exclusive' trade to a limited number of players (and sold at a higher price). In other cases, we witness the opposite, where some members with considerable reputation offer malware but without providing details in public. Considering the value of these transactions, more mature markets employ escrowing to conduct these transactions, where forum administrators act as guarantors for both parties in a transaction [35]. Furthermore, these markets can be particularly harsh against players who do not closely abide by the market rules; selling software that does not match the product description often results in a ban from the platform, causing reputation and financial damage to the offender [16, 274, 171].

Semi-regulated space. When considering the semi-regulated space, malware becomes more tailored and effective, thanks to the presence of particularly resourceful customers: governments. In the case of governments with not fully mature offensive cyber capabilities, they rely on the private sector to conduct their offensive operations; for example, industry suggests that malware similarity among multiple Chinese APTs could indicate that they rely on an external organization that the Chinese government uses to outsource part of the malware development process [174]. Otherwise, nation-states with ample funds and capabilities are able to craft their own malicious payloads; as discovered with Vault 7 by the Shadow Brokers and additional shared information from Wikileaks in 2017, NSA and CIA were developing and stockpiling their own exploits [97, 101]. To have an idea of how impacting this mal-

ware was, one of these exploits, EternalBlue, was weaponized into the infamous WannaCry Ransomware, resulting in an attack of unprecedented scale according to the EUROPOL [31].

2.2.3. Pillar three: Technical Command and Control

	Self-regulated space (Black markets)			Semi-regulated space	
	Free	Pull-in	Segregated	Private - AaaS	Government
Technical Command and Control	○	◐	◐	◐	●

To execute an OCO, malware alone is not sufficient. Most of the time, it is necessary to communicate with the malware to achieve data exfiltration, to deploy additional modules, or to adjust its configuration. The involved infrastructure is not limited to a command-and-control server but includes delivery mechanisms bonded to phishing pages and malvertisement. Command-and-control servers are critical in the execution of an OCO, as they represent a (theoretically) safe gateway for the attackers to monitor the status of the operation, and to collect extracted data. These activities are monitored by cloud providers, which quickly deactivate these servers and may report these activities to law enforcement agencies. Hence, bulletproof hosting is the often go-to solution. Providers of bulletproof hosting services operate servers in lenient jurisdictions against subpoenas or requests for action from foreign law enforcement agencies [193] or even build them in particularly hardened locations (a remarkable example is the Cyberbunker in The Netherlands, a cold-war era NATO bunker, used to host all sorts of illicit content – shutdown in 2019 [102]), often operated by criminals themselves [150]. There is evidence that some of these services rely on compromised websites and infected systems as a cheaper alternative [110]. In addition to hosting, there are a number of services that offer VPN, VNC, and SSH access to compromised hosts, which can be used to increase the anonymity of miscreants.

Self-regulated space. These services are present in the self-regulated space and do not appear evenly distributed across. For example, bulletproof hosting is rare in free-access markets, and it appears more often advertised in pull-in markets in the Russian space [136]. Instead, VPN, socks proxies, and SSH accesses can be frequently found in both free and pull-in access marketplaces [154]. Alongside these, in pull-in markets, there is a growing interest in fraud services to spread malware (e.g., 'traffers' use a combination of compromised hosts, SEO, and advertisement to make the malicious software appear within the top results of popular search engines) [214]. Albeit limited information is available, providers of bulletproof hosting and proxy services in segregated markets should not remarkably differ from what it is available for pull-in markets.

Semi-regulated space. In the semi-regulated space, private actors and governments are able to deploy their command-and-control globally thanks to greater funds. For example, in 2019, a threat actor based in UAE attempted to become a certificate authority, thus becoming able to potentially sign certificates and software to distribute in an offensive operation [37].

2.2.4. Pillar four: Operational Management

	Self-regulated space (Black markets)			Semi-regulated space	
	Free	Pull-in	Segregated	Private - AaaS	Government
Operational Management	-	◐	◐	●	●

Performing offensive operations requires coordination and planning. This is made possible by the existence of social structures within criminal organizations that establish processes, identify suppliers, and set goals for the success of the operation.

Self-regulated space. Organization is fundamental even in small criminal ventures. The unsealed US Department of Justice indictment of Andrey Turchin [75], a member of cybercriminal group FXMSP [203], revealed how the criminal organization executed their frauds: part of the job was to identify potential targets via open RDP scanning, phishing and bruteforce attacks, then to perform lateral movement. Once a stable foothold was gained, these compromised systems were advertised and sold in different underground markets.

With the outbreak of the Russian-Ukrainian conflict in February 2022, a pro-Ukraine employee of the large and successful Conti ransomware gang disclosed the TTPs and conversations between members of the gang, allowing security researchers to reconstruct their organization hierarchy [94]. It emerges that the organization managed different departments conducting the technical side of the operation, and operated with the support of managerial figures like human resources, public relations, and training figures. From our observation, we learn that groups and services that organize and manage complexity to that point are on average more mature than others, and we monitored their presence mostly in pull-in and segregated marketplaces.

Semi-regulated space. The managerial complexity scales up for state-sponsored offensive operations. These attacks require intelligence gathering and precise framing of the targets. Conducting these managerial operations requires years of experience, and tested procedures to achieve their success. Of course, these aspects of offensive operations can be performed with the assistance of the private sector; for example, the management and structure of UAE's cyber surveillance organization Development Research Exploitation and Analysis Department were supported by Good Harbor Consulting [226].

2.2.5. Pillar five: Training and Support

	Self-regulated space (Black markets)			Semi-regulated space	
	Free	Pull-in	Segregated	Private - AaaS	Government
Training and Support	○	◐	◐	●	●

To achieve their goals, threat actors need to rely on skilled personnel or provide them with training to succeed. This pillar is crucial within the context of offensive cyber operations that include multiple threat actors; new employees need to be instructed with the *modi operandi* of the organization, trained, and overseen.

Self-regulated space. In the context of the self-regulated space, training generally comes in the form of tutorials, fraud schemes that show how to cash out stolen credit cards for Bitcoins [76], or how to use off-the-shelf malware available in the same venues [135]. In the case of free-access markets, these tutorials are often outdated and do not pose a significant threat; in more segregated marketplaces, these can be offered alongside material like phishing and exploit kits that attackers can readily deploy to perform attacks. In rare cases, marketplaces can even feature detailed tutorials left by experienced threat actors retiring from business.

Semi-regulated space. The private sector is not an exception in that case as well; the notorious penetration testing framework Cobalt Strike frequently used by APTs [173] hosts free video tutorials [58] on how to use the tool. Organizations in the private sector offering training for governmental agencies provide more than just technical expertise. For example, in the leaks of Hacking Team, a company offering surveillance software, it emerges that the company offered training sessions to Ethiopia's Network Information Security Agency and Sudan's National Intelligence and Security Service to use their offensive capabilities in 2014. Governments instead can invest in education to train potential contractors; for example, the National Security Agency's National Cryptologic School [190], has historically been useful in developing tailored expertise.

2.3. Chapter conclusion

Understanding what criminal markets, governmental agencies, and private AaaS groups offer and how they build state-of-the-art products for conducting offensive cyber operations is a powerful means for stakeholders, policy-makers, and law enforcement to protect their assets and defend the society at large. From this overview, it emerges that OCC providers operating in the self-regulated space are diverse, offering offensive capabilities to threat actors with different degrees of sophistication. These observations suggest that not every underground marketplace should be considered the same way; on the contrary, their differences should be recognized as indicators of different levels of threats and accounted for.

3

A method for stealth monitoring of underground markets

This chapter is based on [Campobasso8]:
M. Campobasso, P. Burda, and L. Allodi
CARONTE: a Crawler for Adversarial Resources Over Non-Trusted, high-profile Environments
2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW) - 1st Workshop on Attackers and Cyber-Crime Operations (WACCO)

From Chapter 2, it emerges that underground markets are diverse in terms of segregation and offer, and products' quality is positively correlated to the markets' level of segregation. To further investigate this relation, it is necessary to study a wide range of underground communities, while keeping our foothold in the markets stable. Especially for segregated markets, obtaining access is non-trivial and may require lengthy infiltration processes that raise the associated costs for investigations. Obtaining access is not the only problem; from literature and first-hand experience, we learn that especially high-profile communities perform network monitoring to detect crawler activity and ban the associated accounts, limiting the activity of researchers and law enforcement operations. In this chapter, we discuss some of the known techniques that marketplaces could employ to identify crawlers and obstruct data extraction. We postulate that using crawlers approximating human behavior could greatly reduce their detection rate. Thus, we devise a general method for stealth underground market monitoring that models human behavior. We experimentally evaluate its efficacy with the implementation of a proof-of-concept, CARONTE, that simulates human behavior during its activity while using an instrumented browser for navigation. We compare its performance against state-of-the-art crawlers and human behavior (Amazon Mechanical Turks) with positive results. Finally, we experiment with a general algorithm to assist researchers in easily creating tailored instances of the crawler with a supervised procedure.

Link to CARONTE:
`https://github.com/michelecampobasso/caronte-crawler`

3.1. Introduction

Underground cybercrime communities are increasingly more important to understanding and measuring the overall threat landscape. Security operators or security service providers scrape them to obtain key intelligence on emerging threats [43]; law enforcement (LE) scrape (and sometimes run) them to monitor cybercrime operations and networks [89]; security researchers are interested, among other things, in understanding the dynamics of attack innovation [15, 21], identify key actors operating in these forums [207], or investigate novel or emergent threats [134, Campobasso7]. Currently, monitoring activities focus on the rapid collection of massive amounts of data [238], which can then be used to train machine learning (ML) models to, for example, extend available parsing capabilities to different forums or underground communities. Indeed, the proliferation of underground criminal communities makes the scalability of monitoring capabilities an essential aspect of an effective, and extensive, data collection, and ML has been the clear go-to solution to enable this. However, this comes at the high price of having to collect large volumes of data for training, raising the visibility of the researcher's activity and interest in the criminal community.

Our previous efforts in Chapter 2 and the scientific literature showed that not all communities are born the same [122]; on the contrary, the majority of underground communities appear largely uninteresting (even when generating massive amounts of data about alleged artifacts [262]), both in terms of economics and social aspects [16, 274], as well as in terms of (negative) externalities for society at large [238, 191]. Whereas there are only a limited number of 'interesting' communities to monitor, gaining access to these may be less than trivial in many cases, particularly for forum-based communities and markets [274, 15]: high entry costs in terms of entry fees, background checks, interviews, or pull-in mechanisms are becoming more and more adopted in the underground as a means to control or limit the influence of 'untrusted' players in the community [274, 15]. Under these circumstances, researchers and LE infiltrating underground communities may face significant opportunity costs whereby increasing monitoring activity may also jeopardize their ability to monitor the very community(-ies) in which they wish to remain undercover: network logs and navigation patterns of crawling tools (authenticated in the communities using the researcher's credentials) can put the real nature of that user's visits under the spotlight, and lead to blacklisting or banning (sometimes hard to obtain) related access credentials [209, Campobasso7, Campobasso8]. This is particularly undesirable in high-profile communities where the cost of re-entry can be high and makes the endeavor of monitoring cybercrime communities time consuming, technically challenging, and expensive to run. A large fraction of this overall 'cost' is constituted by the need to build *ad-hoc* crawlers and parsers capable of correctly navigating different forums, and extracting relevant content, while remaining under the radar [209].

Anecdotal evidence shows that monitoring incoming traffic, for example for robot detection or source-IP checking, is a countermeasure that underground communities may employ to limit undesired behavior. Some communities explicitly acknowledge the adopted countermeasures (see for example Figure 3.1), others explicitly state that they are aware of

Figure 3.1: Example of inbound traffic monitoring from criminal communities

the monitoring operations of LE and other 'undesirable' users; for example, the administrator of one prominent underground forum for malware and cyber-attacks that the authors are monitoring, states explicitly: *'Forums like this are being parsed by special services and automatically transfer requests to social network accounts and e-mails.'* This significantly inhibits researchers' ability to build scalable, reusable parsing modules, as the collection of large amounts of data to train the associated ML algorithms may be slow or carry significant risks of exclusions from the monitored communities. Pastrana et al. [209] lead the way in identifying *stealthiness* as a requirement for systematic underground resource crawlers, with many recent works not explicitly mentioning these aspects [217, 158].

In this chapter, we present a general method for stealthy crawling of underground markets via human behavior modeling, and we experimentally evaluate it by implementing a proof-of-concept crawler named CARONTE. In particular, our method aims at providing a simple user model to mimic human behavior on a webpage, to keep a low profile while performing the data collection. In the same context, we experiment with a procedure that allows a hypothetical tool to semi-automatically learn virtually any forum structure, without the need to write ad-hoc parsers or collect and manually classify large volumes of data. Finally, we evaluate the method by means of a proof-of-concept named CARONTE against four underground forums and compare the network traffic it generates (as seen from the adversary's position, i.e., the underground community's server) against state-of-the-art tools for web-crawling. Our results clearly show that both CARONTE's request patterns as well as the completeness of the downloaded resources page are significantly closer to humans when compared to other state-of-the-art crawling tools.

This chapter proceeds as follows. In Section 3.2 we discuss relevant background and related work; Section 3.3 presents the proposed method and the design of the proof-of-concept, whereas Section 3.4 presents the experimental validation and results. Section 3.5 deepens the impact and limitations of CARONTE, and Section 3.6 concludes the chapter.

3.2. Background

Cybercrime monitoring has mainly been implemented through *ad-hoc* tools to scrape adversarial platforms that do not scale up with the number of sources and the variety of the content; nonetheless, most of these efforts were more concerned with developing techniques that enable underground economy discovery, key hacker identification [1] and threat detection [32], disregarding stealthiness in favor of parsing volumes [158, 217], with few notable exceptions [209]. In this section, we discuss the changing threat model against which these solutions are used and the technical means by which their use can be detected by adversaries.

3.2.1. A changing threat model for cybercrime monitoring

Cybercriminals have demonstrated to be increasingly aware of the mounting interest from scientific and nation-state-sponsored investigations, pushing them to start developing techniques to avoid unwanted actors and data gathering in their communities [198]. Retaliation activities on the side of the cybercrooks have made the news in the past [274, 15, 16], including threatening journalists[1] and academic researchers [123]. The attention posed by cybercriminals to the public sphere is also reflected in the technology and administrative procedures they employ to detect or stop possible 'intrusions' in their communities. Part of these techniques are centered on the evaluation of a prospective member at the act of registration on these platforms [274, 16], or on the continuous monitoring of community participation by each member [15]. Similarly, recent evidence suggests cybercriminals may be monitoring and auditing the traffic on the web servers, both to prevent access from undesired IP ranges (e.g., see Figure 3.1) and to mitigate external threats such as denial of service attacks. These technologies identify patterns and anomalies in network traffic to detect undesired activities, requests generated by robots and crawlers and, if necessary, take action to limit those [219, 91]. Whereas researchers can build profiles to go undercover in certain communities, thereby passing the first access filter enforced by cybercrime community entry regulations [15], network monitoring operations remain an unmitigated threat to automation of data collection once inside the forum, particularly at the face of an evolving adversary (i.e., the cybercrooks behind the platform).

3.2.2. Crawler detection techniques

Recently, crawling has become a conspicuous portion of Internet traffic [27] and an unwanted practice from website owners, due both to network resource consumption and to the lack of explicit permission for a third party to massively download all the website content for unknown goals, often resulting in a privacy violation [276, 96]. Several anti-crawling techniques have been developed, which rely on two different families of anomalies that could be detected in the generated traffic: technical, and behavioral anomalies.

[1] A man accused of trying to frame a blogger with heroin is in big trouble, Business Insider. Visited: April 2019. https://www.businessinsider.com/brian-krebs-heroin-threat-hacker-extradited-2015-10/?international=true&r=US&IR=T

Technical anomalies

Among the first, crawlers generally can be detected from simple anomalies in the headers of HTTP requests, such as the presence of bogus user agents [276, 224], frequent use of HEAD HTTP requests [224], and the lack of a referrer [240]; in addition to these, other anomalies indicating sessions that do not originate from browsers are the lack of cookie management [276, 80] and JavaScript execution [276], malformed requests, and a high number of 404 errors [80, 240]. When multiple of these anomalies are detected in a session, the likelihood that the session is run by command line scripts, commercial crawler solutions, or headless browsers, all of which activities do not involve a real human interacting with the website, is high. Despite their simplicity, these detection strategies are adopted even in modern anti-crawler solutions [55, 56, 50, 183, 200]. However, when considered singularly, these are not able to consistently and reliably detect crawlers; when more of them are monitored, these could still be insufficient in case of a focused and stealthier crawler tampering with information in the requests. In principle, the origin of these anomalies depends on the absence of the rendering engine of a fully-fledged browser, executing and loading dynamic content on a page, together with a consistent management of cookies, and issuing requests like a regular browser would do. In addition, some other anomalies depend more on the nature of the crawler itself; in fact, the goal of a crawler is to extract the most data in the least possible time, thus leveraging on techniques that minimize the 'refetch' of unchanged resources (HTTP HEAD), while building request queues as new links are extracted from pages that are executed without the reference to their page of origin (lack of referral headers).

Behavioral anomalies

For what concerns the behavioral anomalies, crawlers access the *robots.txt* file [224], present high fetch rates [276, 221, 28], show higher text/media ratios compared to humans (i.e., crawlers disregard media content in favor of text) [155, 79, 80], and perform lengthy crawling sessions [79, 28], including working overnight [28]. Apart from detecting these anomalies, additional efforts have been made to create more reliable crawler detection methods; the state-of-art techniques for detecting automated activity on a website include pattern recognition, like loopholes detection and breadth first or depth first strategies, JavaScript fingerprinting and tracking, and Turing tests, on top of other strategies [155, 79, 263, 221]. Nonetheless, the resolution of Turing tests like CAPTCHAs can be outsourced at remarkably low prices [218] or solved via OCR [148], and the production of non-suspicious traffic can be obtained with a focused crawler that acts with some precautions.

3.2.3. Modeling 'regular' user behavior

Apart from the technical and behavioral anomalies discussed in the previous section, we consider the characteristics exhibited by humans when navigating a website. By doing this, we do not simply aim at circumventing the illustrated crawling detection techniques, but we plan to step further into differentiating our tool from bots by modeling the behavior of 'regular' users. Studies on user browsing behavior broadly distinguish between *click patterns* and *time patterns*.

Click patterns

Click models are used to evaluate user decisions in considering a topic or hyperlink relevant to the specific purpose of their navigation or query [85]. Derived approaches consider *single-browsing* and *multi-browsing* models to infer user behavior as a function of the purpose of the navigation, in particular distinguishing between *navigational* and *informational* queries, whereby the user wants to reach a specific resource, or is interested in exploring new information, likely producing respectively one or multiple clicks at a time [85, 114]. These models show that past behavior or user interest are useful predictors of which clicks will happen in the future [85]. In our context, forum-browsing clearly covers both dimensions, depending on whether the user aims at retrieving specific information (e.g., updates in a thread of previous interest to the user), or to explore the content of a forum section.

Time patterns

More broadly, these dynamics are explained in the information retrieval literature as dependent on the user's task [26]. The decision of a user to click on a specific resource depends on its perceived and intrinsic relevance with regards to the user's goal and is bounded by how many topics need to be opened to find the answer to the information the user is interested in [84]. Post-click user behavior (i.e., what the user does once they reach the clicked resource) has been shown to be directly related to the relevance of the document [115]. Post-click behavior includes variables associated with mouse movements, scrolling, and eye-tracking [115, 85], clearly showing that what the user does, and how much time the user spends on a webpage, varies as a function of the relevance of the webpage. Indeed, a user's *inaction* on a webpage has been shown to be relevant to modeling the quality of dynamic systems such as recommendation systems [279]. Part of that behavior can be quantified by considering how quickly users can be expected to process the relevant information [179]. Data around this subject is scarce and quite diverse; some sources refer to the average reading speed to be around 200-250WPM (Words Per Minute) with a comprehension rate of 50/60%[179], others report that for reading some technical content with a good proficiency, the speed can be around 50-60WPM.

3.3. The proposed method: CARONTE

Following the discussion in Section 3.2.2, we define a method to counter the analyzed crawler detection techniques. To achieve this, we identify a set of characteristics as desiderata that will guide us in the definition of the architecture of CARONTE's proof-of-concept. In this chapter, and specifically in this section, we focus on characterizing the proposed method, rather than exploring the technical architecture of CARONTE's proof-of-concept, whose details are provided for its full implementation (named `THREAT/crawl`) presented in Chapter 4; the reader interested to the full details of CARONTE's proof-of-concept implementation can refer to the relevant sections in the Appendix or directly to the original paper [Campobasso8].

3.3.1. Method definition

Technical characteristics

To stay under the radar, a crawler must generate indistinguishable HTTP traffic from a regular browser, both (1) in terms of technical characteristics of the requests themselves, and (2) by requesting the same (dynamic and multimedia) content that a regular would request.

Functional and behavioral characteristics

To covertly extract data from guarded underground communities, a hypothetical tool should be able to act in disguise, granting its user a stable foothold in the target community without being banned. In addition, it would be desirable for the tool to allow a user with no technical expertise to use it. Therefore, a stealthy crawler should be able to: (1) diverge from crawler behavior and, where possible, *mimic human behavior*, by (1a) 'showing interest' for specific areas of forums taught during the training phase like a potential human user would do, (1b) issuing HTTP requests at human speeds, quickly accessing resources already inspected and taking more time on pages with new content in relation to its length, and (1c) by generating stochastic interruptions of the crawling activity and abiding to a user-defined schedule of activity during the week that reflects the needs of a hypothetical human. Apart from stealth, the hypothetical tool should be able to extract data from the pages in a convenient way; to achieve this, the crawler needs to be instructed on what resources should be collected and how to navigate through the forum, using the minimum amount of training data through a guided procedure. Therefore, the tool should (2) semi-automatically learn forum structures to navigate them and parse their content without the need for extensive pre-collected datasets on which to train automated models [217].

3.3.2. Strategy to counter crawler detection techniques

To satisfy the technical characteristics, we propose to use an instrumented session of TOR Browser as the crawler engine. By using a legitimate browser, we completely avoid the problem of generating traffic with abnormal technical characteristics. With regards to the execution of JavaScript, in the context of underground forums disabling it is considered to be a best practice due to privacy concerns; therefore, we can safely disable this feature in the browser without raising any suspicion.

For what concerns the functional and behavioral characteristics, we propose a human behavior model that accounts for these characteristics (e.g., no rapid fetching of resources, navigation of the forum in relation to the volume of displayed content, stochastic crawler interruptions, short crawling sessions over the course of the day, ...). Crucially, these indications set the stage for the definition of a strategy that avoids 'monotonous' and evident robot behavior during the crawler activity[2]. With regards to more advanced crawler detection techniques, like the access to traps and loopholes, the download of only text from a website (ignoring media, styles, and JavaScript), and the incapacity of solving CAPTCHAs,

[2]For more details on the implemented strategy, we refer to the Appendix or to the original paper.

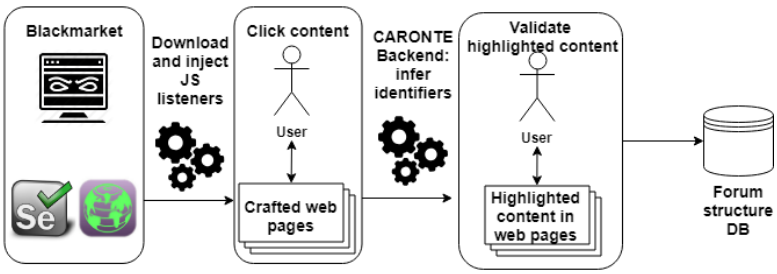


Figure 3.2: CARONTE trainer module structure.

the use of a regular browser comes in help. In fact, browser instrumentation allows us to interact only with the visible and relevant content on a page. By doing so, the crawler will operate without interacting with invisible or fake resources that are hidden on the page, thus ignoring these traps by design, similarly to a regular user (and differently from what a regular crawler would do, by extracting links from the DOM of a page and queuing them for subsequent fetch). Therefore, we design the crawler with a module that informs the crawler about *what* is relevant on a page and *how* to interact with it. Furthermore, using a browser solves the problem of downloading only parts of a website (i.e., the text, at the expense of media content and style sheets), because the browser will issue requests for all the elements needed to render the page correctly. Finally, CAPTCHAs, Turing tests, and browser fingerprinting rely often on JavaScript, which is disabled due to the associated risks (e.g., allowing in the past to bypass completely the anonymization of Tor [244, 231]); at the time of this work, we found no evidence of CAPTCHAs used in cybercriminal forums, whereas their presence has been documented in drug markets [238]. The interface of these markets resembles e-commerce websites like eBay, which greatly differ from forums. Considering our academic interest in cybercriminal products and not in drugs, and the forums being the main venue for their trade, we do not consider drug markets in scope for this work.

3.3.3. Proof-of-concept architecture and implementation

From the described, high-level characteristics, we design CARONTE in a two-tier architecture, separating the *training* from the *crawling* operations. The tool relies on the *tbselenium* library [5], which acts as an interface with the browser automation framework Selenium [130] to access and interact with the target marketplace using Tor Browser.

Trainer module

The trainer module has the task of building a knowledge base for traversing the forum structure. To achieve this, the user needs to provide CARONTE with the required 'sample' pages, which are (at most) five: login page, home page, section page, (optionally) subsection page[3], and a thread page. Nonetheless, identifying these 'sample' pages may not be trivial; to identify robust identifiers, the tool needs to be trained against representative pages that showcase

[3]Some forums may not make use of subsections, making that step of the training unnecessary.

the possible variations that could be featured on a page of that type. During the training procedure, the trainer module will render the operator a downloaded copy of the indicated pages, starting from the login page, and asking to click both the desired content and the necessary navigational items (e.g., login button, next page button, ...). In this way, the trainer will proceed to calculate the identifiers that locate the selected element with multiple techniques, render the result to the operator for confirmation and, in case of a positive outcome, store this information to locate the relevant content for all the pages of this type.

Crawler module

Based on the structural details collected with the trainer module (i.e., where is the relevant content on a page located), the crawler module will traverse the forum to reach the required resources, explore threads, and collect all the required data. The crawler will also embody the requirements of being compliant with the traffic generated from a regular browser via an instrumented session of Tor Browser. Furthermore, it will camouflage its nature by adopting low fetch rates for pages based on the quantity of (unread) text present on the page, compatibly with the average human reading times, and it will perform crawling operations only in certain moments of the day and take random pauses during its activity.

3.4. Experimental Validation

3.4.1. Forum selection

To prove the effectiveness of the proposed method, we implement `CARONTE` and test its capabilities against different forums against four real-world criminal forums built on top of different platforms. The candidates (Table 3.1) correspond to a consistent representation of the most common forum platforms wildly adopted on the Web [1, 32, 217, 238]; as discussed in Section 3.3.2, the selected target forums do not feature JavaScript-based challenges and CAPTCHAs, representing a realistic sample of the target population for `CARONTE`. We first reproduced four live hacker forums by scraping them and hosting their content on a server at our institution. Before reproducing the content on our systems, we inspected the source code and scanned it with `VirusTotal.com` to ensure malicious links or code was not present. Forum mirrors include multimedia content, styles, and JavaScript. To avoid provoking disservice on the server-side while scraping the forums, we avoided aggressive scraping. As our interest is to have an appropriate test-bed to evaluate `CARONTE`'s overall performance, the nature (or quality) of the content of the forums is irrelevant for our purposes.

3.4.2. State-of-art tools selection

To provide a comparison of `CARONTE`'s capabilities against other tools, we select three among the available ones:

- A1 Website Download: shareware crawler specialized in downloading forum content. Through a fine-grained customization wizard, it is possible to use configuration presets that fit better the crawling process against a certain forum software, optimizing its performance;

Table 3.1: Scraped forums for our testbed.

Forum	Time span	Forum software	Obtained with
https://nulled.io	14 Jan 2015 - 06 May 2016	IP Board 3.4.4	Online dump
http://offensivecommunity.net	Jun 2012 - 6 Feb 2019	MyBB (unknown version)	HTTrack 3.49.2
http://darkwebmafias.net	Jun 2017 - 7 Feb 2019	XenForo 1.5	A1 Website Downloader 9
http://garage4hackers.com	Jul 2010 - 4 Feb 2019	vBullettin 4.2.1	A1 Website Downloader 9

- HTTrack: probably the most famous tool for downloading websites, HTTrack provides several features through regular expressions for downloading a website;
- *grab-site*: fully open-source, grab-site is a regular crawler for downloading large portions of the web, powered by the Archive Team.

3.4.3. Training phase

The approach adopted by CARONTE to discover the structure of a forum has proven effective over our tests. In order to get the structure of a forum, we rely on the predictability of the structure of a forum in the future in terms of element locators like XPath and HTML classes. This holds true in the majority of the cases; from the literature analysis and empirical evaluations of the most common forum structures [48, 168, 277], we found no evidence of dynamically-loaded forum structures that would alter the DOM structure at each visit or while being on a page. This seems well in line with environments like the Dark Web, where platform simplicity and functionality, as well as predictability, are desirable [274]. During this phase, all the countermeasures that disable the download and execution of active content and JavaScript are in use as well. As mentioned in Section 3.3.3, for each website it is sufficient to identify at most 5 pages, one of each type, to enable CARONTE to gain complete knowledge of the identifiers necessary to traverse and extract the relevant content from the forums in exam.

Training evaluation

Depending on the peculiarities of the forum against which CARONTE has been trained, different strategies have been adopted to determine the resource identifiers. In particular, as it stands, CARONTE was not capable of detecting an error that would appear later caused by the apparently correct training of `nulled.io`. In this case, we had to 'forcefully' cause the trainer to use a specific resource identification technique to account for variations in the structure of pages not reported in the examined page, for example in the placement of the 'next page' button. Apart from this case, the training on the examined forums concluded positively. For additional details on the problems and solutions, refer to the original paper.

Table 3.2: Experimental features and treatments

#	Exp. variable	Treatment A	Treatment B
readPol	The reader is interested in the content or skims a few posts	Read all the content inside of the thread	Skim thread or read first post
privacy	The user enables or disables JavaScript on Tor Browser	Enabled	Disabled
navPol	Opening resources in parallel or sequentially	Sequential	Parallel

3.4.4. Network patterns and behavior

To evaluate how network traffic generated by CARONTE compares with regards to network traffic generated by humans (i.e., legitimate users) and state-of-the-art crawlers, we performed an experiment employing the Amazon Mechanical Turk platform. This enables us to compare CARONTE against both 'undesirable' and 'desirable' traffic from the perspective of the criminal forum administrator.

Experiment methodology

Human navigation experiment. To generate human traffic to our forums, we rely on Amazon Mechanical Turk (MTurk). From the literature review, we identify three main experimental variables characterizing the habits of a regular user on the Internet:

- **readPol**: The interest raised in the reader by the content may lead them to read carefully all the content of a certain thread or not, resulting in skimming and moving quickly to the next resource [84, 115];
- **privacy**: The desire for privacy of the user, which may be high or low, resulting in the adoption of solutions that prevent JavaScript from being executed or not to avoid fingerprinting techniques [209, 15, 244];
- **navPol**: The propensity of a user to open several resources in parallel before actually browsing them or instead opening them one per time, reading their content first before moving to the next resource [85].

To control for possible interdependencies between these dimensions, we create a 2^{3-1} *fractional factorial experimental design*, that allows us to reduce the number of experimental conditions from eight to four [44]. The experimental treatments and design are reported in Table 3.2, and 3.3 respectively.

Experimental design and setup

An overview of the experimental setup is shown in Figure 3.3. The setup implementation has been carried out in three stages: The selected web forums (ref. Table 3.1) are hosted on an IIS web server (vers. 10) where access logging is enabled. We prepare an Amazon Mechanical Turk task reflecting the experimental design (ref. Table 3.3). The task includes eight questions based on the content of the forum web pages (two multiple-choice questions

Table 3.3: Treatment combination and experiments.

	Exp1		Exp2		Exp3		Exp4	
	A	B	A	B	A	B	A	B
readPol	-	+	-	+	+	-	+	-
privacy	+	-	-	+	+	-	-	+
navPol	+	-	-	+	-	+	+	-

per forum). The task included detailed step-by-step instructions that respondents had to follow. Such instructions serve the purpose of enforcing the treatment in the experiment; for example, *Exp3* requires users to read all content of a thread (**readPol, A**), have JavaScript enabled (**privacy, A**), and open forum tabs in parallel (**navPol, B**):

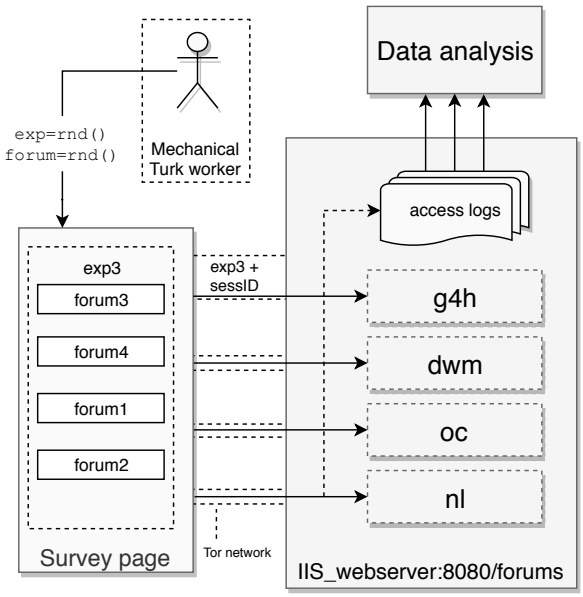
> *[...] open in separate tabs all threads you think are relevant to those two topics (**navPol, B**).*
> *While reading the forum threads, please also skim through to at least the second thread page (**readPol, A**), if present, and even if you already found the answer.*

Notice that, as JavaScript is enabled by default in TOR browser, there is no instruction for **privacy, A**. When a respondent accesses the task on AMT, they are assigned randomly to an experimental condition. Further, each instance of the experiment randomizes the forum order to minimize cross-over effects. The last step consists of enabling us to collect the generated network requests. To avoid limitations imposed by the TOR circuit refresh mechanism[4] that may change the IP address of users every 10 minutes, we set a cookie on the user's browser with a unique session ID. The cookie is used only to distinguish one session from another and we do not collect any personal information associated with the MTurk's account. In addition, in the experiment instructions we inform the participants that no personal information is being collected. Furthermore, the MTurks interact with our servers only via TOR Browser, anonymizing their real IP address. We use the same strategy to track the experimental condition to which the user has been randomly assigned at access time.

3.4.5. Results

Figure 3.4 reports the network analysis for CARONTE compared to the state-of-the-art tools and the MTurks. The goal of this benchmark is two-fold: first, we evaluate the similarity of the generated HTTP traffic between CARONTE and MTurks, and second, we verify whether the technical characteristics of the generated traffic are comparable to those originating from a regular browser (e.g., fetching all the resources needed to render a page such as multimedia and styles, using cache correctly, presence of referrer in HTTP requests headers, etc.). In other words, we aim to evaluate the achieved stealthiness of CARONTE, both in terms of divergence from other crawlers' behavior and similarity to the HTTP traffic generated from a regular browser. To conduct this assessment, we monitor three aspects: *traversing velocity*, in terms of the time elapsing between two requests, *media/text ratio*, as the ratio between

[4] MaxCircuitDirtiness - https://www.torproject.org/docs/tor-manual-dev.html.en



The forums are deployed on an internal system at the university. Resources are accessed by industry standard automated tools (scrapers), CARONTE, and MTurks. All tools access the local resources through the TOR network. Each MTurk is randomly assigned to an experiment setup with different conditions (see Table 3.3). Internal network logs allow us to backtrack user requests to specific experimental setups.

Figure 3.3: Experimental setup

media and text requests, to compare whether CARONTE's activity is comparable to humans, and *thread requests*, as the number of requests issued within a thread. The results are reported in Figure 3.4 respectively for the three mentioned aspects.

For what concerns *traversing velocity*, as emerged from the desired characteristics, CARONTE should not exhibit greed in resource fetching, but should access them with lower frequency, according to the resource content. On one hand, if is true that avoiding request bursts is probably more than enough against an automated monitoring tool, on the other hand, it guarantees an extra layer of stealthiness in case of human verification and could fool ML-based robot-detection systems. Therefore, we compare this traffic throttling model with humans and other tools by monitoring the amount of requests per thread and the time between them. From the comparisons, emerges that the time elapsed between two different requests[5] produced by humans is comparable to CARONTE's and HTTrack's, while the others perform more aggressively. For what concerns the *media/text ratio*, CARONTE together with `grab-site`, perform quite close to humans. Finally, we compare the *thread requests*[6]: CARONTE and `grab-site` perform again better when compared to humans than the other two tools, but their behavior slightly differs from MTurks. Overall, we observe that CARONTE network trace is consistently very similar to human-generated network traffic, whereas other tools are clearly different over one or more dimensions.

[5]A **request** refers to all the calls to a page of a thread, without considering all the linked content downloaded.
[6]**Requests per thread** refer to the set of all the *requests*[5] that have been fired by an actor inside of every thread.

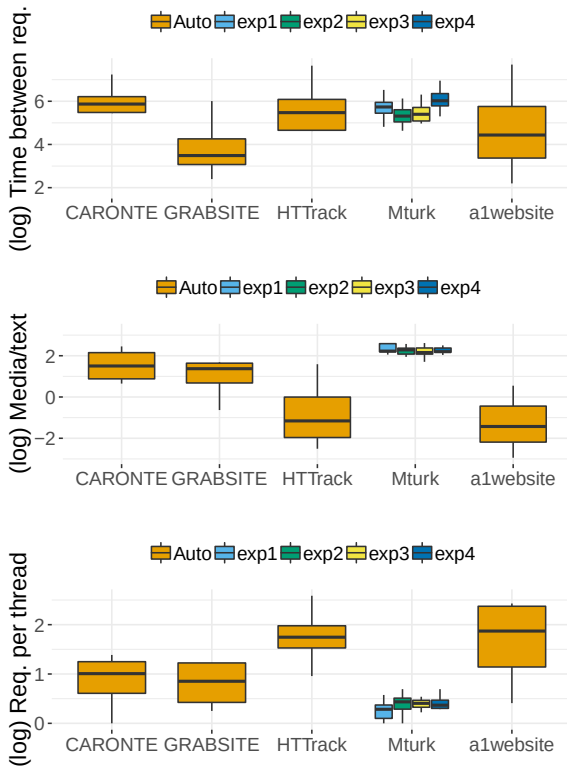


Figure 3.4: Evaluation of CARONTE against state-of-art tools and MTurks

Finally, concerning the technical characteristics of the generated traffic, we do not perform any specific evaluation, as they are enforced conditions by the design of our tool (as in, the use of an instrumented browser nullifies all traffic anomalies caused by command line scripts or headless crawlers).

3.5. Discussion

CARONTE's training module proved effective in flexibly learning diverse forum structures. Differently from ML-based systems, the adopted semi-automated procedure allows the tool to reliably identify relevant structures in the DOM of a page while avoiding entirely the need to collect massive amounts of pre-existent data (for the training and validation) that might jeopardize the researcher activity.

Whereas performing the training procedure does come at the price of additional human-sourced work with regards to fully-automated procedures, CARONTE is meant to be employed over (the few) highly-prominent underground communities where the threat model CARONTE addresses is realistic. The presented proof-of-concept has been tested over four diverse forum structures, and it can be expanded in future work beyond the 'forum' domain (e.g., e-commerce criminal websites).

Table 3.4: Extra features monitored.

Tool	JS	Styles	Cache	Seq./Par.	Referrals
CARONTE	✗	✓	✓	Seq.	✓
grab-site	✓	✓	✗	Par.	✓
HTTrack	✓	✓	✓	Par.	✓
A1Website	✓	✓	✗	Par.	✗

From the network analysis, it emerges that CARONTE reproduces coherently the three investigated features when compared to humans and performs better, on average, than the other tools. Our tool produces the multimedia traffic of a regular human actor, together with grab-site, while the other two tools diverge from this behavior; we suspect that this is to be traced back to some optimization mechanisms that avoid reissuing requests for the same resource, without even issuing a HEAD HTTP request. A regular browser instead will always reissue the request while loading another page, if not explicitly instructed by a server-side caching policy. Nonetheless, we have found no confirmation in the documentation of these tools. With regards to the number of requests generated per thread, there is a noticeable difference when compared to humans. This is probably caused by MTurks skipping some pages in the threads. In fact, in multiple cases, the downloaded forums have plenty of 'useless' replies to threads, which may result in a decreased interest from the reader, possibly leading to skipping the following pages. The observed difference in generated requests per thread between CARONTE and grab-site and the other two tools is caused by the fact that they follow also non-relevant links, such as content re-displacement in the page. In particular, this last behavior represents a well-known traffic feature of a crawler. To improve this, it could be possible to instruct our tool to ignore threads where content is redundant and extremely short.

As mentioned in Section 3.4.4, we have monitored some extra features that may represent a red flag in crawler detection. Nonetheless, they're not part of the experiment since are enforced conditions (like the filling of the referral field) by the design of our tool. These are shown, for reference, in Table 3.4: from the analysis of the HTTP logs generated from CARONTE and the selected tools against our test infrastructure, it emerges that CARONTE explores threads one at a time, sequentially, while other crawlers tend to open multiple resources in parallel. A1 Website Download has never filled the referrer URL in the HTTP requests, highlighting the fact that this request has not been sent from a legitimate browser. In the last analysis, for these monitored aspects, we can say that HTTrack performs better than the others in terms of browser features exhibited.

In conclusion, CARONTE is a prototype that shows the potential as a tool for circumventing passive traffic monitoring and intrusion detection tools; a more sophisticated or active adversary could introduce new techniques to recognize if the user connected is a human (live chat, custom-made CAPTCHAs, ...) or adopt anti-crawling mechanisms that leverage on ML. We are aware that the threat scenario is in constant evolution, and that our study is limited to four underground markets only. Thence, CARONTE should be not considered a stable solution, but rather the starting point for additional studies and more robust implementations.

3.5.1. Limitations

Although our evaluation indicated positive results, CARONTE should not be considered a ready-to-use tool for stealthy automated data extraction. Despite the network logs not highlighting any particular red flag in terms of technical anomalies, more refined crawler detection modules may detect patterns in the traffic that could lead to the termination of the session. Moreover, the human behavior model, albeit nuanced, could raise suspicion over the long run due to the emergence of patterns, especially in the bootstrap phase of the crawler. As a matter of fact, CARONTE has not been tested in a real setting, hence leaving us without conclusive results about its stability and performance toward crawling completeness. As a proof-of-concept, CARONTE showcased some weaknesses: for example, it currently does not provide a guided fallback procedure to execute an alternative strategy to calculate otherwise unstable identifiers in the DOM (as reported in Section 3.4.3). Furthermore, CARONTE lies on the assumption that the target forum accepts the strict no-JavaScript policy imposed by crawler design, or that CAPTCHAs will not appear on forums, both of which may turn untrue in the future. In the case of a forum requiring JavaScript to be enabled, the target website could potentially be unusable or entirely prevent the crawler from browsing it until JavaScript is not enabled.

In conclusion, the proposed approach of relying on an instrumented instance of a browser with the (to some extent simplistic) human behavior model embedded in the implementation is far from being a silver bullet for stealth data extraction in underground communities. Rather, the goal of this work is to explore the potentialities of an approach that overcomes some of the typical weaknesses of crawlers, while trying to differentiate itself from the population of otherwise more studied crawlers.

3.6. Chapter conclusion

Automated tools that gather data in a stealthy way from high-profile cybercriminal forums are a growing need in our society, due to the role that these platforms cover in the attack generation process. The results from the proposed method and associated implementation are encouraging; CARONTE has proven to be effective in learning forum structures and using the obtained information to traverse them and extract the relevant content, with the benefit of instrumenting the crawler without any technical knowledge. Furthermore, relying even on a somewhat rudimentary human behavior model was sufficient to distinguish it from the traffic patterns generated by other state-of-the-art crawlers, while producing comparable features to humans, virtually making it harder to detect. These findings pave the road for the next attempt to develop a more rigorous method and tool for stealth data extraction from underground communities.

CARONTE is available at `https://github.com/michelecampobasso/caronte-crawler`.

4

Scalability of underground monitoring

This chapter is based on [Campobasso3]:
M. Campobasso, and L. Allodi
THREAT/crawl: a Trainable, Highly-Reusable, and Extensible Automated Method and Tool to Crawl Criminal Underground Forums
17th Symposium on Electronic Crime Research (APWG eCrime 2022)

In the previous chapter, we proposed a general method to build a crawler with crawler evasion techniques, and we experimentally validated its effectiveness against other crawlers and humans with success. In that context, we developed and tested a rudimentary procedure to create a tailored instance of the crawler to scrape a specific target. The positive results led us to further focus our efforts to implement a customizable and extensible stealthy crawler supporting the research in underground forums marketplaces. In this chapter, we present an open-source, general-purpose stealthy crawler named `THREAT/crawl`. Aside from its stealthy capabilities, partially drawn from the findings reported in Chapter 3, we focused on the development and implementation of a supervised procedure that allows an arbitrary user of the tool to train a specialized instance of `THREAT/crawl`. This instance is tasked with extracting the content of a specific forum. As a result, this tool allows for performing stealth monitoring (via customizable human behavior modeling) of a diverse population of underground forums. Albeit not final, `THREAT/crawl` is a prototype (TRL-5) available to the scientific community, ready to use and open to further development, to help with the task of extracting fresh data from live underground communities and offering the opportunity to other fellow scholars to engage with research in this discipline.

Link to `THREAT/crawl`:
`https://gitlab.tue.nl/threat-crawl/THREATcrawl`

4.1. Introduction

As discussed in Section 3.1, a large fraction of the overall 'cost' of monitoring cybercrime communities is constituted by the need to build *ad-hoc* crawlers and parsers capable of correctly navigating different forums, extracting relevant content, while remaining under the radar [209]. The positive results on the stealth and rudimentary training capabilities of CARONTE led us to further investigate the identification of the characteristics of a reusable, flexible, and easy-to-deploy crawler for stealth cybercrime community monitoring.

In this chapter, we first derive from the literature and discuss the overall 'foundational' dimensions of this problem, and identify the requirements for a general method over which we design our prototype, THREAT/crawl. We showcase several design choices that can effectively bridge the gap between the dimensions of the problem, allowing to develop a tool that can learn how to crawl a wide range of different forum structures without requiring its users to re-write a parser for each new forum (or, sometimes, forum section). THREAT/crawl provides a simple interface for users to identify specific elements of interest, such as navigation buttons and content, and employs a set of strategies to automatically instrument the crawling engine with the corresponding information needed to traverse the HTML structure of the relevant page(s). Further, the tool can be extended by allowing users to inject (JS) code during the training/crawling phase to perform specific actions (e.g., to adapt the procedure to the specific forum instance), and comes by design with extensive stealth capabilities (some of which already implemented in Chapter 3) to remain under the radar during crawling, if needed. To evaluate the effectiveness of our method and its prototype implementation, we showcase the implemented features of the tool against a set of live, active underground communities and discuss the functionalities and limitations of our implementation.

The contribution of this work is threefold:

1. We analyze the foundational challenges posed by the problem of designing a general, reusable crawler for underground forums; the identified challenges can help frame future contributions in this space;
2. We propose a general method and solution to address the identified challenges, and we provide an implementation showcasing the method and the architectural components addressing each of the identified challenges;
3. We evaluate our method and prototype tool against seven live, active criminal underground forums and identify key aspects for improvement of the tool implementation. We release THREAT/crawl publicly to allow for any uptake and employment of the tool from the community.

The chapter proceeds as follows: in Section 4.2 we identify key dimensions of the problem and derive requirements that a general solution must satisfy; Section 4.3 discusses related work and compares different solutions over the identified requirements. Section 4.4 details the overall design and implementation of our solution, and Section 4.5 evaluates it against seven live underground communities. Section 4.6 discusses results and limitations of our solution, the next steps to take to evolve THREAT/crawl to a mature solution, and Section 4.7 concludes the chapter.

4.2. Problem Space and Solution Requirements

We identify two main dimensions to the problem of designing a general, reusable crawler for underground forums: the diversity of the forums, and the adversarial nature of the monitored environment. A third dimension on ethical aspects is transversal to these.

4.2.1. Adversarial environment for crawling

Crawlers can oftentimes be easily identified due to their high content fetch rates from a target, and from their typical approach to exploring the available content of the target website. These include consumer services such as Cloudflare DDoS protection [257, 69] and other DDoS protection services, provided by underground actors or by so-called Bulletproof-Hosting services [237, 109], specialized in defending onion websites. Browser fingerprinting is an effective strategy requiring the request issuer to execute JavaScript to verify a number of properties of the browser environment, which are hard to mimic with the use of scripts. Other anti-crawler measures include HTTP request inspection, which can provide several indicators of bot activity (e.g., lack of proper 'referer' headers in HTTP requests). As a result of the detection, forums may slow down the crawling process [257], showing CAPTCHAs [69] or throttling traffic, or may ban the forum account used to access the crawled resources [257]. The latter case is particularly concerning in the case of communities enforcing strict access control mechanisms at registration time (e.g., registration on invitation/paywall), potentially jeopardizing months or years of efforts in creating a 'legitimate' identity in the underground to infiltrate [Campobasso7]; similarly, this may pose ethical issues when access fees have to be paid multiple times, potentially compromising the balance between achieving research goals and not providing tangible (economic) support to criminals as a result of multiple payments. Hence, to deal with anti-crawler countermeasures, a general tool has to be:

> **Stealth**: it should avoid generating suspicious traffic while attempting to reproduce the forum navigation activity of a regular user. (R1)

Depending on the community to monitor, researchers may want to finely tune the time at which pages are visited (e.g., according to specific time zones) and more in general the crawling operation as a whole by gathering only relevant information from specific sections. Therefore, a general tool must also be:

> **Configurable**: it should allow the user to finely tune the crawler operation, including the speed and time of the crawling and providing additional (run-time) information for its execution. (R2)

4.2.2. Diverse underground communities

Albeit most forums are generally similar to each other, implementations come with their own peculiarities that make (automated) navigation not trivial [209, 257, 91, 92, 196, 69, 34]. To scrape their pages, developing an *ad-hoc* crawler for each target website is a costly and inefficient procedure, which has to account for different aspects of the target [92, 141, 257]. Forums may be implemented using multiple CMSes [91, 141, 257], with different versions,

flavors, and skins, or even be completely custom solutions [209, 257], making the derivation of a crawling algorithm addressing all the peculiarities for each forum a challenging task. Also, several forums implement anti-crawler mechanisms that randomize HTML attributes like IDs or classes, making the position of content within a page unpredictable. Hence, a general crawler tool should be:

> **Trainable**: it should be capable of learning how to crawl different forums, independently of their structure, DOM properties, and design and deployment solutions. (R3)

Because of non-standard implementations that may appear across different forums (e.g., customization of navigation features such as JavaScript-enabled navigation buttons), a general tool must also be:

> **Extensible**: it should allow users to extend the tool's capabilities by injecting simple procedures in the crawling process whenever a non-standard situation not supported by the tool is encountered. (R4)

Apart from the challenge of crawling different forums, merely downloading pages still poses the problem of content extraction [69], as it may be organized differently across different forums. For example, each post in a thread generally contains information regarding the author (e.g., registration date, popularity, number of posts, ...), but their arrangement and/or identification on the page may vary significantly, some may be missing, and their position may not be constant across sections of the same forum. This forces the implementation of custom parsers for content extraction for every single forum to scrape (and, occasionally, different parsers for pages in the same forum) [257, 196]. This adds additional overhead to the data collection process, increases time-consuming testing requirements, and generates parser software that cannot generally be re-used. Therefore, a general tool should offer:

> **Structured data collection**: it should provide the capability to parse content from crawled pages regardless of how these are structured. (R5)

4.2.3. Ethical considerations

The usage of a crawler should always be subject to ethical considerations. Given the sensible nature of the problem addressed in this work (i.e., monitoring criminal forum communities), we believe these concerns should be addressed at the design level by the tool itself rather than being left entirely for further consideration by the user. As such, requirements over this dimension are to an extent 'orthogonal' to the other requirements, and we therefore label them differently as meta-requirements, `MRx`.

We identify two major concerns that must be addressed. First, the monitoring of adversarial environments may require the user to remain anonymous during (and oftentimes after [Campobasso7]) the crawling activity, which in practice often results in tunneling the traffic over the TOR virtual network [209, 34, 196, Campobasso8, Campobasso7]. Because of the limited bandwidth available to each Onion Router (i.e., a hop in the TOR network), this may compromise or altogether inhibit the experience of other TOR users (which may

Table 4.1: Mapping of tools and solutions from the literature to the requirements of the problem space.

	R1	R2	R3	R4	R5	MR1	MR2	Released?
[91]	◑	◑	○	◑	○	○	○	No
[141]	○	○	◑	○	○	○	○	No
[209]	●	◑	○	●	●	●	●	No
[92]	◑	?	○	○	●	○	○	No
[128]	◑	?	○	○	◑	○	○	No
[34]	◑	?	○	○	●	○	○	-
[Campobasso8]	●	○	●	○	●	◑	○	Yes
[196]	○	○	○	○	◑	◑	○	Yes†
[147]	○	○	○	○	◑	○	○	No
Proposed tool	●	●	●	●	●	●	●	Yes

†: under commercial agreement.

be using it to communicate sensible data or avoid surveillance on their Internet activity). Therefore, a general tool should satisfy the following meta-requirement:

> **Parsimonious**: a tool should limit the bandwidth usage over a private network. (MR1)

Secondly, some content on criminal forums and communities in general may be offensive, or outright illegal even only to access. It is therefore important that a general tool respects the following meta-requirement:

> **Censoring**: a tool should be able to censor by not obtaining, downloading, or saving material that is undesired by the user. (MR2)

In addition, collecting data from underground markets and using stealth techniques may violate the terms of services of a target forum [209] and potentially disrupt their activity [52]. Generally, the societal benefit of studying cybercrime outweighs the remaining risks [176, 34], as also evidenced by the numerous studies in this domain, but ultimately this evaluation has to remain with the final user (and relevant ERB).

4.3. Related Work

Table 4.1 provides an overview of previous work on (or employing) crawlers for underground forums, alongside some open-source implementations available. Early studies collecting data from live underground communities can be traced back to the late 2000s [280, 91]; almost immediately, two problems emerged: the development of complex crawling infrastructures for that specific purpose, and the need to implement strategies to circumvent the target's (at that time, rudimentary) anti-crawler measures [91]. The overhead of creating non-reusable software for each target platform quickly became an evident problem; Jiang et al. proposed a supervised learning tool to teach the crawler how to predict the URL structure of links in a forum [141] (R3). However, this solution disregards the critical aspects of preserving stealth operations, and there are no clear suggestions of data parsing capabilities.

Other studies on cybercriminal activities with the use of crawlers to scrape content from underground communities followed, but remained limited in the scope of the analysis to one or few forums, using *ad-hoc* solutions (R4), and not or only partially accounting for stealthiness [147, 196, 92, 141, 128] (R1). Other studies rely on not-clearly defined crawling infrastructures and focus on the development of different data extraction strategies [92, 147] or the identification of relevant products [196] and actors [128] using natural language processing (R5). Campobasso et al. [Campobasso8] developed a software showcasing a supervised procedure to teach the crawler where to find the needed elements to crawl, parse, and save within the pages of a target forum (R3, R5), while accounting for some anti-crawler techniques and trying to remain stealth by modeling human behavior [209, 34, Campobasso8] (R1). A more remarkable example of reusable software was proposed by Pastrana et al. [209], providing support for the addition of new modules (R4) to enable crawling and data extraction (R5). They developed a crawler accounting several aspects to conduct stealth operations such as human behavior modeling (R1), similarly to [Campobasso8], including also the possibility to enable or disable specific behaviors (R2). However, the software does not offer a guided procedure for the creation of new modules. In the panorama of open-source commercial solutions, we mention some of the most famous general-purpose crawlers, such as *scrapy* [228], *Apache Nutch* [159], and *Heritrix* [184], albeit none of them is designed to stealthily crawl underground communities. Some of these solutions aim at rapidly extracting content from pages [228, 159] (sometimes at the cost of stumble into rate limiting or ban from the target [229]) or can be used for archival purposes [184], they offer their users a great degree of flexibility [228] while requiring to code the business logic of the crawler (including any anti-detection strategy), and necessitate of additional libraries to support browser capabilities, like session handling or JavaScript execution [228, 159, 184].

4.4. Overall Method and Solution Design

We name our method and tool THREAT/crawl. Table 4.2 provides an overview of issues associated with each (meta) requirement and the corresponding strategy employed by THREAT/crawl. The overall THREAT/crawl design is described in the following.

4.4.1. Solution architecture

The different architectural components are mapped to one or more requirements; the mapping is summarized in Table 4.3.

A1. Training module

The training process is summarized in the process diagram in Figure 4.1.

Bootstrap. Before crawling a new community for the first time, it is necessary for THREAT/crawl to learn how to navigate it. To achieve this, the user starts THREAT/crawl and defines a new configuration, then starts the training. In the configuration interface, the user provides a list of example URLs (login, home page, section, optionally

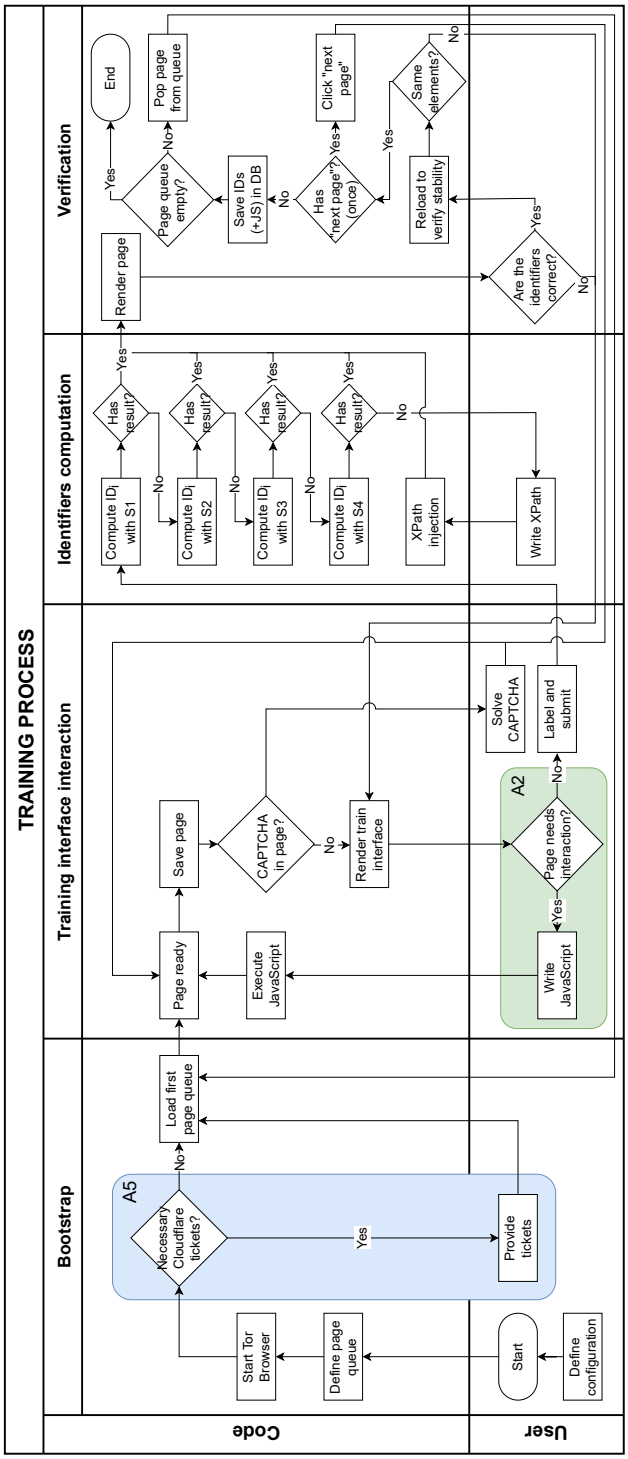


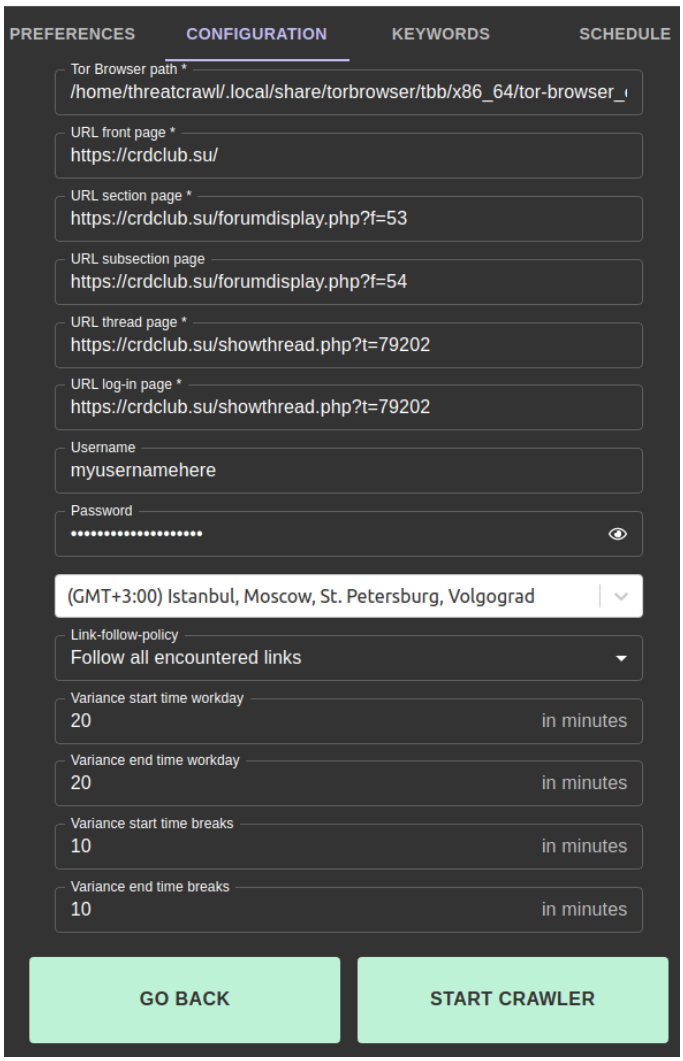
Figure 4.1: Trainer execution process diagram.

Table 4.2: Summary of issues and strategies addressing each identified requirement in THREAT/crawl.

Requirement	Issues	Strategy
R1. Stealth	Crawler traffic is easily identifiable compared to 'user' behavior, (absence of) specific information in HTTP requests, or by fingerprinting the device [257, 69].	Use a real browser for the crawling; mimic human behavior adopting the strategy used in [209, Campobasso8] to interact with buttons and links in the displayed page, and regulate timing according to the amount of text in the page [Campobasso8].
R2. Configurable	Crawling must be limited in time during the day and should not show consistent patterns over time; not all content is relevant to crawling.	Allow the user to define a crawling schedule for each day of the week, as well as when to pause. To limit the onset of patterns, a number of randomly generated pauses are integrated into the schedule, and random noise is added to the start and end times. Specify the 'reading speed' of the simulated user to calculate the time between accesses to subsequent resources. Define (white- and black-listed) keywords to select (or avoid) specific content. See also MR2.
R3. Trainable	Flexible identification of relevant HTML elements in a page.	Identify strategies to account for diverse forum structures and provide different solutions to infer the necessary identifiers, while offering a guided and simplified procedure to its user.
R4. Extensible	Underground communities increasingly feature modern CMSes, supporting dynamic content generation and requiring JavaScript, making the localization of content hard to predict.	Allow the user to input pre-generated identifiers to find HTML element identifiers. In some circumstances, content on a page may be overshadowed or can be dynamically generated, requiring to interact first with the page; for this reason, THREAT/crawl provides the option to perform preliminary operations by executing JavaScript on the page.
R5. Structured data collection	Crawled web pages need to be parsed to extract and structure their content in an underlying database.	Allow users to label specific content to be saved from each page in a structured way.
MR1. Parsimonious	Limit the use of shared bandwidth in TOR while keeping crawling functionalities sufficiently fast.	Limit content to be crawled to focus on what is necessary (R2) and throttling traffic (R1).
MR2. Censoring	Avoid the download of unwanted material.	Keyword white- and black-list matching (R2).

subsection, and thread pages) on which the training will be executed (Figure 4.2). In addition, the configuration interface allows to select the desired timezone and variance ranges for the time to start and end of both crawling and breaks. During the configuration, it is also possible to specify the policy of content exploration, which can filter out content when matching a list of blacklisted keywords (e.g., avoid exploring threads), or exploring only content containing specific keywords of interest. THREAT/crawl provides an interface to define the blacklisted keywords to filter from the crawling or those of interest (MR2, Figure 4.3), and the execution schedule (Figure 4.4).

THREAT/crawl allows to modify a configuration or a training after these have been defined. The identification of representative example pages is not necessarily a trivial task, and changing one of the selected URLs to train again for that page type while preserving the other correct trained pages should be a possible alternative. This is desirable in the case of a minor update in the forum structure breaking the crawling procedure, or when finer tuning to deal with exceptions is necessary.



When creating a new configuration, in the configuration interface the user can provide the list of example URLs that will be used during the training, credentials, timezone to which the schedule will apply, keywords policy (i.e., open only threads containing specific keywords or all), and the start and end variances for both crawler session and breaks in minutes.

Figure 4.2: Configuration interface of `THREAT/crawl`.

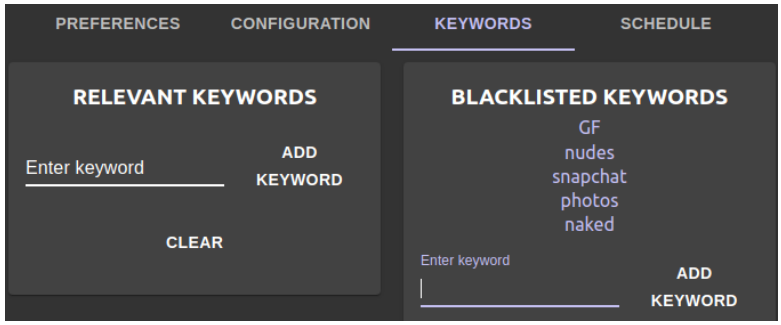


Figure 4.3: Keywords definition in the configuration interface.

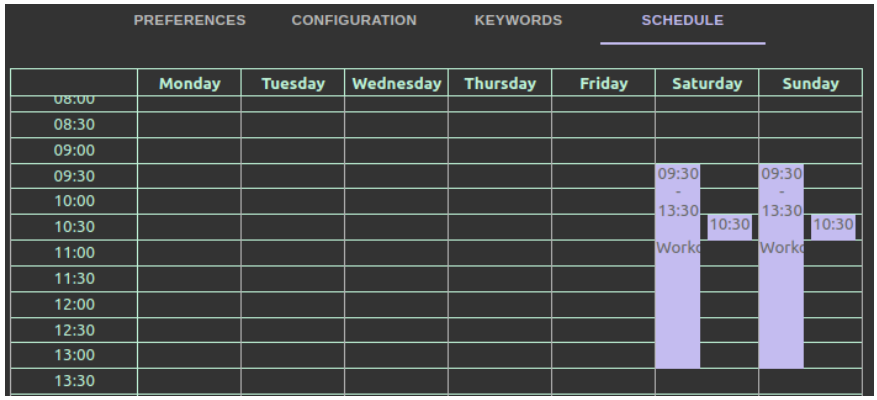


Figure 4.4: Overview of the scheduler configuration interface.

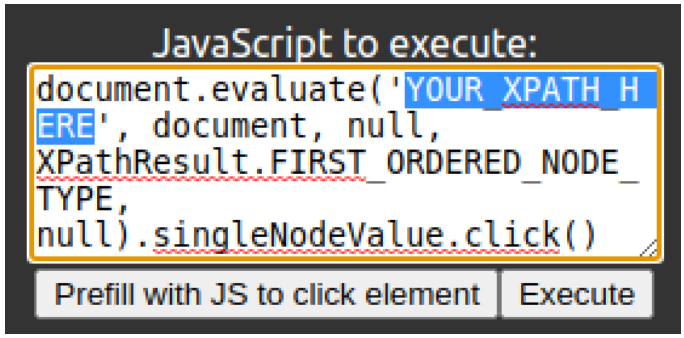


Figure 4.5: JavaScript injection module prefilled with a 'click-on-element' script.

Table 4.3: Requirements to arch. components mapping.

	R1	R2	R3	R4	R5
A1 Training module	✓	✓	✓		✓
A2 Javascript injection module				✓	
A3 Scheduler	✓	✓			
A4 Crawling module	✓	✓			✓
A5 Privacy pass module	✓	✓			

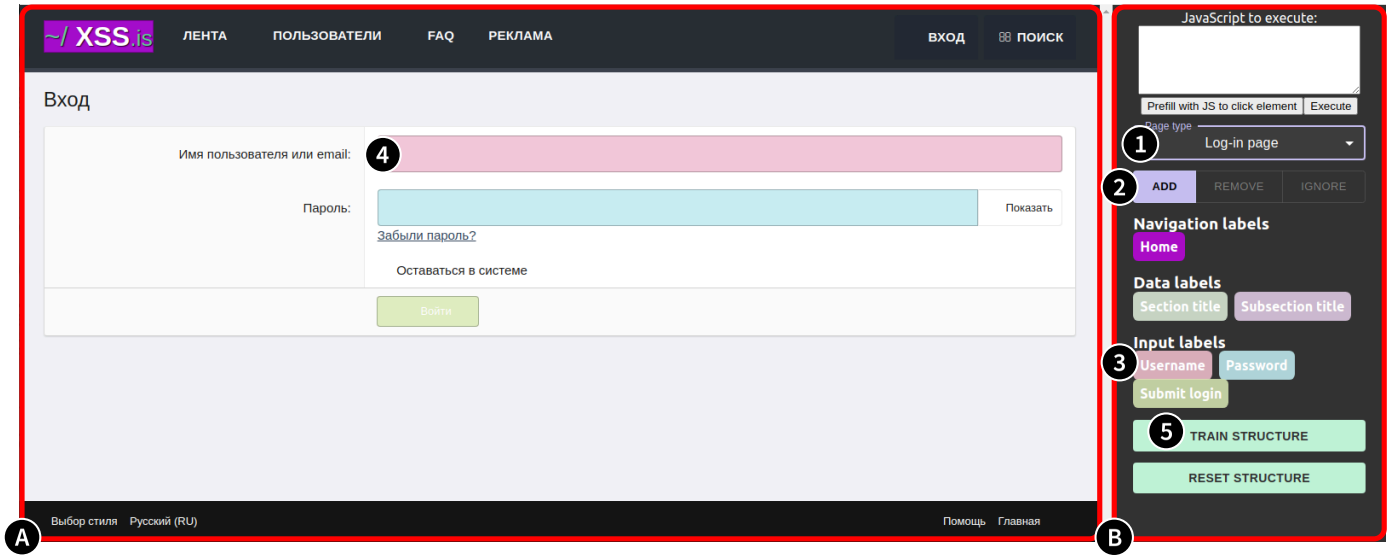
For this reason, THREAT/crawl offers the possibility to select a previous configuration and untick the 'skip training' in the configuration interface; starting will result in a new training session pre-labeling the previously identified elements. When started, the tool prepares a queue with the URLs provided during the configuration and starts TOR Browser. The tool asks the user if the target website needs Cloudflare tickets to avoid encountering their DDoS protection page. This will be detailed in the discussion of A5. The training module can now proceed to load the first page in the queue.

Training interface interaction. When the page is ready in the browser, it is saved and checked for CAPTCHAs. If any, the user is notified via the CLI, asking to solve it. The page is then rendered in the training interface, as shown in Figure 4.6. The login page is then saved and, if no CAPTCHA was encountered, rendered in the training interface, as shown in Figure 4.6. Before the labeling, the user has the possibility to execute JavaScript on the page to interact with it, for example, to show the content of interest otherwise hidden (ref. A2, Figure 4.5). The training interface allows the user to specify the type of the current page (login, homepage, section, optionally subsection, and thread), which is used to learn the target structure. After indicating the correct page type, THREAT/crawl will show the specific labels relevant to that page. By clicking on a label, it is possible to apply it to the relevant element(s).[1] Once the user identified each relevant element on the displayed page, they can confirm the selection.

Identifiers computation. The identified elements are processed to produce *XPaths* identifiers. *XPaths* are calculated using four strategies, each working as fallback to the previous one. Each element's XPath is independently calculated with the first strategy yielding a correct result; the four strategies are:

- The *first strategy* is an implementation of the algorithm proposed by Leotta et al. [162], that prioritizes a set of attributes stably identifying HTML elements (e.g., id, name, class, ...), avoiding those who do the opposite (e.g., src, href, height, ...). However, in the context of anti-crawler measures, IDs and names may be randomized, resulting in an apparently valid training that is no longer effective in a new session or on page reload. Also, this strategy can only calculate the identifier for a single element, while in many cases we may need to have an identifier matching several elements.
- The *second strategy* accounts for these limitations and tries to calculate XPaths by constructing the absolute XPath (i.e., fully characterizing all the descending selectors start-

[1] To keep the training procedure as flexible as possible, THREAT/crawl does not mandate the training of any specific element(R4).



To the left, there is the rendered page (A), and to the right, there is the training pane (B). From the latter, the user first has to specify what type of page is being trained (1) (in this case, login page). The user selects 'add' (2) and clicks on the label to assign (in this case, 'Username') (3) in the training pane (B). In the rendered page (A) it is now possible to click on the corresponding input field (4). This will color the input field with the same color as the selected label. In the case the user misclicks the box, they can click on 'remove' next to 'add' to remove the wrong selection and try again. When the user labels all the relevant elements, they can confirm by clicking on 'Train structure' (5).

Figure 4.6: Training window for a login page.

ing from the root element `/html`). If two or more elements are provided, the strategy calculates the absolute XPaths for all of them and derives the common XPath matching them all by considering their common roots.

- The *third strategy* attempts to extract an element's `class` attribute in the case of randomized identifiers and full XPath changing (e.g., the number of navigation buttons changing per thread depending on their number of pages). While this can be also performed from the first strategy, it is not capable of deriving the common classes for different elements. This method therefore attempts to calculate the common `class` attribute of different elements.

- The *fourth strategy* tries to calculate the XPath using Selenium; for the obtained identifiers from the training, Selenium finds the related `WebElement` object and extracts their XPaths. In the case of multiple elements, this strategy uses the second strategy's approach to derive a common XPath.

As a last resort, in case all four strategies fail to deliver an XPath, `THREAT/crawl` will ask to provide an XPath identifying the problematic element(s) in the displayed page. This requires the user to manually calculate a stable XPath.

Verification. Once all the submitted elements yielded an XPath, a new confirmation window will render again the page, color-coding each relevant element for visual verification. The user can verify if the training was correct; if one of the XPath calculation strategies yielded a wrong result it is possible to adjust the learned structure, by correcting the wrongly labeled elements. This will cause the next strategy to run and calculate new XPaths. In the case of XPaths matching more elements than desired (e.g., only specific sections identified by their `tr` element, rather than all `tr` elements inside of a `table`), `THREAT/crawl` offers the possibility to (de)select all elements that should not be included in the current selection. Once the page is correctly labeled, the training for the current page is considered complete; the page is then reloaded, and `THREAT/crawl` checks if the calculated identifiers are stable. If not, a prompt will ask if the user can still see the element that `THREAT/crawl` could not find. If this is the case, `THREAT/crawl` considers the element identifiers in the page as unstable (i.e., likely randomized) and the training continues until a strategy provides stable identifiers. When the identifiers are deemed stable, it is possible to move to the next page in the queue. For pages containing navigation items (i.e., next or previous page buttons) `THREAT/crawl` will load also the next page by clicking the next page button, to verify that the right strategy was used and the training was successful; these elements are particularly sensitive as their position and number varies when moving to the next page. When one page contains multiple elements of the same category (e.g., thread titles), it is possible to click multiple of them; in this case, the training module employs the first strategy capturing the XPath matching all the identified elements. Finally, for elements containing dates (e.g., post date), it is possible to specify a date format for parsing. Once the training for all pages is terminated, the procedure is completed.

A2. JavaScript injection module

Some underground platforms are particularly difficult to crawl, as they dynamically load content on the page upon interaction without affecting the URL, making the desired content unreachable at that stage. In the training page, THREAT/crawl offers the possibility to inject and execute JavaScript in the loaded page via the execute_script function of *tbselenium* [5], a browser instrumentation library extending the popular browser instrumentation library *Selenium* [130] to support TOR Browser. After the script execution, THREAT/crawl proceeds to render the updated page again in the training interface. Upon confirmation of the training, the JavaScript code is saved in the database. Every time a page of the same type is loaded during the crawling, THREAT/crawl will execute the script before interacting with it. Among other use cases, the JavaScript injection module allows to remove elements hindering the interaction with the page (e.g., closing a popup) or to show the list of sections of interest. The module offers a button that pre-generates the needed code to click on an element, and the user needs only to identify its XPath and replace it in the code (Figure 4.5). More advanced cases include accessing 'private' forum sections, where a page requiring an additional password may be displayed, or showing hidden post content after multiple interactions with the page.

A3. Scheduler

During the setup of THREAT/crawl it is possible to define a schedule for the crawler. The scheduler allows to specify when the crawler should start and end its execution over each weekday and to schedule pauses in between a crawling session. Also, it is possible to specify how strictly the schedule must be followed, by defining ranges that alter the start and end time of both crawler activity and pauses, and the timezone to which the schedule applies. When the crawler is allowed to start, the scheduler compiles a list of time spans for the crawler to run or pause.

A4. Crawling module

The crawling process is summarized in the process diagram in Figure 4.7. The crawler uses an instrumented instance of TOR Browser [251], maneuvered via *tbselenium*. We decided to use TOR Browser to improve the anonymity of THREAT/crawlwhile granting access to underground communities available over TOR. Selenium uses geckodriver to hook to TOR Browser's APIs; however, it discloses that the current browser is controlled by automation by setting a read-only variable navigator.webdriver to true. To avoid this, we create a profile for TOR Browser coming with the extension TamperMonkey, which allows to create and execute scripts during the lifecycle of a webpage.

Bootstrap. The crawler can start in two different ways: it can begin after completing a training procedure for a new forum, or from a pre-existing configuration of interest for which training already happened. In both cases, the GUI spawns the crawler process by passing the relevant configuration. The scheduler calculates when the crawler should start and end the activity, and schedules both the breaks defined during the configuration and a number of random interrupts. When it is time to start, it creates a queue of pages to visit (namely, the login

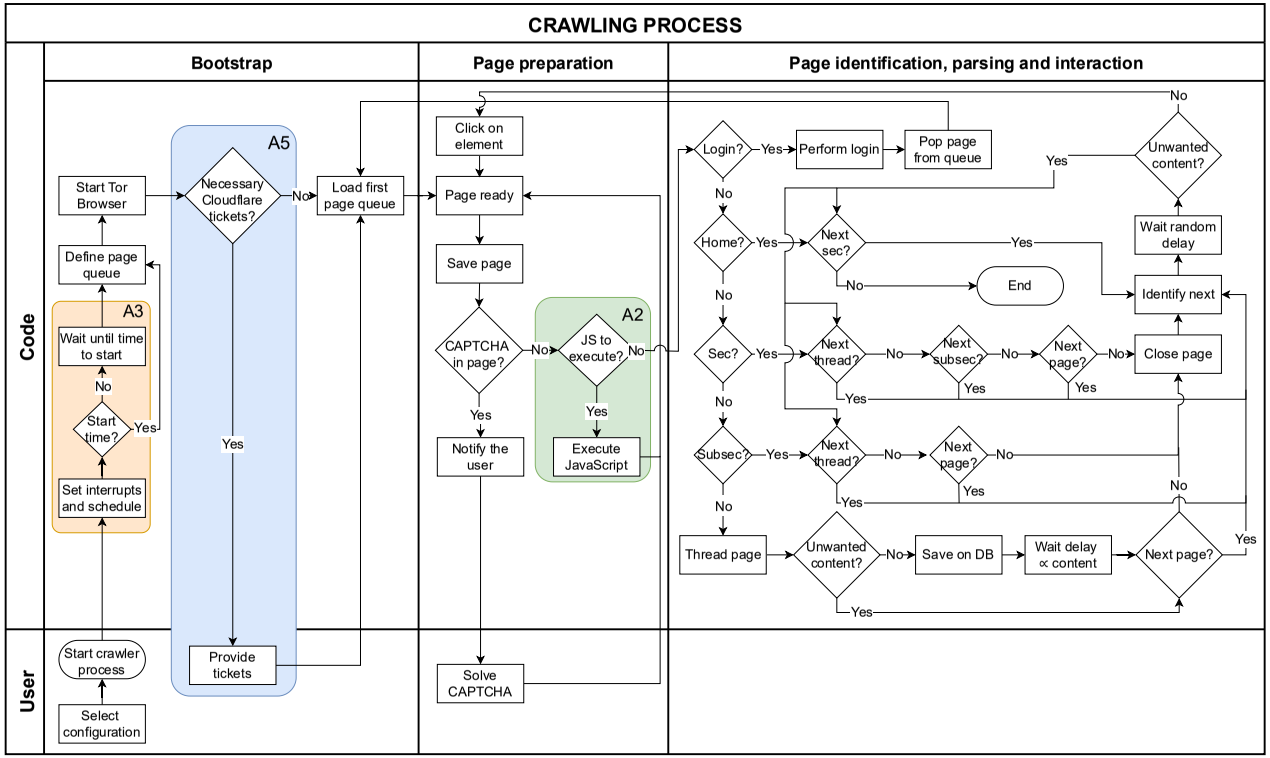


Figure 4.7: Crawler execution process diagram.

page and the home page only) and starts TOR Browser. Similarly to `A1`, it asks the user if Cloudflare tickets are necessary. Then, it loads the first page in the queue.

Page preparation. When the (login) page is ready, the crawler saves it and checks for the presence of CAPTCHAs; if present, the user will be prompted to solve it. Then, the module checks if for that page there is some JavaScript to execute. If so, the page is saved again and it is ready to be parsed.

Page identification, parsing and interaction. In the case of a login page, login is executed and the next page loaded is the home page. When moving across pages that are not thread pages, `THREAT/crawl` waits a random number of seconds between 5 and 15. For each section (and optionally subsection, if specified), `THREAT/crawl` checks if there are threads yet to crawl. If so, each thread of interest (that is, not containing blacklisted keywords in the title) is opened in random order, one at a time, and traversed to extract its content, which is sent to the database. Before moving to the next page of a thread, `THREAT/crawl` calculates a waiting time based on the WPM (words-per-minute) speed defined in the configuration and on the length of the text content of the current page. When the thread crawling is completed, the crawler returns to the parent section (or subsection) and looks for a new thread to crawl. If all threads in the page have been crawled, the crawler will attempt to reach the next page of the current (sub)section. When completed, it moves to the next section to crawl, if any. The crawler can suspend its execution for the scheduled pauses (randomly) planned during the definition of the schedule. When all sections have been fully crawled, the crawler will terminate its execution.

`A5`. Privacy Pass ticket injection module

An increasing amount of underground communities are adopting Cloudflare DDoS protection to mitigate attacks against their infrastructure. Lately, Cloudflare CAPTCHAs have become particularly obtrusive when trying to access a protected website via TOR, due to the low reputation assigned to IPs of TOR's exit-nodes. This causes the presence of very long sequences of CAPTCHAs. In 2018, a group of researchers developed a security-enhancing protocol and an extension in cooperation with Cloudflare that allows users to solve CAPTCHAs in exchange for so-called 'tickets' that can be used to bypass Cloudflare's CAPTCHAs [68]. `THREAT/crawl` comes with a TOR Browser profile with the *Privacy Pass* extension installed, which allows one to legitimately earn tickets from Cloudflare's website `captcha.website`, and to store them in the extension. The only caveat is that earning tickets is not possible via TOR Browser because `captcha.website` is protected from the same DDoS protection mechanism, requiring browsing the website from the clear web. The user can provide the obtained tickets to `THREAT/crawl` via the dedicated interface.[2]

[2]Privacy Pass functionality is suspended by Cloudflare when Cloudflare customers declare an ongoing attack ('I'm under attack!' mode); this provides various mitigation techniques to DDoS attacks, among which disabling the Privacy Pass protocol [57]. This causes the extension not to be effective and the interstitial CAPTCHA page to be displayed.

Table 4.4: Descriptive statistics for the seven selected underground forums.

	Focus	Language(s)	CAPTCHAs?	Sec	Subsec	Threads	Posts	First activity
crdclub	Carding, documents, fraud	EN, RU	No	4	47	86′537	395′276	Jul 8, 2016†
nulled	Leaks, accounts, fraud	EN	Yes	48	45	1′203′886	35′177′498	Apr 22, 2015†
xss	Malware, spam	RU	No	48	3	50′610	394′486	Sep 19, 2018‡
altenen	Ewhoring, malware, accounts	EN	Yes	55	68	970′023	6′840′943	Mar 22, 2010†
nulledbb	Accounts, hosting, Ewhoring	EN	No	23	61	$\sim 206K$	$\sim 1.5M$	Jan 01, 2015†
deeptor	Carding, fraud	EN	No	31	8	14′132	98′631	Jul 24, 2015 †
darknetcity	Accounts, proxy, fraud	EN	No	53	0	3′089	15′018	Oct 26, 2017†

Information fetched on June 18th, 2022. †: oldest staff registration date. ‡: domain registration date.

Table 4.5: Summary of THREAT/crawl performances across the selected forums.

	Train	Crawl	Time_{total}	Time_{breaks}	WPM	Download img	JS exec	CF tickets	Threads	Posts
crdclub	✓	✓	4:05:52	40:45	180 – 240	✓	✗	✗	1	330
nulled	✓	✗	-	-	180 – 240	✓	✓	✓	-	-
xss	✓	✓	3:51:40	42:46	180 – 240	✓	✗	✗	1	580
altenen	✓	✓†	3:31:12	28:29	180 – 240	✗	✓	✗	94	1′691
nulledbb	∼✓	✗	1:04:54	00:00	180 – 240	✓	✗	✗	4	13
deeptor	∼✓	✗‡	08:12‡	00:00	180 – 240	✓	✗	✗	1	10
darknetcity	✓	✓†	3:31:29	44:15	600 – 800	✗	✗	✗	6	1′451

†: premature termination due to connectivity issues with the target; ‡: manual termination of the tool due to wrong behavior during crawling.

All the parameters specifying the page loading and download duration timeouts, the timezone adopted, the variance intervals to apply when calculating the start and end of workday and breaks, as well as interruption duration and minimum time between two interruptions are set to default.

4.5. THREAT/crawl evaluation against live, active underground forums

We tested THREAT/crawl against seven live, active criminal underground forums to which we have access. In this section, we start by providing an overview of these forums and the overall capability of THREAT/crawl to adapt to the different environments to crawl. We then discuss in detail the capabilities of the tool's architectural components in relation to the most interesting challenges posed by the selected forums. Finally, we provide an overall description of the performance of THREAT/crawl across the forums. The present evaluation serves two purposes: (1) to evaluate the effectiveness and performance of the proposed core functionalities of the tool; and (2) to identify weaknesses and points of improvement for future iterations of the tool, as some edge cases not considered at design and implementation time may not be fully supported yet. This data collection was performed under ERB approval ERB2021MCS1.

4.5.1. Selected underground forums

Table 4.4 provides an overview of the selected forums for the evaluation. All forums have been active for at least four years, with the oldest recorded activity in Jan 2015 for nulledbb. Almost all are organized in sections and subsections, although numbers vary widely across forums as do the number of posts. Selected forums are also well-varied in terms of number of posts and cover English and Russian locales. nulled and altenen employ a CAPTCHA system at login time. Below we discuss the forums' relation to the problem space defined

in Section 4.2. crdclub provides a baseline for the performance evaluation. This forum does not come with any specific anti-crawler measure and its structure is rather straightforward. However, it features two inconvenient aspects: it shows a popup when a direct message is received, altering the interaction with the page, and it opens the last page of the thread if we click on its title. The former problem is solved by disabling this option in the user control panel of the website, while the latter, shared with nulledbb, is tackled with a specific solution implemented in A1. nulled represents the benchmark to test both A2 and A5 capabilities; respectively, marketplace sections can be dynamically loaded after clicking on a button on the homepage, and the website is protected by Cloudflare DDoS protection. xss implements ID randomization in the DOM and some elements such as thread title in the subsections and post author details do not come in predictable positions all the time, which is interesting from A1 perspective. altenen is a forum requiring to interact multiple times in a thread to show the hidden content (A2), causing threads to be long and rich in spam, offering interesting considerations during the execution of A1. nulledbb presents an interesting marketplace section, and we configured THREAT/crawl to target content presenting malware-related terminology. Finally, darknetcity is a forum hosted on TOR, and comes with a non-trivial layout for the user details in a post. In addition, altenen and darknetcity present significant performance issues due to the large amount of high-definition GIFs, worsened by TOR's bandwidth; we present a solution to mitigate this issue. We also discuss the problems encountered during the training (nulledbb) and the crawling (deeptor, nulledbb and nulled).

4.5.2. Overall performance

We test THREAT/crawl functionalities by performing training and crawling, for a session of four hours, for all forums. Rather than completing the data collection, our goal here for the presented prototype is to test whether the core functionalities of THREAT/crawl work across all forums and are sufficient, to identify those who need further refinement and to discuss the possible alternatives, while providing an estimate figure of the volume of crawled pages in a unit of time. To provide a range of estimations for different use cases, we customize configurations for each forum depending on the expected quantity of content, long delays from the platform, and desired stealth. Table 4.5 details the chosen configurations and summarizes THREAT/crawl performance. The relatively low number of threads and posts visited for crdclub and xss compared to altenen and darknetcity can be traced back to the verbosity of their posts. Often, members quote the original post of the author, increasing the delay before moving to the next page. Considering that a regular user would notice the repeated content, a higher WPM range could be defined, as we did for darknetcity. Both altenen and darknetcity suffered from premature termination of the crawling session due to connectivity issues with the websites. However, the results were interesting; both performed well, collecting 1′691 and 1′451 posts across 94 and 6 threads respectively in approximately three hours and a half. As mentioned, both forums feature a large amount of high-definition GIFs, making the complete loading of the page extremely long and causing timeouts. In the configuration, THREAT/crawl allows to disregard loading images and, albeit representing a possible suspicious behavior, the solution proved effective. altenen's training had a visual glitch, where clicking on the 'next

page' button would highlight it temporarily, but the training was not negatively affected. For this forum, we also created some JavaScript to show hidden content in the threads by liking and replying to the original post. Finally, we could not perform an adequate crawling session on `nulled`. Despite the presence of valid Cloudflare tickets, the forum set the 'I'm under attack!' mode, disabling the functionalities of the extension [57] and making crawling impossible for more than a few minutes. The training for `nulledbb` was complicated due to an improper rendering of the identified elements, making the adjustment process tedious until the problematic label was identified, and a stable identifier was manually created and provided via the fifth strategy. Whereas `deeptor` successfully completed the training session, its structure seemed to change during crawling, thus making the tool incapable of accessing previously visited threads.

Overall, `THREAT/crawl` was successfully trained for all the forums, with some imperfections for two of them, and managed to crawl four of the seven live forums employed for the evaluation. A fifth one, `nulled` could be theoretically crawled, but the defenses in place at the moment of the benchmark blocked TOR IP addresses, obstructing our operations. The problems encountered in the two remaining forums will be discussed in relation to the appropriate architectural component in Section 4.5.3.

4.5.3. Technical rundown

In this section, we provide an insight into the involved processes from the tool's perspective and detail the most interesting cases.

A1. Training

The underground forum `xss` implements randomized IDs for several elements across its pages as an anti-crawler countermeasure. In the login page, the username and password fields have `id=_xfUid-1-`**`timestamp`**, where `timestamp` is expressed in seconds from epoch. In this case, these elements also present a stable attribute `autocomplete`, which is identified from the first strategy and used as a reliable identifier. On the home page, the user proceeds to label the sections and subsections of interest. In this case, we are only interested in the subsections 'Malware' (`XPath: /html/.../`**`div[5]`**`/.../h3[1]/a[1]`) and 'Cracking' (`XPath: /html/.../`**`div[6]`**`/.../h3[1]/a[1]`) under the section 'Underground'. When two or more elements of the same type are defined, `THREAT/crawl` attempts to infer a common XPath that matches all the selected elements of that type using the second strategy. This is beneficial from a user perspective, as it allows one to identify only a few examples to infer the identifier, instead of clicking them all (e.g., the list of all threads in a page). However, in that case, we are interested in only that specific set of elements; it is possible to click on the button 'Ignore' and select the elements to blacklist (i.e., the uninteresting subsections). The strategy in charge (strategy two) obtains the correct XPath (`XPath: /html/.../`**`div`**`/.../h3[1]/a[1]` while keeping a list of the ignored XPaths, and thus telling `THREAT/crawl` to access only those matching the XPath that are not blacklisted. In subsections, threads can be generally identified by using the XPath `/html/.../div[`**`thread_id`**`]/div[2]/div[1]/a[1]`. However,

sometimes they present a tag before their name, resulting in tagged threads being identified by the XPath selector /html/.../div[thread_id]/div[2]/div[1]/**a[2]**. Selecting both types of threads would generate the common XPath /html/.../**div**/div[2]/div[1]/**a** with strategy two (note that the div in bold matches all the thread_ids), which matches both the thread links and tag links. Once we acknowledge that every subsequent strategy fails to identify the list of threads correctly, we are prompted to provide an XPath that we can calculate by inspecting the page (R4). From manual inspection, it is possible to note that thread titles consistently present the attribute data-xf-init="preview-tooltip". It is possible to use this attribute to generate the XPath //*[@data-xf-init="preview-tooltip"] and to provide it to the trainer, thus solving the problem.

The training of nulledbb resulted challenging in the thread page. One or more wrong identifiers matched wide areas of the page; for example, the post date identifier was wrong and resulted in a verification window without any labeled element. This made the training complicated, as it was not possible to correct the labeling for the single wrong identifier, leading the tool to use new strategies even for the correct identifiers. After some attempts, it became clear that also the identifiers for the post content were creating problems during the rendering of the identified elements. After reiterating the training for a few times, THREAT/crawl asked to provide the XPaths for the problematic elements. From manual inspection of the page, we identified the attribute data-original-title='Original post time' for the post date, and class='post-message flex-fill for the post content, and we created the corresponding stable identifiers.

A summary table with the used strategies used by THREAT/crawl for the identification of each element is reported in Table 4.6.

A2. JavaScript injection module

altenen allows one to see the content of a post after the user 'likes' the post and replies to it as an anti-crawler measure. To solve this problem, during the training of a thread page, the user can write a script to perform these actions. The like, quote, and send reply buttons come in predictable places, //post_footer/div[1]/div/a[1]/span/bdi, //p_footer/div/div/a[2], and //form/../button[1]/span respectively. The user needs to prefill the JavaScript injection box with the 'click on element' code (ref. Figure 4.5) and to provide the relevant XPaths. Considering the time required to submit the like, quote the post and, send the reply, it is necessary to introduce some waiting time between actions, by using the await new Promise(r => setTimeout(r, *millis*)) function. Once ready, the script is executed on the page in TOR Browser via *tbselenium*, and the page is downloaded and rendered again. Upon confirmation of the training, the JavaScript code is saved in the database.

nulled organizes the content in several sections accessible from the home page by clicking on the 'topic' of interest, which shows the relevant sections. This content cannot be accessed directly from a URL but rather requires the user to click on the topic of interest. Similarly as seen in altenen to click on the 'Leaks' topic button, it is sufficient to prefill the JavaScript injection box and provide the button's XPath.

Table 4.6: Summary of the strategies used to derive identifiers during the training.

	Login page				Home page			Section page						Subsection page					
	Home	User	Pass	Login	Home	Sec(s)	Subsec(s)	Home	Sec	Subsec(s)	Threads	Next	Prev	Home	Sec	Subsec	Threads	Next	Prev
crdclub	-	S1	S1	S1	S1	S2	S2	S1	S2	S2	-	-	-	S1	-	S2	S2	-	-
nulled	S1	-	-	S1	S1	S2	S2	S1	S2	S2	S2	S1	S1	S1	S2	S2	INJ	S1	S1
xss	S1	S1	S1	S1	S1	S2	S2	S1	S2	S2	-	-	-	S1	S2	S2	INJ	S1	S1
altenen	S1	S1	S1	S1	S1	S2	S2	S1	S2	S2	S2	S1	S1	S1	S2	S2	S2	S1	S1
nulledbb	-	S1	S1	S1	S1	S2	-	S1	S2	-	S2	INJ	INJ	-	-	-	-	-	-
deeptor	S1	S1	S1	S1	S1	-	S2	S1	S2	-	INJ	S1	S1	-	-	-	-	-	-
darknetcity	S1	S1	S1	S1	S1	S2	S2	S1	S2	-	INJ	S1	S1	-	-	-	-	-	-

	Thread page											
	Home	Next	Prev	First page	Thread title	Thread sec	Post author (PA)	PA # posts	PA popul	PA registration date	Post date	Post content
crdclub	S1	S1	S1	S1	S2	S2	S2	S2	S2	S2	S2	S2
nulled	S1	S1	S1	-	S2	S2	S2	S2	S2	S2	S2	S2
xss	S1	S1	S1	-	S2	S2	S2	S2	INJ	S2	S2	S2
altenen	S1	-	-	-	S2	S2	S2	S2	S2	S2	S2	S2
nulledbb	S1	S1	S1	S1	S2	S2	S2	S2	S2	-	INJ	INJ
deeptor	S1	S1	S1	-	S2	S2	S2	S2	S2	S2	S2	S2
darknetcity	S1	-	-	-	S2	S2	S2	INJ	S2	INJ	S2	S2

It is possible to see that nor S3 or S4 ever appear to be used. However, S3 exists because of `crdclub`; at the beginning of 2022, we could not identify navigational items consistently with the first two strategies, and we developed this strategy to solve this issue. S4 instead has been tested in a synthetic environment, but it did not yield any useful result in the real world as of yet.

A3. Scheduler

The scheduler execution does not have noteworthy details to report for the evaluation set.

A4. Crawler

altenen contains unwanted material, such as so-called revenge pornography material. We set the crawler to explore all links but to avoid threads containing the keywords 'GF', 'nudes', 'photos', 'snapchat', and 'naked' in any of their posts or title (Figure 4.3). To achieve that, the crawler parses the current page (i.e., a section), seeking threads to explore. From the threads list, it checks if any of these should be excluded based on the blacklisted keywords. The same process applies while browsing a thread: if any of the posts mentions any of the keywords, the thread is closed, the posts discarded, and the tool moves to the next thread. Similarly, nulledbb features potentially interesting content among a large amount of spam, and we set the crawler to explore only threads matching one or more relevant keywords related to malware trade. However, crawling for nulledbb failed due to Selenium being unable to detect if the browsed page was successfully loaded or not. This resulted in a page not successfully loaded and Selenium not raising a timeout error to let THREAT/crawl reload the page and try again, ultimately stalling the crawler.

deeptor crawling failed in the section page. The problem is that the position on page for thread titles mutates when a thread is accessed for the first time; when THREAT/crawl returns to the section page after crawling the first thread, it fails to identify all the threads on the page, resulting in an error and prematurely terminating the crawling of the current section. A solution could be to manually inspect the structure of the page to derive XPaths matching both cases. XPath syntax includes a UNION operator, which could be used to derive the list of threads for both cases. Therefore, running again the training and voluntarily falling back to the XPath injection strategy (ignoring the currently correct training not accounting for the future DOM of the page) is a possible workaround. nulledbb crawling stalled when a page failed to load; this is not an uncommon issue within the crawling context, especially when using TOR, and THREAT/crawl manages this issue by interpreting the errors arising from Selenium. However, during this run, we encountered a case in which Selenium 'hangs' indefinitely, and our tool manages the situation as a network issue, and attempts to refresh the page. Despite that, Selenium remains unresponsive, and our tool cannot proceed in the crawling. This issue would require to inspect Selenium and to extend its functionalities.

Although we have not tested this feature, THREAT/crawl naturally handles parallel instances to scrape the same forum over multiple accounts. During execution, when THREAT/crawl reaches a (sub)section page, it selects the next thread to crawl based on its title and link. Before starting to crawl the selected thread, the tool verifies whether this information is already present in the database: in the positive case, the crawler moves to another candidate thread. In fact, when THREAT/crawl begins to crawl a thread, name and link of a thread are immediately stored in the database, minimizing the risk of a race condition. This evaluation is performed every time THREAT/crawl needs to crawl a new thread, thus minimizing the risk of accessing a thread that has already been crawled from another instance, while keeping records of the threads that have been parsed from current or past/concurrent instance(s).

A5. Privacy Pass ticket injection module

To both train and crawl nulled, we need to prevent Cloudflare from showing the CAPTCHA page. To do so, the user has first to earn tickets on captcha.website in the case of Cloudflare and then export them. The user needs to access the Firefox debug mode (about:debugging#/runtime/this-firefox) and click on inspect for Privacy Pass. By browsing the 'Storage' tab, under 'Local Storage' they can find the tickets in the form of two key-value pairs (cf-commitment-2.58 and cf-tokens) to copy and paste into THREAT/crawl's dialog, (Figure 4.8). After submitting all the key-value pairs,

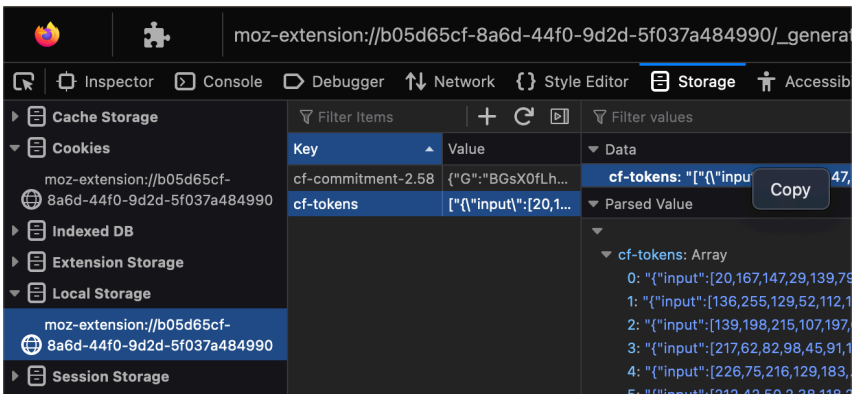


Figure 4.8: Screenshot of the tickets available in the source browser.

THREAT/crawl will trigger a sequence of actions via Selenium, opening the same page and executing JavaScript in the browser console, thus loading the tickets in the extension. Once this operation is completed, the browser will successfully load the target website, bypassing the CAPTCHA. This procedure allowed us to perform the training of nulled; however, when we tried to crawl some days later, the forum set the 'I'm under attack!' mode and stopped accepting tickets (considering the timing of the episode, and the negligible volume of traffic generated by our training we consider it unlikely that our training session caused the state change). A possible solution for a future release of THREAT/crawl would be to allow the option of using either TOR Browser or Firefox for the execution of THREAT/crawl with a dedicated proxy, thus avoiding to use of the IP addresses of TOR exit-nodes which generally suffer of bad reputation.

4.5.4. User interface

A1. Training module

In the *(Bootstrap)* phase of this module, the GUI offers the user the possibility of creating a new configuration for a new forum or using an existing one. The configuration interface is shown in Figure 4.2. When the crawler is allowed to start, it asks for Cloudflare tickets and finally loads the login page. Additionally, the 'Keywords' tab allows the user to specify keywords indicating unwanted content or relevant keywords, allowing the crawler to respec-

tively avoid opening threads containing any of these in the title or any of its posts, or to open only those containing them in the title in the case of relevant keywords (Figure 4.3). The user can now proceed to label the page, by selecting the relevant label and clicking on the corresponding element.

Upon submission, a new window will appear to confirm the current selection, showing the identified elements obtained from THREAT/crawl with the learned information. The user has the chance to manually remove the wrong labels or to reset completely the training to start from scratch and to add them once again. The user can reiterate the process until the result is satisfactory. As a last resort, THREAT/crawl may ask the user to provide an XPath to identify the specified element(s). Once an XPath is provided, THREAT/crawl renders again the page to confirm the selection.

A2. JavaScript injection module

When a script is defined during the training for the specific page, it is executed before interacting with the page. Figure 4.5 shows the interface to inject the desired JavaScript as discussed in Section 4.5.3. The currently displayed code is generated from the 'prefill' button, which allows to identify an element in the page via XPath and to click on it; the user is then left to replace 'YOUR_XPATH_HERE' with the correct XPath.

A3. Scheduler

The schedule for a crawling activity is defined at configuration time. Figure 4.4 provides a view of the interface. The indicated times apply to the specified timezone defined in the configuration tab. In our case the crawler should perform its activity during weekdays from 17:00 to 20:00; during weekends, the crawler should work from 9:30 to 13:30 with a scheduled pause between 10:30 and 11:00 (this is the schedule used for the benchmark), and another crawling session from 15:00 to 20:00.

A4. Crawling module

During the crawling process, one of the two interactions with the user is the notification of a CAPTCHA on the page. During the login on nulled and altenen the crawler informs the user that there is a Google reCAPTCHA on the login page and asks to solve it. When the CAPTCHA is solved, the user confirms by typing 'solved' in the terminal and the crawler proceeds to login. The second interaction available to the user is the manipulation of the execution of THREAT/crawl by typing in the console the commands 'resume' to skip the current break, interruption, or delay, resuming the crawling immediately, 'pause' to suspend its execution or 'terminate' to end the current crawling session.

A5. Privacy Pass ticket injection module

nulled is protected by Cloudflare DDoS protection. Therefore, accessing it via TOR Browser will cause Cloudflare to show the CAPTCHA page and deny access due to TOR's exit-nodes poor reputation. To solve this issue, at startup, the crawler asks whether the user

wants to provide Cloudflare tickets before execution. To earn tickets, it is sufficient to open a regular Firefox instance with the Privacy Pass extension installed and browse the website `captcha.website`. Solving a challenge will grant the extension with 30 tickets necessary to bypass the CAPTCHA page (which could be displayed several times during one session) and consume more than a single ticket to bypass it. Solving the challenge multiple times allows to earn more tickets, granting access for a longer session. These tickets can be exported and provided in the prompt of `THREAT/crawl`, as shown in Figure 4.8.

4.6. Discussion

In this chapter, we presented `THREAT/crawl`, a general method to learn and stealthily crawl arbitrary (underground) forums. We present the foundational challenges, proper of the problem space, such a solution must address, and design our method and overall solution around those. We provide a prototype implementation and test it against live, underground forums. The results show that `THREAT/crawl` successfully managed to learn all the forum structures and to crawl four out of seven forums proposed for validation. The tool successfully learned the structure and content layout of the forums using different strategies accounting for possible anti-crawler measures. `nulled`, `xss` `darknetcity` and `nulledbb` showcase one of a module allowing the user to provide a manually created identifier when all the identification strategies of the tool fail. The crawling sessions, configured to last four hours terminated successfully in four cases. Out of these four, `THREAT/crawl` was configured to not download pictures from two targets, `darknetcity` and `altenen`, mitigating the long loading times during the crawling and reducing TOR network stress; in addition, we modified the WPM parameter to shorten the delay between the crawling of two pages, considering the large amounts of spam in the posts of `darknetcity`. `darknetcity`, `nulledbb` and `deeptor` do not feature any subsections, and the crawler manages this case naturally. `crdclub` and `nulledbb` present a special case in which clicking on a thread opens its last page; the tool offers the possibility to handle this case by training the button to open the first page of the thread when landing on it. In another two forums, `nulled` and `altenen`, we showcased another feature supporting the extensibility of the tool, the JavaScript injection module; for the former, we wrote a script to reach the sections of interest from the front page, while for the latter we wrote a script to interact with the page to reveal the hidden content of a post. Furthermore, during the training of `altenen` we opted not to train the next page button, causing `THREAT/crawl` to seamlessly terminate the crawling of the current thread and move to the next one, to avoid crawling pages containing spam. Moreover, `nulled` imposed a strict policy for visitors, nullifying our efforts of bypassing Cloudflare's CAPTCHAs during the benchmark. However, this is a problem affecting every user of the platform that tries to access it via TOR. A possible solution to this problem would be to offer `THREAT/crawl` the possibility of switching to a regular Firefox browser, using an appropriate proxy for anonymity, and benefiting from a 'clean' IP address. `deeptor` had an apparently correct training, but during the crawling the tool could not find the location of threads within the section; this was caused by a change in the DOM after the thread was opened. However, the problem could be solved by rejecting the apparently correct training for the problematic element until the XPath injection module is triggered and

providing an XPath accounting for both cases. Finally, `nulledbb` had problems in rendering the calculated identifiers during the training due to one or more wrong identifiers; this could be solved by implementing a feature that allows to visually verify identifiers one at a time, narrowing down the retraining to the incorrect identifiers only and avoiding to run the training several times to guess what are the problematic labels. On the same platform, we experienced a limitation of the Selenium framework, which turned out to be incapable of detecting whether a page was successfully loaded or not, stalling the execution.

4.6.1. Final remarks and `THREAT/crawl` release

A significant portion of today's research on cybercriminal communities relies on leaked data and old datasets, allowing to perform *post-mortem* analysis on them. In other cases, researchers develop *ad-hoc* crawlers and parsers to tap data from each community of interest, which is a burdensome procedure. This software is rarely shared among the community, because its purpose is limited to the scope of the research. The cost of developing such software discourages research and limits its scope, whereas the (un)success of extracting data from a community can make the whole research unfeasible. Risks are higher when the target is a prominent community, where the costs (e.g., pecuniary) of losing access are substantial, and obtaining access again may be ethically hard to justify.

Therefore, we propose and release `THREAT/crawl`, a prototype crawler that aims to address different problems related to the crawling of criminal communities, offering a supervised procedure to learn the structure of a target community, supporting manual intervention in particular cases, and enabling dynamic interaction with the page via JavaScript to circumvent several custom anti-crawler mechanisms. Together with a (potentially more supported) CAPTCHA bypass mechanism and the modeling of a seemingly legitimate user, `THREAT/crawl` proves that it is possible to crawl across a number of different underground communities, without the burden of creating scarcely reusable software for both crawling and parsing their content while remaining stealth, and giving the user possibility to tune the tool to reach the desired trade-off between stealth and throughput.

4.6.2. Future work

As the tool is currently a prototype meant to showcase the overall approach and its viability, from the evaluation we identify a number of key improvement points to address in the future, as well as possible uptakes from the community.

Training procedure

The training of `nulledbb` was particularly complicated due to the improper rendering of the calculated identifiers. The training interface could be improved to ease the troubleshooting of these problematic training scenarios, allowing to highlight one family of identified elements at a time. Detecting only CAPTCHAs hinders the crawling of a target platform. A fully-fledged solution would require to detect and forward CAPTCHAs to an operator in charge of solving them to resume the crawling of the target platform. Commercial solutions

managing CAPTCHA resolution exist and could be used for the same purpose. At its current state, the crawler does not support the crawling of threads that have been already visited. This could be solved by training in the section/subsection `THREAT/crawl` to recognize the count of posts within a thread and to keep track of it; if this number does not correspond to the stored value during the previous crawling session, the thread should be visited again. This could be an optional feature working only when this information has been provided to `THREAT/crawl` during the training.

Crawler robustness

In the case of `nulled`, albeit `A5` offers a solution preventing CAPTCHAs being displayed and Privacy Pass being potentially adopted by different DDoS protection services in the near future, it falls short when stronger access policies are enabled, preventing access to users connecting via TOR. To solve that, `THREAT/crawl` should offer the possibility to use a regular Firefox instance tunneling traffic through a proxy different from TOR, benefiting from a non-blacklisted IP address. Another encountered problem is connected to the use of Selenium as the browser instrumentation library of choice. Selenium is designed to be a web application testing framework, and some edge cases regarding connectivity issues that may arise are not taken into account in the current state of the library. Therefore, it is necessary to extend the library to allow its usage as a browser instrumentation framework for crawling over unreliable networks.

4.7. Chapter conclusion

In this chapter, we showcased the advantages and limitations of the current iteration of `THREAT/crawl`. This prototype shows the potential to achieve the identified goals for a reusable and extensible automated crawler for underground communities, highlighting the strengths of an extensible training process tackling different anti-crawler measures. In the near future, we plan to deploy an enhanced version `THREAT/crawl` in our institution to start a longitudinal data collection across different criminal communities for further research in the current and active underground threat scenario.

Publication, Development, and Licensing

The development of `THREAT/crawl` was partially supported by a team of BSc students, composed by Sam Baggen, Akam Bilbas, Steven van den Broek, Yuqin Cui, Milan van Gool, Wouter Haneveer, Peter Heijstek, Samar Jameel, Pim van Leeuwen, Michel van de Looij, and Jeroen Oerlemans, as part of their final graduation project. `THREAT/crawl` is released under GNU Affero General Public License v3.0. Source code and documentation are available at `https://gitlab.tue.nl/threat-crawl/THREATcrawl`.

II

Investigation and evaluation of a prominent, emerging threat from underground markets

5

Identifying emerging threats: the Impersonation-as-a-Service case

This chapter is based on [Campobasso7]:
M. Campobasso, and L. Allodi
Impersonation-as-a-Service: Characterizing the Emerging Criminal Infrastructure for User Impersonation at Scale
2020 ACM SIGSAC Conference on Computer and Communications Security (CCS 2020)

During the preliminary investigation conducted in the context of Chapter 2, we had to interact with multiple and diverse underground communities: to gain access, some of them required as little as a simple registration, while others asked for a 'credential', a proof of our intentions or commitment. Since our goal is to study these communities and not to participate in criminal ventures, in some cases it was necessary to create and 'cultivate' identities across different underground communities, showing interest in products or services, and interacting with the community to build some reputation [1]. With a more stable foothold on different underground communities, we began examining said communities to identify relevant threats and exploring advertisements across multiple underground communities. During the exploration of a longstanding underground market showing convincing evidence of trade of mature offensive capabilities, we identified the advertisement of a new and innovative criminal service. In the advertisement, the administrators provided details about the offered service, available in a one-stop-shop platform called Genesis Market. Apart from the convincing advertisement, we noted a remarkable interest in this service within multiple affiliated communities. Hence, we applied for an invite to the platform to investigate it.

After a lengthy process of infiltration to obtain multiple accounts, and following the ban of one of our accounts used by a non-stealth crawler we initially designed, we employ the methods discussed in Chapters 3 and 4 to develop specialized versions of the tools proposed

[1] Apart from product details requests, the only interactions were in beginner sections on generic topics like online privacy and networking that could be otherwise retrieved from a Google search.

in the same chapters to covertly extract data from Genesis. From the extracted data, we could derive and characterize Genesis Market's novel threat model, for which we coined the Impersonation-as-a-Service (IMPaaS) term. In short, Genesis solves the limitations of credential stuffing attacks against websites protected from Risk-Based Authentication systems, which proceed to fingerprint the authenticating system to further verify its identity.

In relation to the breakdown of Access-as-a-Service cybercriminal operations in the five pillars (Chapter 2), Genesis Market represents an excellent example of how sophisticated criminal services are built. Currently, the cybercriminal ecosystem is dotted with criminal services that can support other criminal operations, and we found evidence that Genesis Market relies on third-party services. We found evidence that Genesis Market relies on Malware-as-a-Service (MaaS), capable of offering 'off-the-shelf' malware on a subscription basis; in addition, our tests on AZORult (one of the adopted malware strains) revealed that Genesis customized it to implement some additional functionalities. By doing so, Genesis' operators partially outsourced the first two pillars of offensive cyber operations (1. vulnerability research and exploit development, and 2. malware payload generation). The third pillar, technical command and control, is very likely to be at least in part outsourced; in fact, it is best practice for criminal services to rely on Bulletproof-Hosting-as-a-Service, services that offer hosting in jurisdictions with lax regulations towards legal complaints from other countries' law enforcement agencies. Finally, pillars 4. operational management and 5. training and support are most likely to be managed directly by Genesis. In fact, Genesis' operators are active across multiple affiliated communities to attract more customers to join their market, actively develop and update the tools necessary to conduct the impersonation attacks, interact with customers, provide support, and detailed documentation.

Therefore, the innovation of Genesis in bypassing state-of-the-art authentication mechanisms, together with its peculiar service composition to support its operations, makes it a particularly interesting criminal service to investigate. In this chapter, we will first discuss the infiltration and the data extraction process (in accordance with the guidelines as per Part I, to then study the platform's threat model, provide insights on the targeted population, and the platform's pricing model. Finally, we discuss our findings to make assumptions about the platform's maturity.

Link to the created datasets (available under license):
`https://security1.win.tue.nl/doku.php?id=artefacts#data_sharing`

5.1. Introduction

In recent years there has been a surge in criminal infrastructures supporting cyberattacks and cybercrime activities at large [113, 15, 47]. For example, *exploitation-as-a-service* and *pay-per-install* provide a set of attack technologies generally aimed at infecting systems or controlling bots that are then employed to launch, for example, DDoS attacks, or subsequent malware and phishing campaigns (e.g., to harvest credit card numbers or steal credentials). An important problem in any venture, let alone a criminal one, is the ability to *systematically* monetize the effort that goes into it [120]. In criminal enterprises, monetization is not

necessarily an easy feat: whereas re-selling or giving access to infected systems to fellow criminals alleviates the problem for who generates the infection (e.g., the *bot herder* [126, 38]), the problem of assigning a price to each bot remains [22]. Whereas the dynamics of demand and offer in the underground are likely to play a role in this setting (and remain an important open question to investigate in this domain), another key factor in determining the value of an infected system is the information it manages and/or processes; for example, access to the email account(s) of an Internet user may have a different value, to attackers, than access to a user profile with a server-stored credit card number (e.g., an e-commerce website). On the other hand, it is not yet clear how (and if) attackers can *systematically* employ those credentials to impersonate Internet users at large, particularly in the presence of multi-factor authentication systems whereby a username and password alone are not sufficient to gain access to an Internet account.

Credential theft and re-selling in underground communities have been studied multiple times in the literature; for example, recent studies provide an in-depth view of what happens to credentials after they have been stolen [201] and their employment for final attacks [248]. Similarly, several studies investigate the attack vectors that allow attackers to obtain the credentials in the first place, ranging from (targeted) phishing and phishing kits to malware infections at scale [42, 46, 201]. On the other hand, systematic employment of the stolen credentials remains out of reach for most attackers: credentials stolen from the underground may be accessed by multiple criminals, effectively destroying their value for later accesses [120]; similarly, the effort required to monetize access to stolen or hijacked user accounts does not scale well with the number of available accounts [120, 119]. In particular, protection systems such as multi-factor and risk-based authentication systems severely limit the capabilities of attackers to effectively employ stolen credentials, requiring the employment of more sophisticated attack vectors than a simple credentials dump [248]. Risk-based authentication systems receive user authentication requests and are responsible for deciding whether additional multi-factor authentication is required for that session, or if the provided (valid) password suffices to grant access to the user requesting it. The idea behind risk-based authentication is that, by 'measuring' certain characteristics of the user environment (i.e., its fingerprint [13]), the authenticating system can build a 'risk profile' associated with that request as a function of the distance between the current fingerprint and the profile associated to the requesting user. If the mismatch is too large, the risk-based authentication system will defer the decision to a multi-factor mechanism (e.g., requesting a code sent to a trusted device or account, such as a mobile phone or an email account); on the other hand, if no anomaly in the user profile is detected, the risk-based authentication system will – in most cases – grant access just with the password.

This mechanism is a significant obstacle to a successful impersonation attack, as the very high dimensionality of a user fingerprint makes it impossible, for an attacker, to *systematically* reproduce it for arbitrary users from scratch [248, 13]. A recent study by Thomas et al. [248] highlights how modern phishing kits [199] are equipped with fingerprinting modules that, together with the user credentials, obtain a measurement of the user's environment that can be re-used to circumvent risk-based systems. On the other hand, obtaining these user profiles requires systematic efforts to phish targets, perhaps across different platforms, and may not provide reliable and stable measures of a user's fingerprint as the victim's inter-

action with the attacker's website may not accurately reflect the victim's interaction with the legitimate website (e.g., for behavioral fingerprinting [242, 49]). Overall, traditional attack strategies seem unsuitable for reliably obtaining, updating, and enforcing user profiles.

In this chapter, we provide evidence of a new emerging criminal infrastructure for *Impersonation-as-a-Service*, that relies on custom malware and a marketplace platform to systematize the delivery of complete *user profiles* to attackers. A user profile on an `IMPaaS` service comes complete with stolen credentials for multiple platforms, the ability to either reproduce or re-generate a user fingerprint from the stolen data, and a software bundle to enforce the user profile during an authentication session. To study the presence of the `IMPaaS` model in the wild, we provide an in-depth analysis of a large criminal platform (Genesis Market) providing, at the time of writing, more than 260′000 profiles of Internet users, globally. Genesis Market is an emerging, invite-only, Russian `IMPaaS` platform currently operating in the underground. To evaluate the nature of `IMPaaS` operations, we dissect the process behind the *acquisition, selection, and enforcement* of stolen user profiles enabled by the `IMPaaS` model, and provide a detailed evaluation of the characteristics of Genesis Market, its extension, the characteristics of the user profiles it provides to final attackers, and the relative effect of different user profile characteristics on its value.

5

Scope and contribution

The contribution of this work is three-fold:

1. we provide the first characterization of the `IMPaaS` model for the systematization of impersonation attacks at scale;
2. we provide an evaluation of a large, invite-only, emergent Russian `IMPaaS` platform that automates the collection, provision, and enforcement of user profiles collected worldwide;
3. we provide insights on the relative effects of different user profile characteristics on the value of the user profile, and quantify these effects.

A detailed technical analysis of the malware for the user profile exfiltration and enforcement is out of the scope of the present work.

This chapter proceeds as follows: Section 5.2 sets the background for impersonation attacks and their relation to existent countermeasures; Section 5.3 introduces the `IMPaaS` model for impersonation attacks at scale, and Section 5.4 describes Genesis Market implementing it, our infiltration and data collection strategy. Genesis Marketoperations are analyzed in Section 5.5. Section 5.6 discusses our findings, and Section 5.7 concludes the chapter.

5.2. Background and Related Work

5.2.1. User impersonation attacks

With the rise of sophisticated web applications, much of a user's Internet activity happens by accessing a multitude of remote services, from banking to e-commerce and social net-

work platforms, through the browser. Most of these services will have authentication mechanisms that are meant to grant access to the underlying service to the authorized user(s) only. From an attacker's perspective, user impersonation provides a large portfolio of additional attack opportunities, ranging from economic gain [15, 98] to more targeted scenarios such as targeted-phishing [124] and violent crimes [118].

Password-based authentication (PBA) is the most common (first) barrier attackers have to overcome to perform an impersonation attack. Whereas passwords have proven difficult to securely handle, are prone to leaks and off-line attacks [187, 273] and still present severe usability problems [241], they represent the most widespread means of authentication online [41, 42]. PBA requires users to create a non-trivial secret, not to reuse it across several services, and to memorize both the secret and where it has been used; nonetheless, several studies indicate that up to $\approx 90\%$ of users reuse passwords or small variations thereof across several services [67, 138].

Whereas this leaves room for password-guessing attacks, additional attack vectors (such as malware and phishing [248, 46]) can be used to obtain user passwords, regardless of their complexity. In general, hijacked accounts can allow adversaries to tap into the social connections of victims to compromise additional accounts [247, 104], by creating targeted social-engineering attacks against their circle of trust or by spamming malicious content [222], liquidate financial assets [137], steal sensitive information with the aim of blackmailing users [46, 222] and sextortion [271]. Additionally, stolen user credentials are oftentimes made available to the cybercrime community through underground markets [201, 248]. These markets generally provide 'dumps' of stolen credentials obtained from data leaks from an affected platform, or as a result of an extensive phishing campaign targeting its users [248]; common target platforms include banking or trading websites, cryptocurrency services, pornographic websites, and other internet services. A recent estimation calculates that, between March 2016 and March 2017, 1.9 billion phished credentials have been sold through the underground markets [248].

5.2.2. Countermeasures to attacks against PBA

Multi-Factor Authentication

To mitigate the shortcomings of authentication mechanisms relying solely on passwords, web platforms have started adopting additional authentication measures such as Multi-Factor Authentication (MFA). MFA moved the authentication paradigm from (solely) something that the user *knows* (e.g., a password) to something the user *has* (e.g., a token) [248, 78]. This is achieved mainly with a combination of a pair of valid credentials and a One-Time Passcode (OTP) received via some trusted component such as a mobile phone, email, or a hardware token [78]. Albeit possible attack scenarios exist where the attacker can obtain the information required for the authentication almost in real-time (stolen token generator, compromised email, SIM swap attacks [189], etc.), MFA dramatically increases the costs for an attacker, and it is widely regarded as an effective countermeasure to password-based impersonation attacks [248]. Nonetheless, MFA is not devoid of security problems, perhaps most notably related to its usability [181], concerns on token-recovery mechanisms, and third-party trust [41].

Risk-Based Authentication

Partly to mitigate the usability problem, *Risk-Based Authentication* (RBA) is oftentimes adopted as a means to evaluate whether the authenticating user is (likely to be) the one that has, historically, access to a specific account. RBA is an adaptive security technique aiming to strengthen password-based authentication by monitoring how unexpected or suspicious a login attempt is from the perspective of the authenticating service [268, 181, 248]. During the authentication, the RBA system monitors both behavioral and technical characteristics of the user and the device, producing a *fingerprint* of the authenticating user [268]. RBA computes a risk score associated with the ongoing authentication by comparing the existent profile of the authenticating user against the features collected for that instance of the authentication. The features vary from basic information such as User-Agent, system time, and OS, to environmental or behavioral features, such as system language, keyboard layout, fonts and plugins installed, mouse movement, geolocation, and keystroke speed [13, 268, 99, 248]. Whereas the high dimensionality of this data generates, with high probability, *unique* 'fingerprints' of a user, these are not necessarily *stable* in time (as, for example, users may access the service from multiple or new systems, may update software configurations, or authenticate from different locations). Depending on the computed risk score for that transaction, the authenticating service may grant access to the user with only a valid password (if the risk level is low), or require additional authentication factors (e.g., codes sent to associated email accounts, SMS verification) or even deny access for higher risk levels [268, 181]. This mechanism relies on the assumption that attackers cannot systematically re-create the profile of the victim, unless the attacker is already in control of a user's system.

Following the implementation of RBA techniques across critical services, adversaries developed sophisticated solutions aiming to impersonate the user profile of the authenticating user. Recent literature has shown that phishing kits have developed capabilities to obtain user profiles that can then be re-used by the attacker; similarly, recent malware has been specifically engineered to report user activity back to the attacker [248]. In particular, Thomas et al. [248] highlight the improved capabilities of phishing kits in collecting information related to victims, including geographical location, browser metadata, and answers to security questions; they found that attacks relying on user profile information collected from phishing kits are 40 times more likely to be successful than 'regular' attacks based on leaked credentials. On the other hand, the collection of user profile information does not scale well across users and platforms as user profiles may vary with time, across services, and must be collected by the attacker through additional attack means (e.g., phishing).

5.2.3. Analysis of current attack strategies

Attack capabilities

From the analysis above, we identify six capabilities required to systematically bypass RBA.

Password authentication. At the very minimum, an attacker needs the authentication credentials of the victim.

User profiling. To attempt to circumvent RBA systems, an attacker should have an accurate measurement of the victim's profile/fingerprint for that platform.

Multi-platform. The attacker may need to access multiple web platforms to bypass some MFA controls (e.g., tokens or OTPs sent to an email account of the victim). Authentication credentials and user profiles need to be collected for these additional platforms as well. The capability of impersonating the victim on multiple platforms further increases the attack surface in the scope of the attacker.

Profile updates. User profiles are unique but not necessarily stable. For example, a user may update a password, change software configuration, or access the service from a different geographical region. These changes may invalidate previously collected profiles for that user, which may therefore require updating.

Infection infrastructure. The attacker requires an infrastructure to infect users, and collect and update the collected user profiles. This has to be maintained as defensive capabilities evolve (e.g., blacklisting of an employed phishing domain) and may require the acquisition of external services (e.g., for an infection update [113, 47]).

Automated profile enforcement. Once a profile is collected, the attacker needs to enforce it when authenticating on the platform. Whereas some aspects of the profile are easy to reproduce (e.g., user agent, screen resolution), others are not (e.g., installed fonts/plugins, keystroke speed, mouse movements, etc.). As profiles change across users and platforms, attackers likely need a system capable of enforcing those profiles in an automated fashion.

Analysis across attack strategies

Kurt et al. [248] identify three main strategies for impersonation attacks. Table 5.1 provides an overview of their capabilities.

Leaked credentials. credentials derived from data breaches on a platform. Leaked credentials are generally traded in bulk in underground forums; the leaked data oftentimes only contain associations between usernames and (hashed) passwords, with no user profile information. The data is static and if a user changes the password, the information owned by the attacker loses all value. As the leak concerns only one platform (and multiple leaks are likely unrelated to each other), cross-platform attacks against one user are not enabled by this attack strategy. However, password-reuse attacks may provide the attacker with access to additional platforms on top of the one that suffered the leak.

Phishing kits. attackers can employ kits to deploy phishing websites aimed at stealing user credentials. As users directly interact with the phishing kit, user profiling can be achieved by injecting fingerprinting code in the phishing webpage [248]. The profiles derived through phishing kits are however limited to only one occurrence of the authentication (on the phishing website) and may be incomplete or inaccurate. For example, the employment of password manager software may hinder the realism of the derived fingerprint (e.g., in terms of input time or user behavior on the page) when compared to the one measured by the original platform. To achieve multi-platform capabilities, the attacker must develop or acquire a

Table 5.1: Overview of impersonation attack capabilities.

● indicates full systematic capability; ◑ indicates systematic capability only after specific engineering effort from attacker; ○ indicates no systematic capability.

	Leak	Phishing kits	Malware	ImpaaS
Password auth.	◑	●	●	●
User profiling	○	●	◑	●
Multi-platform	○	◑	◑	●
Profile updates	○	○	●	●
Infection infrastructure	○	○	◑	●
Automated profile enf.	○	○	○	●

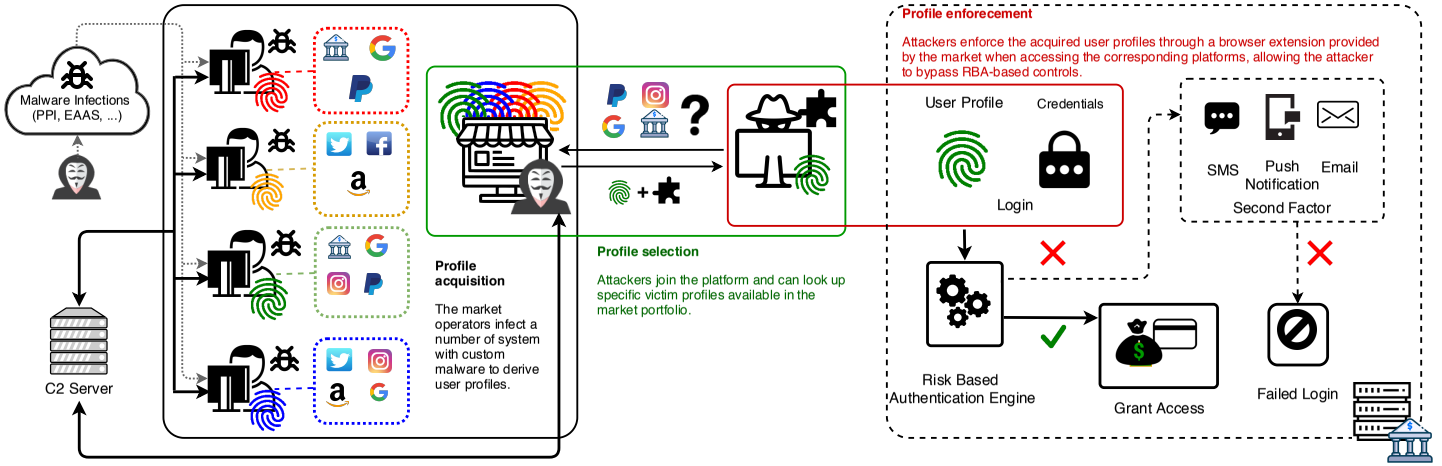
phishing kit for each of the phished platforms, and collect the relevant data through separate attacks against the same user.

Malware. the attacker has access to the system through a keylogger or trojan/bot. This requires the attacker to either purchase/rent the infected system [113], or create the infection themselves (e.g., through malware attached to a phishing email, or through Pay-per-Install services [47]). Due to the specificity of the attack, custom malware is likely needed to collect and update the profiles. As the attacker is virtually already in full control of the user system, they can collect user profiles related to any platform accessed by the victim. However, due to the position of the attacker, most of the impact (e.g., email access or web session hijacking) can be achieved through malware without the need to collect the user profiles to then replicate them at a later stage.

5.3. The Impersonation-as-a-Service Model

In this chapter, we describe evidence of a new emerging attack model, namely *Impersonation-as-a-Service* (`IMPaaS` for short), and the criminal infrastructure supporting it.

`IMPaaS` directly addresses the main limitations of the 'traditional' impersonation attack strategies highlighted above by moving the acquisition and enforcement of victim profiles from an *ad-hoc* process to a *systematic* one. An overview of the comparison between `IMPaaS` and current vectors for impersonation attacks is summarized in Table 5.1. Figure 5.1 provides a bird's eye view of the attack process, from profile *acquisition*, to *selection* and *enforcement*. `IMPaaS` operators rely on widespread malware infections to acquire 'user profiles' globally and provide these profiles as 'goods' via the underground economy through a dedicated marketplace. As a result, attackers can acquire systematic access to a large set of user profiles spanning multiple platforms (social networks, email, corporate accounts, banking/cryptocurrency, etc.), alongside associated credentials and cookies; attackers can select the profiles they are most interested in based on a number of features, including the geographic location linked to the profile, the platforms for which impersonation data is available, amount of stolen cookies, date of profile acquisition, and others. The user profiles available



By only owning the credentials of the victim, the attacker cannot bypass the MFA as the Risk-Based Authentication (RBA) system will detect an anomaly in the profile of the authenticating user. By relying on *Impersonation-as-a-Service* (`IMPaaS`), the attacker can reliably impersonate that profile by providing the values the RBA system expects for that user. `IMPaaS` obtains user profiles from a (large) botnet, and provides them in *bundles* as user profiles. An attacker purchases a user's profile(s) on the `IMPaaS` platform together with a browser extension that, provided the victim's profile as input, reproduces it when accessing a service.

Figure 5.1: Diagram of Impersonation-as-a-Service operations.

on the IMPaaS platform are automatically updated by the underlying infrastructure (e.g., as users change software configuration, or update passwords); furthermore, the attacker can easily enforce and switch across the acquired user profiles by means of a dedicated browser extension provided by the IMPaaS platform, effectively commodifying the systematic impersonation of Internet users at large across multiple platforms.

Profile acquisition

The IMPaaS infrastructure is fueled by a botnet whose goal is, rather than solely collecting credit card information or banking credentials, to provide the information needed to replicate the user profiles of the infected victims across the online platforms on which affected users are active. The malware distribution is independent of the IMPaaS model: it can be delivered through phishing campaigns, targeted attacks, pay-per-install [47] or exploitation-as-a-service infrastructures [113]. Through the chosen attack vector, the attacker installs on the victim system custom malware engineered to collect user credentials and cookies from the victim's browsers; the custom malware further collects a large set of technical and (user) behavioral information that can be replicated, by means of the infrastructure itself, to fully emulate the user; these include the fingerprint(s) of the victim's browser(s) and other behavioral metadata that uniquely identify the user, such as network activity, browser history, cookie data, and interactions with the user interface of the platform. As profiles are fetched through a persistent malware infection, the infrastructure can provide updates of the profile data and credentials for each affected user. The harvested profiles and the respective updates are then pushed to the IMPaaS servers.

Profile selection

An IMPaaS operator provides the harvested user profiles to interested attackers via a dedicated marketplace. The marketplace provides an overview of the characteristics of the collected profiles available for purchase, such that the attacker can select which profiles best fit their goal by searching for victim profiles that show specific features, such as a certain geographic location, web services for which stolen credentials are available, presence of cookies, etc. Albeit less targeted than allowed by a spear-phishing attack scenario, the selection procedure allows for a high degree of precision on the characteristics and/or environment of the user. For example, by browsing through the available credentials it is possible to identify users operating in a specific environment (e.g., a specific corporation, university, or other organizations), or with profiles on platforms of interest to the attacker. Once an attacker has identified their victim(s), the attacker can then proceed to buy the selected profiles. This can be achieved through the usual payment methods adopted in the cybercrime markets, such as via cryptocurrency payments to the marketplace, and/or by relaying the payment through a third-party escrow service. Importantly, as each profile can be purchased individually, the IMPaaS platform is in the position of removing purchased profiles from the marketplace listings, thus potentially reassuring the customer that they are the only one (next to the platform operators) with access to that profile.

Profile enforcement

The IMPaaS platform provides their customers with a customized software bundle that includes a custom browser (based on open-source projects) and a browser extension that allows attackers to fetch and 'enforce' the purchased user profiles during the attacker's browsing session on that platform. Based on the profiles selected and purchased by the attacker, the software provided by the platform recreates a browsing environment that replicates the victim's environment by instantiating exact copies of the stolen cookies and user credentials, and by spoofing other information on the victims' systems (e.g., installed fonts/plugins, browser agent, …). Further, the profile enforcement system provides cookies that embed behavioral metadata derived from the victim [49] without requiring explicit action from the attacker, and it provides SOCKS5 proxy solutions to spoof the usual geographic location of the victim.

5.4. Characterizing ImpaaS in the Wild

In this section, we describe the operations of an emergent, invite-only IMPaaS platform, Genesis Market. The platform has operated since late 2016 and grew considerably, in terms of available user profiles, in 2019. At the time of writing, Genesis Market provides approximately 260′000 (and growing) user profiles available for impersonation attacks against Internet users worldwide. Genesis Market is a Russian IMPaaS platform reachable from the surface web. This platform is, to the best of our knowledge, the first, large IMPaaS operator operating in the underground. On Genesis, a user profile contains information coming from user systems infected with a credential stealer custom malware acting as a man-in-the-browser. The custom malware enables the exfiltration of cookies, credentials, and sniffing of keystrokes, alongside additional environmental and device information that uniquely characterize the user. The IMPaaS platform states user profiles are updated and pushed to the attacker's system in real-time, and that sold user profiles are removed from the listings of profiles available for purchase, although this is difficult to verify empirically, and ethically[2]. An overview of the profile characteristics is provided to browsing customers; profiles with specific characteristics can be searched through the marketplace interface. From the platform, it is possible to access the list of bought profiles and download the related fingerprint. Further, Genesis provides their customers with a custom Chromium-based browser plugin and a pre-built version of Chromium for both macOS, Linux and Windows. This bundle can be accessed only after having bought at least one user profile on the platform. The plugin comes with the capability of loading fingerprints previously obtained from the acquired profiles and can tunnel the traffic through an attacker-specified SOCKS5 proxy to spoof a victim's geolocation.

Malware customization

The latest known custom malware employed by Genesis Market is based on the AZORult malware [66, 105, 39]. Genesis reports a recent update (Nov 2019) in AZORult address-

[2]A proposition is to infect one's own system and purchase back the generated profile to verify its disappearance. As the malware employed by the platform is custom, reproducibility is non-trivial. See also the 'Malware customization' paragraph.

ing changes introduced in the Chrome browser that appear to have affected the malware functionality. Confirmation of massive phishing campaigns in that period associated with AZORult come independently from Kaspersky and other researchers [105, 39, 146]. Note that, start of 2020, AZORult was abandoned by the market in favor of a new (and, at the time of writing, still unnamed), custom malware. Due to the changing nature of the adopted malware, we here only provide a high-level overview of AZORult operations from samples available (at the time of data collection) in the underground and malware repositories. For our analysis, we replicated the latest three versions of AZORult (at the time of writing 3.3, 3.4.1, and 3.4.2) in a virtual environment, with the aim of evaluating its overall functionalities and their relevance to Genesis Market. Malware customization happens through two modules, namely the *builder* and the *C2 server*. The *builder* has the purpose of generating the custom build of AZORult including the URL of the C2 server. The *C2 server* module is a ready-to-deploy web service providing an overview of the harvested data and a page for setting up the features of the malware; these features are user-defined, and include the collection of browser history, saved passwords, cryptocurrency client files, Skype history, a customizable regexp-based file grabber targeting user-defined folders on the infected host, and an additional setup for the deployment of a second stage infection on the victim system: as AZORult removes itself from the system after execution, the second-stage mechanism can allow Genesis Marketoperators to obtain persistence on the infected system and further refine the data collection (e.g., to harvest behavioral data over time, see profile updates analyzed in Section 5.5).

5.4.1. Platform infiltration

Access to Genesis Market is invite-only, and a valid account is needed to access the listings of available user profiles. Access to the registration procedure is provided through invite codes available to members already active on the platform, provided they spent at least 20 USD in purchased user profiles. To gain access to Genesis Market, we probed several underground forums in which we have a pre-existent foothold, and we identified users that claim to be involved with the market. As recent evidence suggests that underground platform operators are actively monitoring and blacklisting 'rogue' accounts (e.g., performing scraping activities) [Campobasso8], we aimed at the collection of several valid accounts prior to data collection to distribute the activity and have 'backup' identities to use if some of our accounts were to be blacklisted. Our search led us to six members in `Torum` and one member in `Crdclub` (who claimed to be one of the operators of Genesis Market) that were offering free invite codes between December 2019 and March 2020. We contacted them through the private messaging facility of the forums as well as on the messaging board and obtained valid invitation codes from three of them in `Torum`. In the case of `Torum`, we relied on a single account, considering the private nature of the conversation and the impression that those members were merely customers of the market. From `Crdclub` we gained access to an additional eight valid invitation codes using separate (and active) identities on the forum, for a total of eleven Genesis Market accounts overall. The process was eased by the free registration policy of `Crdclub`, which allowed us to create additional identities to those already in our control. To minimize suspicion, with each account, we interacted in the newbie sections of the market and asked for further information about the offered products in the market-

place sections of `Crdclub`. Albeit without a clear and set strategy, we broadly defined a 'personality' for each of the created profiles; some used a more informal language with slang, while others had a more professional tone; some profiles would have typos in their posts, while others featured recurrent phrases at the beginning or the end of posts. When a profile counted around 6-10 messages, and after at least 2 weeks from the registration, we applied for an invite. In no instance, the invite was denied.

5.4.2. User profiles on Genesis Market

Genesis Market offers an overview of the available profiles, highlighting the information bundled in that user profile. A view of the interface accessed by attackers is provided in Figure A5.2 and Figure A5.1 in the Appendix. It is worth noting that, whereas Genesis' listings do not readily provide identifying information on the user, the information available on a listing is detailed enough to identify users operating in specific target environments such as a specific organization (e.g., to then perform lateral attacks [125]). Genesis distinguishes between the following information in a user's profile: cookies, resources, and fingerprints.

`Cookies`. Cookies captured by the custom malware and available for injection toward the respective platforms once the user profile is purchased and enforced by the attacker.

`Resources`. `Resources` are collections of data derived from keylogging activity and probing of the browser's local resources, such as the database of stored passwords, and browser history. Some well-known resources (e.g., related to social media platforms, home banking, etc.) are highlighted as `known resources` by the platform, suggesting that the type of extracted `Resources` is an important piece of information for the attacker to consider. A resource can include multiple data reporting login credentials, answers to security questions, detailed balance info for bank accounts, credit/debit card numbers, and holder details. Genesis Market states that the malware extracts `Resources` from infected systems through three main modules: `FormParser` reads the contents of the form data inputted by the user; `SavedLogins` gathers credentials saved in the browser's local database; `InjectScript` implements code injection on the victim's browser on behalf of the attacker, but its operation is unclear and most of the listed profiles do not appear to rely on it.

`Fingerprints`. `Fingerprints` provide a collection of the features exposed by a browser when interacting with RBA systems, ranging from technical metadata (user-agents, browser version) to more finely grained features (geolocation, latency, system language, fonts installed, website device access permissions, etc.)[3]. Depending on the specific RBA implementation, a service may probe a specific subset of the features characterizing a browser or system. Differently from `Resources` (which are tied to a specific service, e.g., a username/password combination on Amazon), the features collected in a Genesis's fingerprint are not bound to a specific service, but to the browser environment itself (e.g., available system fonts, or installed plugins). Therefore, these constitute a *pool* of features that can be

[3]Whereas a full list of the probed features is not available from Genesis Market nor from our analysis (see Section 5.4), a number of commercial and free solutions could be employed by the Genesis Market's malware to implement reliable fingerprinting of the infected systems.

requested by any service, when available. Genesis Market distinguishes between two types of available `Fingerprints`:

1. **Real fingerprints**: these are directly collected from the victim's device, providing an accurate identity of the impersonated device; albeit rarely available in bots, they appear to be sought after by market users;
2. **Synthetic fingerprints**: these fingerprints are generated on the basis of the data collected by the malware. However, accurate 'synthetic' fingerprints cannot be generated without user data (e.g., system fonts, plugins installed in a browser, etc.). For this reason, we consider the availability of `Resources` *and* of browser data in a user profile as an indication that the malware is in the position of collecting the necessary data to generate a reliable synthetic fingerprint.

5.4.3. Data collection strategy

To collect data on Genesis operations, we first consider a number of structural limitations at the core of our sampling strategy:

`Lim1` To avoid disclosing our identity to the market's operators, we perform the scraping via TOR. This poses technical limits (as well as ethical concerns) for bandwidth usage.

`Lim2` We have a limited number of accounts to perform our measurements; aggressive probing risks exposing our accounts to the market's operators, and lead to blacklisting.

`Lim3` Information on `Resources` cannot be accessed in bulk via an API or other requests to Genesis Market, but rather have to be requested in limited bundles with separate requests. This explodes the number of requests necessary to obtain `Resources` information on all user profiles on Genesis Market.

To address `Lim1` and `Lim2`, we employ an *ad-hoc* crawler. Initially the crawler was set to work $\approx$ 24h/day issuing, on average, 15 requests per minute; despite the relatively low requests volume, this strategy led two of our accounts to be blacklisted, suggesting that Genesis Market operators may be employing network monitoring solutions to avoid measurement activities. Following the takeaways from Chapters 3 and 4, we progressively reduced the crawling activity to $\approx$ 6h/day. In the process, an additional three accounts were banned, for a total of five banned accounts. It is interesting to note that three of the five blocked accounts were not linked to each other in any way[4], suggesting that market operators have kept their crawling-detection efforts high during our activities. To mitigate this problem, we employed different strategies to access specific pages and resources to crawl on Genesis Market: as already noted in Chapter 3, accessing URLs directly (as opposed to via website navigation) may generate anomalies in crawler monitoring systems. For this reason, we operationalized all crawling activities through browser instrumentation and configured the crawler to mimic activity patterns compatible with those of a human user (e.g., timeouts between requests proportional to the length of the visited webpage, taking breaks, ...), following the guidelines

[4]The first two accounts were obtained from a single member of Genesis Market and active on `Torum`. The other accounts all came through invitation codes generated by either different market members, or released by Genesis Market operators themselves to different and unrelated accounts we control on `Crdclub`.

drawn in Chapters 3 and 4. With this final setup, we finally managed to silently crawl the market avoiding the detection and ban of the remaining accounts in our possession.

While necessary, the above strategy makes it impossible to gather complete information on `Resources` due to the exploding number of requests (`Lim3`). This results in two datasets:

- `Full database` includes information on approximately $262'000$ user profiles on Genesis Market, including (infection, update) dates, prices, number of browsers for which resources are available, number of collected fingerprints for that user profile, and number of stolen cookies.
- `Sampled database` adds `Resources` information to a random selection of approximately 5% ($n = 13'512$) of the user profiles available on the market.[5]

The collected data is available for sharing to the research community at `https://security1.win.tue.nl/doku.php?id=artefacts#data_sharing`

Analysis procedure

The data analysis in Section 5.5 is split into two sections: in Section 5.5.1 we provide an overview of the data collected in the `Full dataset`, and characterize Genesis Market operations by looking at its evolution, victim profile characteristics, profile updates, and pricing; in Section 5.5.2 we analyze the distribution and effect of `Resources` on pricing, as reported in the `Sampled database`. Standard sanity checks (e.g., on the regression results presented in Section 5.5.2) are performed on all analyses. Reported logarithms are natural logarithms unless otherwise specified.

Manual resources classification. To factorize the type of resources reported in `Sampled database` in the analysis, we provide a classification of each resource in one of six categories. Table 5.2 lists the employed categories and their corresponding definitions. The classification was done manually by one of the authors over 454 unique platforms for which `Resources` are reported in the dataset. The other author independently classified a random sample of 100 platforms, reaching an agreement score of 89%; after review, conflicts were resolved and the classification was updated accordingly. Additional random checks did not reveal any remaining mismatch.

Ethical considerations and limitations. No personally identifiable information is reported in our dataset. IP addresses of victims are masked on the platform, and no detailed information about the victims is available without purchasing a user profile. For obvious ethical concerns, we did not purchase any. Whereas this limits our analysis in that we do not have access to the software bundle provided by Genesis, and cannot ascertain in detail the quality or operative aspects of the `IMPaaS` service provided by Genesis, we are in the position of providing a full evaluation of the data is available to the attacker when browsing for victims.

[5]This fraction was originally set to 10%, however approximately half of the selected profiles were removed from the market during the data collection process.

Table 5.2: Categories of resources.

Category	Definition	Examples
`Services`	Platforms providing the delivery of physical (e.g., goods, postage, etc.) or digital (e.g., content streaming, cloud, mail, etc.) services to final users.	Google, PosteItaliane
`Social`	Platforms to share user generated content.	Twitter, Skype
`MoneyTransfer`	Platforms enabling direct payments between people or organizations using traditional payment circuits.	CreditUnion, Transfergo
`Crypto`	Platforms enabling direct payments between people or organizations using cryptocurrency circuits.	Coinbase, Bittrex
`Commerce`	Platforms whose sole purpose is to purchase or book goods/services from one or multiple vendors.	Amazon, SaldiPrivati
`Other`	Platforms that do not match any of the previous categories.	Auth0

5.5. Data Analysis

Table 5.3 provides an overview of the collected datasets.

`Full dataset`. The data collection spans from Dec 2017 to March 2020, involving approximately $262'000$ user profiles. Most user profiles available on Genesis Market target only one browser, with the top 5% targeting three browsers. Only 35 user profiles report data for more than six browsers in our data. Cookie distribution is similarly skewed. Profiles are distributed globally across 213 countries[6], and prices range from 0.7 to 96 USD; 50% of the profiles cost at most 5 USD, whereas the priciest 5% are priced above 20 USD.

`Sampled dataset`. This dataset reports data on 5.2% of the profiles available on Genesis spanning from March 2018 till March 2020 ($n = 13'512$). For this dataset, we collected detailed information regarding the available `resources`. As this is a random sample, values are distributed similarly to `Full dataset`. Additionally, we extract information on the number and type of resources available for each profile. The average profile has upwards of 30 resources; most resources are of type `Services`, whereas `Social` and `Commerce` are less common. `Crypto` and `MoneyTransfer` resources appear to be the least numerous in a profile.

5.5.1. Overview of Genesis Market's operations

To provide an overview of the `IMPaaS` operations conducted in the market we first look at the full dataset summarized at the top of Table 5.3. Interestingly, we find that approximately 12% of all profiles are *not* associated with a browser on the victim's system.[7] As

[6]Although there are only 195 recognized countries worldwide, Genesis reports ISO 3166-1 codes, which do not distinguish sovereign nations from dependent territories.

[7]Note that all profiles without browser data also do not, by definition, report any data on cookies or fingerprints.

Table 5.3: Summary statistics of the collected datasets.

	Variable	min	mean	max	sd
Full dataset (n : 262'080)	N° browsers	0	1.58	10	1.02
	N° cookies	0	1719.56	125198	1773.57
	N° *real* fingerprint	0	0.06	17	0.32
	Date infection†	12-12-17	20-11-19	16-03-20	157.81
	Date updated†	01-02-18	23-11-19	15-09-19	156.69
	Country	*char*	*char*	*char*	*char*
	Price (USD)	0.7	7.83	96	7.62
Random sample (Dec'17-Mar'20, n : 13'512)	N° browsers	0	1.57	8	1
	N° cookies	0	1782.01	26981	1735.24
	N° *real* fingerprints	0	0.12	9	0.54
	Date infection†	28-03-18	08-01-20	16-03-20	40.04
	Date updated†	12-11-19	14-01-20	13-06-20	37.62
	Country	*char*	*char*	*char*	*char*
	Price (USD)	0.7	8.84	63	8.17
	N° resources	0	31.13	1322	46.63
	Crypto	0	0.07	18	0.6
	Money Transfer	0	1.66	385	6.23
	Social	0	7.95	1322	18.73
	Services	0	16.64	560	24.44
	Commerce	0	4.66	296	11.12
	Other	0	0.15	16	0.88

†: dates are reported in dd-mm-yy format. *sd* in days.

these profiles do not allow for impersonation attacks under the `IMPaaS` model, we remove those from further analysis. Relative to the number of user profiles, the number of available *real* fingerprints is surprisingly low, with only 4.3% of the available profiles having at least one. Note however that this refers only to `Real fingerprints` collected by the malware, not the `Synthetic fingerprints` that can be synthesized from user data (ref. Section5.4.2). Nonetheless, this suggests that (real) fingerprints, available browsers, cookies, and resources could be the driving force behind Genesis Market's activities.

Figure 5.2 provides an overview of the geographic distribution of the user profiles available on Genesis Market and their median price per country. Most of the profiles belong to users in the United States of America and Europe, with a high fraction of EU countries showing volumes similar to those of the US. Users in Asian and African countries are comparatively less affected. As commonly seen in Russian cybercrime markets [15], Genesis Market does not provide profiles for users in Russia, Ukraine, Belarus, and Kazakhstan (CIS countries). Furthermore, with the exception of Chad and the Central African Republic, the CIS countries appear to be some of the only unaffected countries, globally. Overall, median prices appear to vary from country to country rather than at a macro-regional level. For example, EU median prices seem to be higher in Spain ($m = 9.55, sd = 9.07$) and GB ($8.3, sd = 7.5$) than in Germany ($m = 7.21, sd = 8.21$) or Finland ($m = 6.96, sd = 6.68$). A set of Wilcoxon Rank-Sum tests evaluating the alternative hypothesis that SP and GB profile prices are higher than DE and FI ones confirms this observation ($p < 0.0001$). The high median price in Mauritania (26 USD) is caused by only one profile (with no fingerprints, two browsers, and four thousand cookies) available for that country.

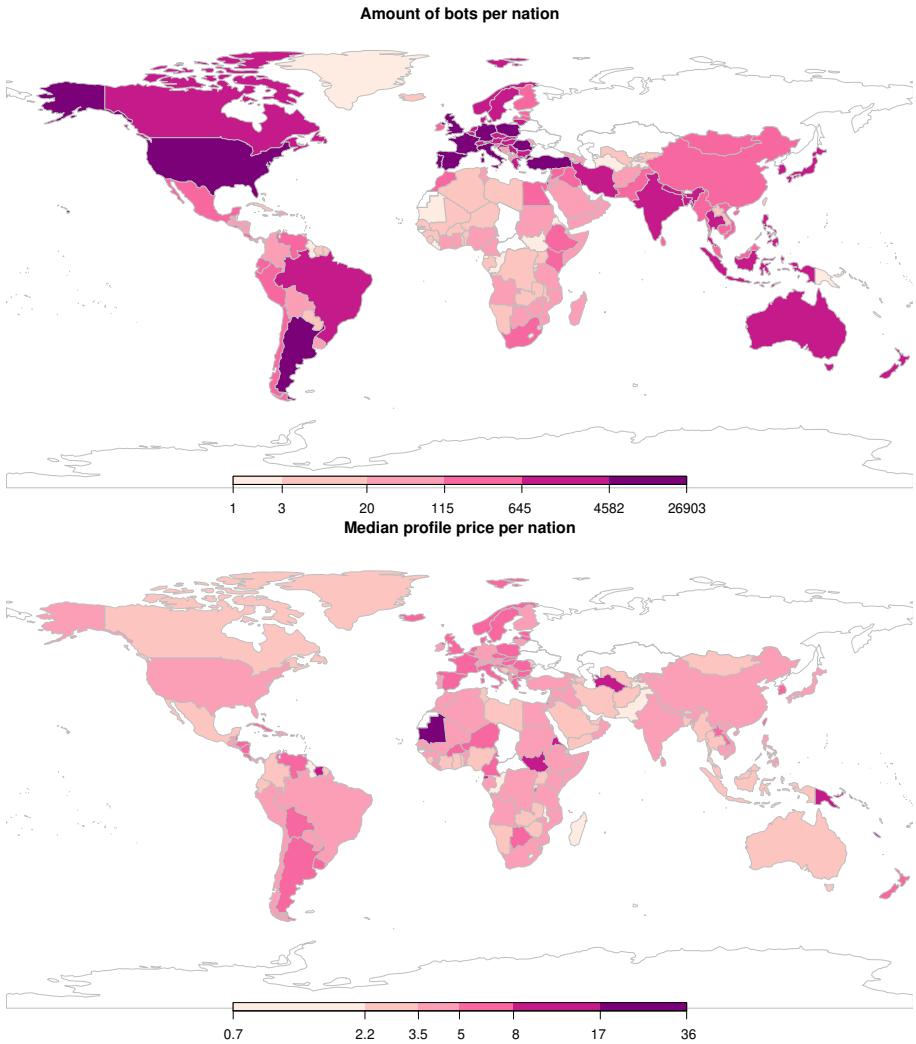


Figure 5.2: Global distribution of user profiles (top) and their median price (bottom) on Genesis Market

The rate of appearance of new and updated user profiles on Genesis Market is depicted in Figure 5.3. A clear upward trend in terms of the number of available profiles is visible, with a large jump in available profiles in November 2019 (coinciding with the 2019 spike in phishing campaigns distributing AZORult [105, 39, 146]). Overall, in Figure 5.3 we observe a sustained rate of new (black bar) and updated (orange bar) profiles, suggesting that the platform is systematically updating existing profiles while adding new ones to the platform portfolio. We further investigate the time passing between the time of infection and (last) profile updates; Figure 5.4 shows the boxplot distribution of time passed between the infection and the last update received by the platform, plotted against the date of installation; in red, the upper

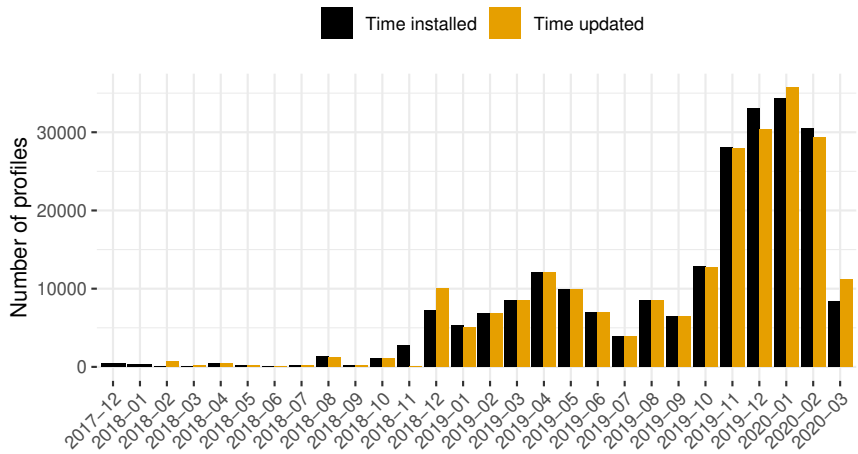


Figure 5.3: Progression of available user profiles over time.

5

bound of the maximum possible time in between. Overall the distribution appears relatively stable, with a median update time ranging between ten hours and four days. Unsurprisingly, recently acquired profiles are updated only after a few hours from acquisition; overall, the distribution suggests that profiles are kept updated on average over an extended period of time, ranging from a few days to several months at the extreme of the distribution.

Analysis of profile values

Figure 5.5 reports the moving average of user profile prices as a function of time. The value of the traded profiles steadily increases as time passes, a signal of growth of the platform. In particular, profile prices seem to have doubled since November 2019, perhaps as an effect of the updated malware released in that period discussed at the start of this Section. Figure 5.6 reports the relation between the number of available `Fingerprints` in a profile and its price. The effect of an increased number of available fingerprints is, albeit positive, very limited. The average price seems to stabilize around the median value of 5 USD regardless of the number of fingerprints available in the profile, suggesting once again that other variables could be at play. We find no correlation between the number of available browsers, cookies, and prices. This is not surprising, as these dimensions express little in terms of *which* identities of the victim the attacker may affect.

To further look at factors that may determine the value of a profile, we look at the impact of the geographic location to which the profile is linked. To do so, we investigate the relation between (log-transformed) profile prices and the wealth of the country in which the profile is located, expressed in terms of (log) GDP per capita (as reported by the World Development Indicators [30]). The intuition is that the more 'valuable' a target is perceived to be, the greater the value of the corresponding profiles might be. Figure 5.7 reports the analysis. A positive and statistically significant correlation emerges, suggesting that profile prices are

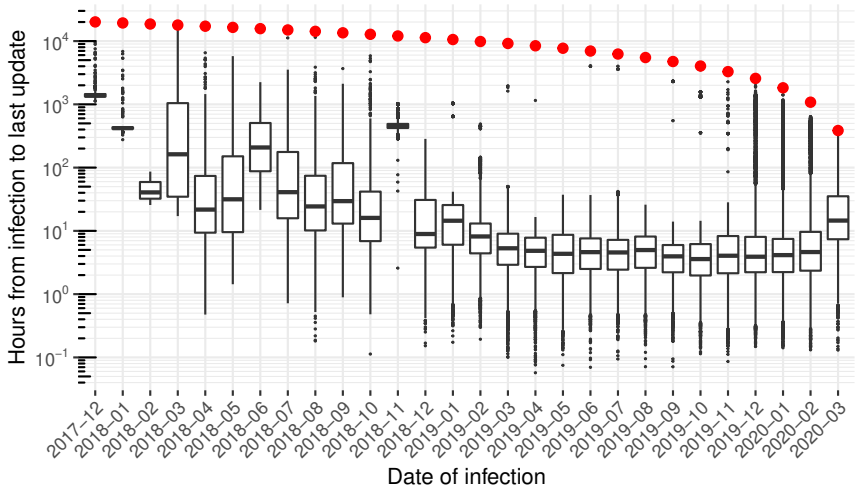


Figure 5.4: Time between infection and last profile update (in log scale).

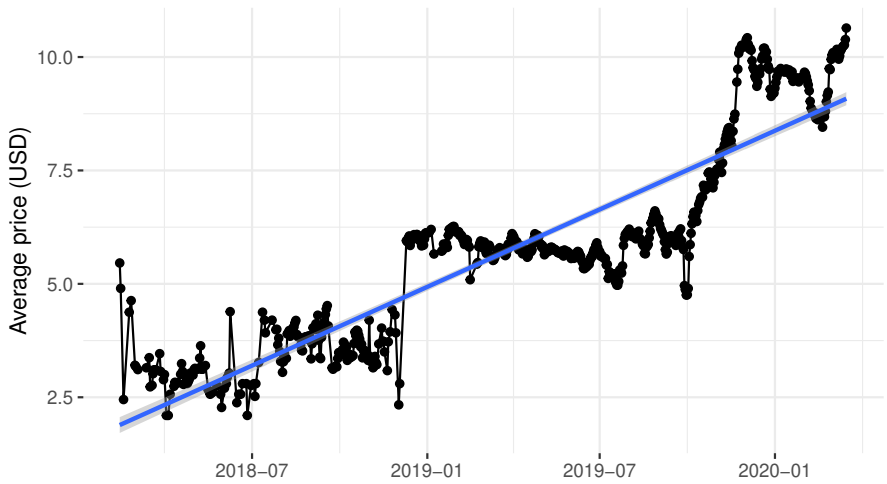


Figure 5.5: Weekly moving average of user profile prices.

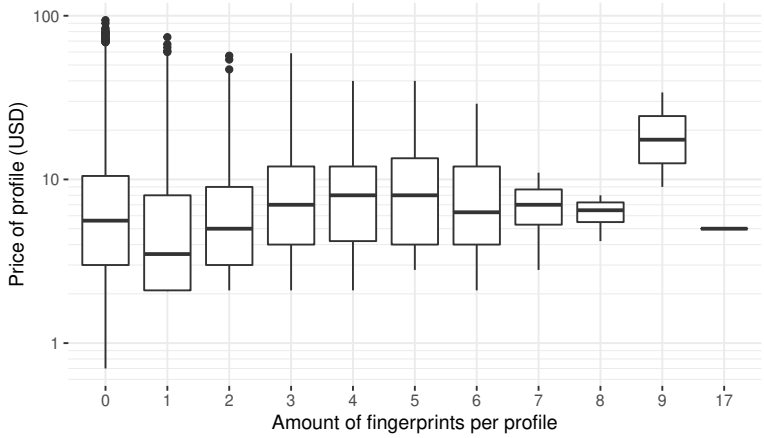


Figure 5.6: Relation between the amount of fingerprints available and the average price of user profile (in log scale).

Table 5.4: Type of resources per user profile.

Resource type	no. profiles ($n = 12'052$)
Cryptocurrency	236
Money Transfer	3109
Commerce	5'066
Social	8'111
Services	11'167
Other	548

indeed correlated to the wealth of the respective country, perhaps a sign of the perceived value of that user profile ($corr = 0.4, p < 0.001$).

We note that some user profiles on Genesis Market appear to be discounted at a rate of 30%. We do not find a clear-cut effect explaining which profiles are likely to be discounted (Table A5.1 in the Appendix).

5.5.2. The impact of `Resources` on profile pricing

We first look at the distribution of resource types in the `Sampled dataset`. As for the `Full dataset`, we remove from further analysis profiles that are not associated with at least a browser of the victim's system and, in addition, profiles that do not contain any stolen resource, limiting the size of the dataset to $n = 12'052$. Table 5.4 provides an overview of the distribution of user profiles per category. Note that a profile can have resources that belong to more than one category. Overall, `Services` is the most commonly available resource type across user profiles. Resources in the `Social` and `Commerce` categories are also common, with respectively about 70% and 40% of user profiles with resources in these

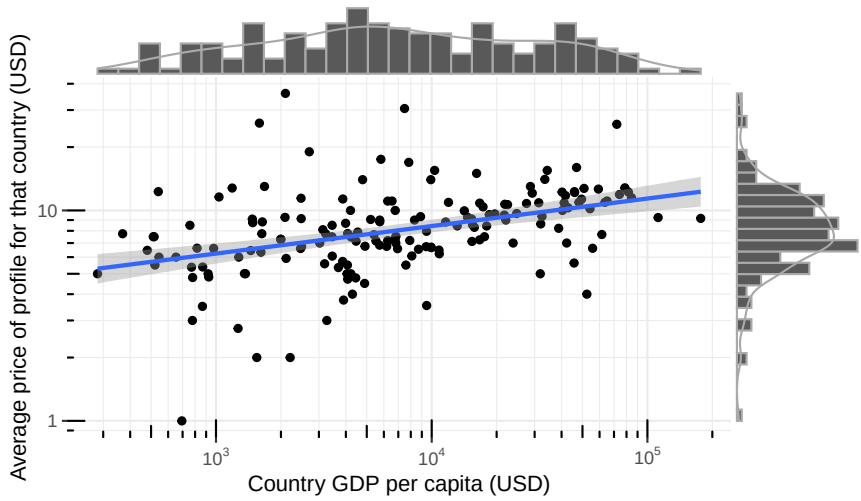


Figure 5.7: Relation between GDP per capita and average price of user profiles in that country (in log scale).

categories. Approximately 25% of the profiles have data for banking and payment accounts; by contrast, less than 2% of user profiles have resources in the `Crypto` category. Only 4.5% of the resources in our dataset were classified as `Other`, indicating that the proposed classification covers the vast majority of resource types in Genesis Market.

Figure 5.8 provides a first overview of the relation between the number of resources available in a profile and the associated price. A clear correlation emerges. The depicted linear log-log relation indicates negative marginal returns for each added resource, meaning that every additional resource added to a profile provides an increasingly smaller, albeit positive, added value to the profile. Further exploring the impact of resources on pricing, Figure 5.9 shows the impact of the presence of resources in any specific category on the value of a user profile. Note that, because each profile can contain resources of more than one category, one cannot isolate the relative importance of each category here. However, the comparison shows how, on average, a profile that contains (also) resources in that category is priced versus other profiles that do not have it. This is meaningful as the categories show relatively low correlations (reported in Table A5.2 in the Appendix). On average, profiles that include `Crypto` resources seem to be the most valuable. `MoneyTransfer` and `Commerce` resources belong to profiles of approximately the same value, whereas profiles with `Social` and `Services` are the least valued in Genesis Market. By comparing the relative 'jump' introduced by the addition of each category, one can further evaluate the added value, on average, of having a resource of that type. In this respect, `Other` appears to be the least 'impactful' category, as the appearance of a resource of this type is related to the smallest relative increase in price, on average, in a profile. On the contrary, `Crypto` and `MoneyTransfer` resources cause the highest jump in profile value, passing from a median value of approximately 7 USD to more than 20. Other categories show less extreme price changes. Overall, we find that resources associated with financial platforms and services appear to have the highest impact on the value of a profile, with `Social` and `Services` being the

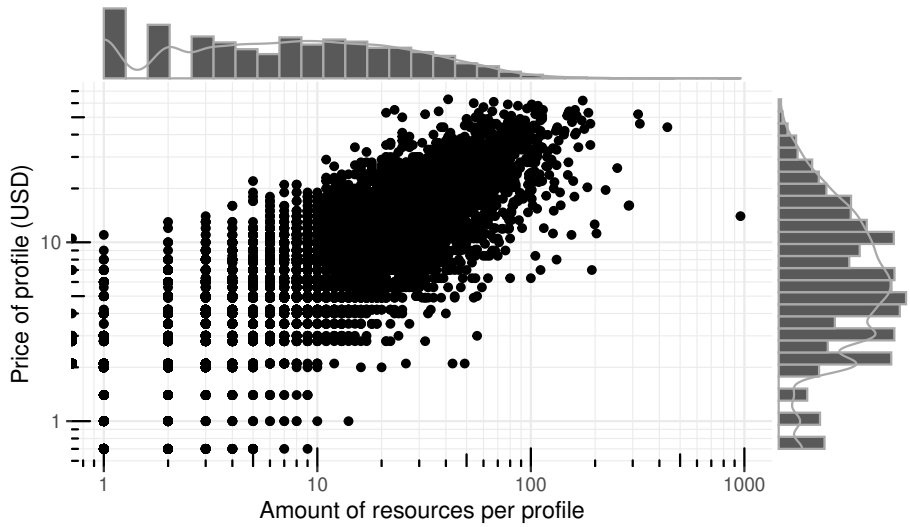


Figure 5.8: Relation between the amount of resources available and the average price of user profile (in log scale).

least valued. On the other hand, the addition of resources of any category appears to have a positive impact on the value of a profile.

To formally evaluate this relation, we build a set of linear regression models to quantify the effect of different profile features on profile values in Genesis Market. To evaluate the effect of each factor independently, and monitor its relation to other characteristics of a user profile, we define the following nested models with response variable $y = price$ (the error term ϵ_i is omitted for brevity):

$$
\begin{aligned}
M1 &: y_i = \beta_0 + \beta_1 \texttt{Real fingerprints}_i \\
M2 &: y_i = ... + \beta_2 \log(\texttt{GDP}_i) \\
M3 &: y_i = ... + \beta_3 \texttt{Crypto}_i + \beta_4 \texttt{MoneyTransfer}_i \\
&\quad + \beta_5 \texttt{Commerce}_i + \beta_6 \texttt{Social}_i \\
&\quad + \beta_7 \texttt{Services}_i + \beta_8 \texttt{Other}_i \\
M4 &: y_i = ... + \beta_9 \texttt{resources}_i
\end{aligned}
$$

where β_0 is the intercept, `Real fingerprints` is the number of fingerprints embedded in that user profile, log(`GDP`) is the natural logarithm of the gross domestic product per capita for the country associated with the user profile, {`Crypto` ... `Other`} are dummy variables representing the presence of resources of the corresponding category, and `resources` is the overall number of resources in that profile (irrespective of category).

Regression results are summarized in Table 5.5. To evaluate the effects of profile characteristics on full prices we remove profiles 'on sale' from the dataset. Table A5.3 and Table A5.4 in the Appendix report, respectively, a full breakdown of the variables' impact on the prediction,

Figure 5.9: Profile price variation according to the presence of not of resources of a certain category (in log scale).

and the regression results for all data points including profiles on 'sale'; both tables report results quantitatively and qualitatively in line with those reported in Table 5.5. Overall, the coefficient estimates appear stable across the models, with the exception of β_2 (log(`GDP`)), which becomes less important on the estimation of the dependent variable `price` as the types of `Resources` are added to the model. The change ranges from an expected increase of 0.2 USD in profile value for every 10% increase in GDP ($\beta_2 = 2.29$ in M2, $2.29 \times \log(1.10) = 0.22$), to a relatively smaller (0.04 USD) price increase when all resource categories are added in the model. This indicates that some resource categories may appear more frequently for high GDP countries than for others; with reference to Table A5.3 in the Appendix, it appears that resources of type `MoneyTransfer` and `Commerce` tend to appear more often in wealthy countries, as most of the effect of the GDP variable disappears when this category is accounted for, while the opposite effect emerges when `Social` resources are included in the model. Additional resource categories have modest effects on the GDP coefficient estimate. As resource categories are added to the model, the impact of the number of fingerprints increases, passing from a 0.55 USD increase in expected profile value for each additional fingerprint in the profile ($\beta_1 = 0.55$) to a 1.31 USD increase estimated by M3. This suggests a positive joint effect of the number of fingerprints in a profile and the number of platforms with resources an attacker can employ to impersonate a victim. All resources have a positive effect on the value of a user profile with `Crypto` and `MoneyTransfer` having the highest impact, increasing the expected value of 13.62 and 8.86 USD respectively when available. Following this trend, `Commerce` shows a relatively large effect as well, increasing the profiles' expected value of 5.06 USD. These findings may not come as a surprise and may indicate that Genesis customers may be primarily aiming at economic profit (supporting insights from observing Genesis customers discussing on a dedicated Telegram channel, see Section 5.6.2 for an informal report). Finally, the effect of the number of resources in M4 is significant and positive; interestingly, its addition decreases the effect of the single resource categories, confirming the intuition that the more platforms

Table 5.5: Regression analysis on prices of user profiles.

	Model 1	Model 2	Model 3	Model 4
β_0	10.41^{***}	-12.11^{***}	-5.57^{***}	-3.70^{***}
	(0.11)	(1.21)	(0.81)	(0.63)
Real Fngrpr	0.55^{***}	0.69^{***}	1.31^{***}	1.11^{***}
	(0.16)	(0.16)	(0.10)	(0.07)
$\log(GDP)$		2.29^{***}	0.46^{***}	0.42^{***}
		(0.12)	(0.08)	(0.06)
Crypto			13.62^{***}	10.12^{***}
			(0.44)	(0.34)
Money Transfer			8.86^{***}	6.20^{***}
			(0.16)	(0.13)
Commerce			5.06^{***}	3.22^{***}
			(0.15)	(0.12)
Social			3.44^{***}	1.68^{***}
			(0.15)	(0.12)
Services			3.95^{***}	2.31^{***}
			(0.29)	(0.22)
Other			4.22^{***}	0.89^{***}
			(0.31)	(0.24)
Resources				0.10^{***}
				(0.00)
R^2	<0.01	0.05	0.65	0.79
Adj. R^2	<0.01	0.05	0.65	0.79
Num. obs.	7123	7123	7123	7123

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

an attacker can impersonate, the higher the value of the profile.[8]

In all, resource types appear to explain the majority of the variance in the model, with MoneyTransfer accounting for a jump in more than 30% in the model (adjusted) R^2 when compared with the previous model. The complete model explains most of the price variance in our dataset ($R^2 = 0.79$), suggesting that the model provides an appropriate description of the features determining user profile values in Genesis Market.

5.6. Discussion

In this chapter, we presented the IMPaaS model as a novel threat enabling attackers to perform user impersonation at scale. IMPaaS is supported by an emergent criminal infrastructure that controls the supply chain of user profiles, from system infection to profile acquisition and commodification. Whereas traditional impersonation attacks relying solely on stolen credentials are greatly mitigated by risk-based and two-factor authentication systems, the capability of seamlessly reproducing a user's 'appearance' to an authentication system

[8]Driven by observations in Chen et al. [49], who identified cookies as having a key role in behavioral fingerprinting practices, we find that in terms of profile pricing the availability of cookies does not show a statistically significant effect (Anova $F1.92$, $p = 0.17$) in our dataset, suggesting that cookies do not play a central role in impersonation attacks as driven by Genesis Market.

allows attackers to systematically compromise accounts of users across multiple platforms.

Whereas Thomas et al. already suggested that user profiling could be used to bypass modern authentication systems [248], in this work we provide evidence of an emergent *as-a-service* impersonation model that appears to be rapidly expanding. The profile value analysis provided in Section 5.5.2 suggests a mature pricing model, which may indicate that the analyzed platform operations are of stable, predictable quality, and likely to expand in number. Overall, the analysis of the available user profiles on Genesis Market and the reportedly widespread adoption of *info-stealer* malware such as AZORult in phishing campaigns [105, 152] provide further supporting evidence of the growth of this threat model.

Our analysis of Genesis Market allows us to further quantify the relative effects of different resources on the value of a user profile. Interestingly, albeit perhaps not surprisingly, we find that profile values show a significant correlation with the wealth of the country (expressed in terms of GDP) associated with that profile; this suggests that attackers looking to impersonate and, likely, monetize user profiles assign a greater value to profiles likely to give access to greater financial resources (e.g., bank balances or valid credit cards). Interestingly, this effect is significantly reduced by the presence of `Commerce` resources in a profile, perhaps due to the prevalence of e-commerce platforms in wealthy countries. Nonetheless, other resource categories have a clear impact on the overall valuation of a user profile, with `Crypto` and `MoneyTransfer` resources driving most of the value. Another potential factor explaining the high price tag of profiles originating from North America could be related to the regulations in the banking system of the USA. As noted by Sinigaglia et al. [235], the number of MFA protocols implemented by North American banks during payments is remarkably lower than those of their European counterparts. According to the same study, it emerges that the regulatory body dictating the technical standards on strong customer authentication enforced in Europe is more stringent than the one in place in the USA. This difference in industry security standards may indicate that financial frauds using home banking systems are easier to accomplish at the expense of North American citizens and that Genesis Market's toolkit could facilitate fraudsters in the process. As a consequence, Genesis Market's operators could be aware of the higher expected value of North American profiles, adjusting the price for those profiles accordingly. *Real* fingerprints (those derived directly from the device, rather than being synthesized by the `IMPaaS` platform using the profile's metadata) available in a profile also add value to the user profile. Our analysis suggests that each *real* fingerprint adds about 0.55 USD to the value of a user profile, and up to 1.31 USD when considered jointly with the available resources, suggesting that the *modus operandi* enabled by `IMPaaS` described in Figure 5.1 is supported by the platform operations.

Importantly, our analysis allows us to put a number on the value of user information to attackers, contributing to the literature on the subject. A user's 'virtual identity' seems to be worth between less than 1 USD and approximately 100 USD. This value changes significantly depending on the wealth of the country where the user (appears to be) located; a rule-of-thumb indication seems to be that for a tenfold increase of the user's 'expected wealth' (approximated by a country's GDP), a profile value increases on average by approximately 1 USD. Cybercriminals seem to particularly value profiles with access to `Crypto` and `MoneyTransfer` platforms, whose prices are respectively 10 and 6 USD higher than profiles with no access to platforms of these types. To put this in perspective, these represent respectively

a 150% and 90% markup over the price of the average profile, a clear sign of the relevance of resources of these types for cybercriminal activities. By contrast, access to `Social` and other services does not seem to be (in comparison) highly valued by cybercriminals.

5.6.1. Implications for victimization

The systematization of impersonation attacks enabled by the `IMPaaS` model allows attackers to select and target specific victim profiles, and to automate the attack procedure by means of dedicated software bundles replicating a victim's browsing conditions on the environment of the attacker. Differently from traditional phishing-based attacks, `IMPaaS` provides an attacker with access to several platforms on which a user is active, effectively allowing the attacker to both mitigate security measures (e.g., by monitoring email for authentication codes or activity notifications), as well as extending the attack surface to different services (e.g., banking, social, etc.).

Attackers leveraging an `IMPaaS` platform can rely on an automated source for credentials to conduct sophisticated attacks at scale. In addition to obtaining access to banking websites, cryptocurrency exchange platforms, and e-commerce websites, an attacker may compromise multiple accounts to gain control over the identity of the victim. The capability of selecting victim characteristics before the acquisition of a profile is also a potential enabler of targeted attacks against organizations or communities for which a victim is an employee or a registered member. The attacker may employ that advantage point to facilitate lateral movement attacks, for example targeting colleagues or family members of the victim by using their legitimate contact details. Furthermore, the attacker could integrate additional information about a victim gathered through the accessed platforms (part of a corporation, subscription to meeting websites, etc.) to further escalate the attack to other victims.

5.6.2. Examples of (alleged) criminal operations enabled by Genesis Market

To informally investigate how attackers are weaponizing capabilities enabled by `IMPaaS`, we collected a number of examples provided by users of Genesis Market on a Telegram channel linked to the platform, to which we have gained access through Genesis. Many attacks reported there appear to focus on `MoneyTransfer` and `Commerce` services. For example, a user shared that they were (allegedly) able to cash out from a US bank using a synthetic fingerprint acquired on Genesis Market, and with the support of a geographically accurate SOCKS5 proxy. The user further suggested to rely on `911.re` as the marketplace where to buy proxies linked to specific ZIP codes and/or ISPs. On a similar line, a second user reported to have managed to issue a new debit card on behalf of the victim, with the aim of cashing it out through ATMs. Interestingly, some Genesisusers report performing multi-stage attacks deployed through the obtained user profiles and exploiting multiple platforms. For example, an attacker describes setting filters to a victim's email mailboxes accessed through the victim's user profile, with the aim of hiding notifications from Amazon related to purchases the attacker made using the victim's Amazon account.

Overall, whereas of course none of these examples can be verified and the threats described above are not new per se, the mix of infrastructural support for profile acquisition, selection, and enforcement enabled by `IMPaaS` opens to the *systematization* of threat scenarios such as the ones described above, on a global scale.

5.7. Chapter conclusion

In this chapter, we presented the emergence of the *Impersonation-as-a-Service* criminal infrastructure, which provides user impersonation capabilities for attackers at large. `IMPaaS` allows attackers to bypass risk-based authentication systems by automatically simulating the victim's environment on the attacker's system. In this study we characterize the largest currently operating `IMPaaS` infrastructure, Genesis Market, by performing an extensive data collection spanning more than 260 thousand stolen user profiles collected worldwide. Genesis Market infiltration and data collection required substantial efforts to collect multiple accounts and needed to fine-tune the data collection as platform operators seemed to monitor crawling activities and blacklist related accounts. From our analysis, Genesis Market emerges as a mature, expanding infrastructure with a clear pricing structure, suggesting a well-established criminal business model. *Impersonation-as-a-Service* represents an additional component of the cybercrime economy, providing a systematic model to monetize stolen user credentials and profiles. Finally, our data collection efforts provide supporting evidence that underground platform operators are actively monitoring crawling activities, and take measures to limit them. This may prevent future research activities and significantly impact the possibility of designing large-scale studies studying cybercriminal online venues. Specific sampling strategies and analysis techniques will have to be devised to further develop research in this domain.

6

Deriving attacker preferences and overall market activity from live market data

This chapter is based on [Campobasso1]:
M. Campobasso, and L. Allodi
Know Your Cybercriminal: Evaluating Attacker Preferences by Measuring Profile Sales on an Active, Leading Criminal Market for User Impersonation at Scale
32nd USENIX Security Symposium (USENIX Security 2023)

In the previous chapter, we assessed that Genesis Market is a mature market posing a serious threat, providing an innovative and effective service for user impersonation at scale. Among its peculiarities, Genesis is a platform selling a single kind of product with a single vendor and lists every profile currently available. The presence of profiles dated close to the birth of the market and the presence of discounted prices for very old ones, together with the verification that disappeared profiles do not reappear in the future (apart from reserved ones, more in the Appendix), led us to assume that profiles offered on Genesis Market disappear after purchase and do not reappear in the future. This creates the rare opportunity to directly measure the supply and demand of an underground market and allows one to draw conclusions on the customer's (attacker's) preferences. This information is particularly precious as it allows us to better understand what are the characteristics of victims more exposed to this threat and to quantify the affected population. Furthermore, we use the information on attacker preferences to model the risk posed by IMPaaS by contextualizing our findings into the larger risk model proposed by Woods and Böhme. Finally, as an effect of our measurements, we are able to provide an estimate of the market revenues.

In this chapter, we will delve into the hurdles of gathering live data about supply and demand from an underground market. Aside from the problems discussed in Part I, we will discuss the additional challenges, unique to this kind of study, which we faced when gathering and producing data on market sales, and how we managed to use the incomplete data we scraped

to obtain solid, representative results. To achieve this, we deployed a series of crawlers to extract live market data and devised a rigorous data analysis pipeline to simulate missing data, interpret attacker preferences, predict sales, and estimate market revenue.

Our study indicates that Genesis Market is indeed a mature threat, targeting the majority of the globe, and that monitoring the economic activity of a market like that is possible; by doing so, stakeholders could better inform their risk models, allowing them to take proactive measures to limit threat and to evaluate their exposure.

Link to the created datasets (available under license):
`https://security1.win.tue.nl/doku.php?id=artefacts#data_sharing`

6.1. Introduction

Studying underground communities can provide important insights into cybercriminal actions and threat levels [15, 238, 14]. In particular, the evaluation of underground markets can help quantify the risk to final users posed by cybercriminal activities. For example, the observation of criminal ecosystems has been employed in research to identify innovative or emergent threats, and the monitoring of trade activity to evaluate their associated impact on final users [36, 14, Campobasso7]. On the other hand, obtaining reliable data from criminal marketplaces is an increasingly challenging activity [257] as platform administrators start deploying anti-crawling measures [Campobasso3] and access control measures vetting accounts requesting access to their community(-ies) [15]. Furthermore, data collected in these underground places is often censored or missing, for example, due to infrastructural failures at certain crawling times. This is particularly challenging for longitudinal studies (of any length) aiming at monitoring market/community evolution over a period of time, measuring differences in outcomes or, for example, product provision [238]. Data is hard to interpret as well, as generally only indirect signals of events are available for inference (e.g., user feedback as a proxy variable for product sales). Exceptions exist for leaked databases, although this generally allows studying markets that have already died or collapsed, oftentimes as a result of the leak itself. In other words, the opportunity to reliably study threat levels posed by active underground markets, their relevance globally and over time, and the overall size of the underlying economy supporting those threats is rare.

6.1.1. Research gap and contribution

In this work, we study a unique data collection of sale volumes and trends on Genesis Market, a leading, invite-only Russian underground platform currently active and operating as the main provider for Impersonation-as-a-Service in the criminal underground [Campobasso7] to evaluate the overall threat levels it poses globally to the Internet users, quantify the size of the underlying market economy supporting these attacks, and evaluate attacker preferences when choosing a profile to purchase. This work's contribution is multi-fold:

1. We present a thorough data collection methodology addressing the key challenges of monitoring the evolution of specific products in the market while avoiding anti-crawler technologies and under the constraints introduced by monitoring closed-access marketplaces. We discuss the necessary trade-offs and present the respective solutions.
2. We devise a robust data analysis methodology addressing uncertainties in the data collection resulting from those trade-offs; the proposed methodology handles high dimensionality data while capturing all variance in the original variables and maintaining full transparency on the relation between dimensions and outcome;
3. We provide an extensive analysis of the size and relevance of `IMPaaS` as a global threat model, estimating volumes of acquired profiles across regions, profile characteristics, and time, hence providing a realistic proxy measure of actual victimization rates.
4. We provide a characterization of attackers' purchasing decisions and their price sensitivity across profile types. Whereas limited to the setting of Genesis Market, our characterization provides novel insights on criminal purchase decisions and associated trends;
5. We provide a robust estimation of the revenues of the analyzed criminal market. We analyze sale trends and derive market economic size employing a mixture of real, predicted, and simulated sale data;
6. We discuss our findings on attacker preferences and their relation to attack surface evaluation, and to the identification of possible countermeasures in response to market observations.
7. We share all the datasets and the crawling infrastructure at `https://security1.win.tue.nl/doku.php?id=artefacts#data_sharing`.

This chapter proceeds as follows: Section 6.2 discusses related work; Section 6.3 breaks down the problem at hand and presents our methodology for data collection and analysis, whereas Section 6.4 first presents an overview of the data, to then delve in sale activity in Genesis Market. Section 6.5 discusses our findings and concludes the chapter.

6.2. Background and Related Work

As discussed in Chapter 5, Genesis Market specializes in offering *user profiles* to attackers (i.e., Genesis Market's customers). We refer to Section 5.3 in the previous chapter for further details on Genesis Market's product. For context, from that section, we recall how the profile selection process works: customers of Genesis Market can browse across the portfolio of offered profiles and evaluate them by inspecting the list of websites for which stolen credentials are present, the country of origin of the profile, when the information was first harvested and last updated, etc. (refer to Chapter 5 for a full enumeration). When buying a profile, the customer can download the bundle of information within that profile together with a Google Chrome browser extension developed by the Genesis Market operators. Importantly, upon purchase of a profile, the profile is unlisted from the market. On one hand, this assures a

profile is purchased only once; on the other, it provides a method to precisely measure sales. Interestingly, recent work by Lin et al. [169] proposes techniques to evade risk-based authentication (RBA) services similar to those originally introduced by Genesis Market (including stealing information from the victim's environment, and re-producing these in the attacker's by means of a browser extension), and find that authentication services are indeed vulnerable to these attacks. The threat posed by impersonation attacks against RBA demonstrated in [169] was first described in [Campobasso7], together with a description of the `IMPaaS` threat model, Genesis's pricing model, and features of the traded product (i.e., the user profiles). Differently from these works, in this work we study attackers' profile purchasing behavior by monitoring patterns in product offering and sales from the market activity itself, to derive insights on attackers' decisions when selecting targets to impersonate under the `IMPaaS` model. Further, by analyzing actual sales data from Genesis Market, we evaluate the overall relevance of the `IMPaaS` threat worldwide.

To contextualize this work, we refer to the cyber risk model proposed by Woods and Böhme [272] (depicted below, in gray). The risk model presented in [272] identifies a num-

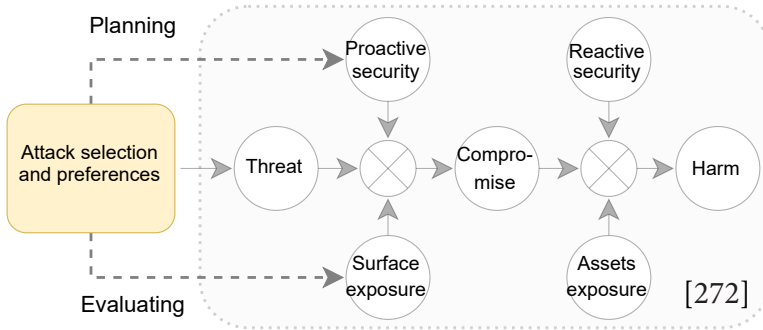


Figure 6.1: Overview of the proposed theoretical framework.

ber of latent variables whose interplay characterizes the overall risk picture, starting from 'Threat' and leading to 'Harm'. On the other hand, threats do not materialize 'out of thin air'; rather, they are generated by (human) attackers that, whether through access to the criminal ecosystem or by their (or their organization's) own means, consciously choose their targets and suitable attack technologies or methods [62, 15, Campobasso5]. Critically, being able to characterize attacker preferences before the threat materializes can help defenders in better devising their 'proactive security', and can provide insights on the actual exposure of an organization to said threats. To capture this, we propose to extend Woods and Böhme's model by including '*Attack selection and preferences*' as a precursor step to the arrival of a 'Threat'. Specifically, by studying Genesis Market sales, in this work we reconstruct the attacker preferences leading to the actualization of the `IMPaaS` threat, and we discuss implications on defenses and attack exposure.

Table 6.1: Relationship between challenges and mitigating step(s) of the methodology.

		Ch1	Ch2	Ch3	Ch4	Ch5	Ch6
Meth. step	Data collection & enrichment	×	×	×	×		
	Feature extract & orthogonalization					×	
	Data diagonalization						×
	Sales predict & listing reconstruct	×	×		×	×	

Related Work

Gathering data to study cybercriminal ventures is a longstanding problem. Often, data comes from manual collection [33], incomplete or partial crawling [217], or relatively outdated leaks of underground marketplaces [261, 33, 188, 217, 205, 36]. The objective difficulty linked with the collection and analysis of this type of data results in multiple studies looking at the same or similar (e.g., updated) data [261, 12, 36, 262]. As discussed in Chapters 3 and 4, several authors develop specialized crawlers to scrape the target infrastructure [238, 278], produce tools capable of obtaining fresh data over time across underground communities [209] and tackle the problem of developing general crawlers flexible enough to target multiple criminal forums or marketplaces [141, Campobasso3]; some of these solutions propose anti-crawler detection techniques to avoid detection from the administrators of the crawled communities [209, 217, 238, Campobasso3].

Aside from the data collection, the analysis of this type of data presents foundational challenges: the processes behind its generation are oftentimes at least partially unknown [20], and estimates (particularly of an economic nature regarding sales and purchase activity) can only be approximated [238, 262]. *Post-mortem* analyses of cybercriminal revenues based on data provided by law enforcement following takedowns of markets [192, 259] or leaked data [14, 45, 144] are often among the most accurate estimates one can derive, albeit generally on criminal marketplaces or communities that no longer exist. The difficulty of this data collection and analysis process sometimes results in contrasting and/or disputed estimates [178, 194]; [20] provides an additional commentary. Live data collection with a clear data generating process aiding its analysis is rare in criminal settings, albeit crucial to obtain reliable estimates of still-alive and evolving cybercriminal activities and to develop tailored countermeasures to operating threats [25].

6.3. Methodology

As part of our approach, we first identify critical aspects of the data collection and the problem at hand. These challenges are posed by the nature of the data and of the problem we address; therefore, it is useful to detail these challenges upfront. Table 6.1 shows which methodological steps address them.

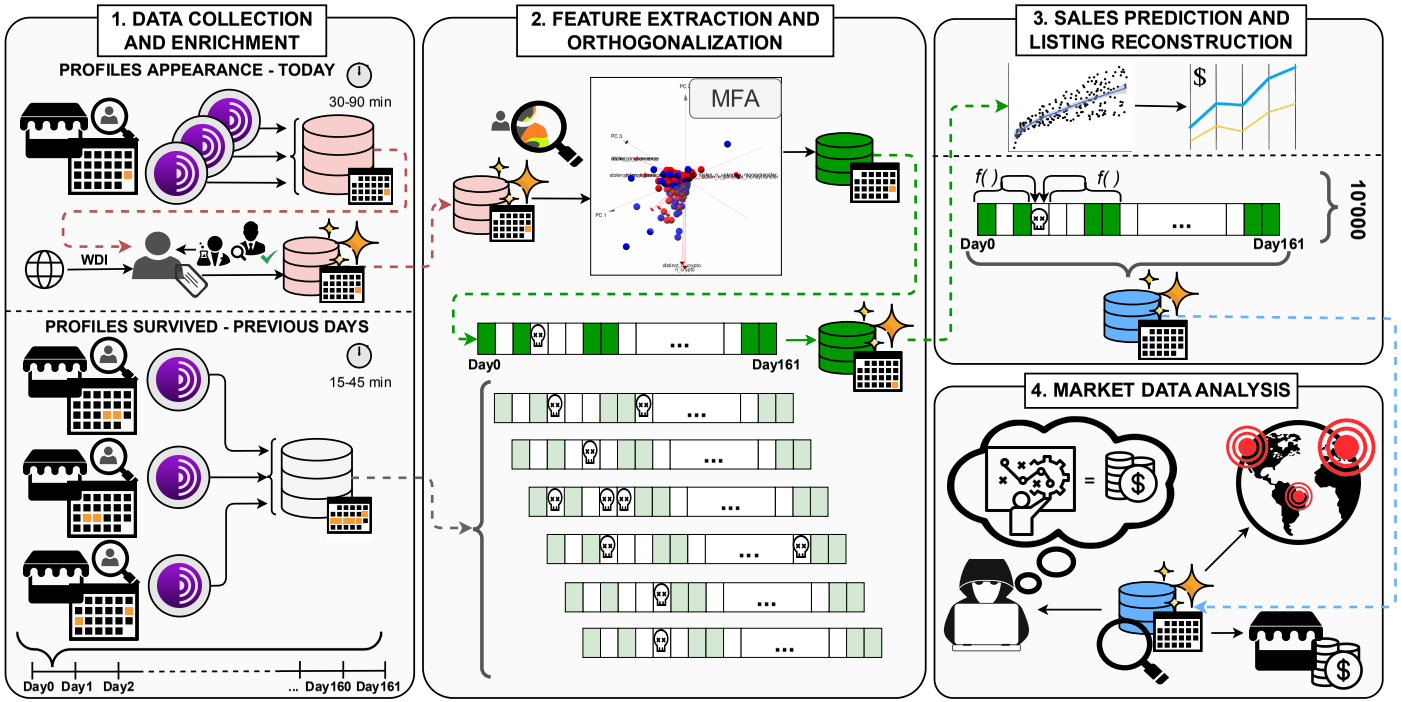


Figure 6.2: Methodology overview. Acronyms: WDI = World Development Indicator; MFA = Multiple Factor Analysis.

6.3.1. Challenges

Ch1. **Reliability of criminal infrastructures.** Connectivity to criminal infrastructures (Genesis Market included), critical for prolonged crawling activities monitoring market evolution such as the one performed for this research, is often unreliable.

Ch2. **Bandwidth of TOR network.** To minimize exposure of the crawling activity, it should be performed over TOR. It is critical for the data collection to use as little bandwidth as possible so as not to compromise other TOR users' experience.

Ch3. **Crawling prevention measures.** Prior work showed that Genesis employs anti-crawler measures that can lead to user banning; as obtaining access to Genesis can require up to ≈ 1 month, it is critical that the crawler accounts for the countermeasures in place.

Ch4. **Repeated measurements.** To monitor product evolution on Genesis Market we must monitor their (dis)appearance as time progresses. This requires repeated (re-)measurements of the platform at different moments in time and within sufficiently small time windows. However, these time windows cannot be too small due to the risks connected to **Ch3**, meaning that a trade-off exists between sampling completeness and persistence of market access.

Ch5. **Measurement of aggregate, high-dimensionality effects.** Uncovering the decision process of attackers operating on Genesis Market to acquire a profile requires transparently linking highly-dimensional data [Campobasso7] with sale observations while preserving as much as possible (or all) of the original variance in the observations. Further, only being able to observe the aggregate effect of customers' purchase decisions increases uncertainty in the model.

Ch6. **Accurately measuring sales.** We must distinguish a profile 'disappearance' during a crawling session caused by its 'sale' rather than by temporary glitches or effects.

6.3.2. Methodology steps

We devise a multi-stage methodology for collecting and enriching Genesis Market data, with the aim of modelling market sales and gaining quantitative insights into the market economy and customer purchase decisions. Figure 6.2 gives a bird's eye of our methodology.

Methodology overview. Our overall methodology is divided in four steps. In the first step, we devise a method to scrape the market daily supply at regular intervals and to monitor it for a week. By doing so, we are able to monitor both daily supply (with their full profiles' characteristics, pink database, Figure 6.2) and their presence in the market for the following week (grey database). This procedure serves to collect the data necessary to evaluate characteristics in the supply and to monitor sales, which in turn will inform us on the attackers' preferences. We then proceed to enrich profiles with additional information to obtain further insights ('shiny' pink database). In the second step, we transform the enriched profiles to reduce noise and data dimensionality to model sales (green database). Without this, it would be virtually impossible to obtain a sales prediction model that accounts for the multifaceted

profile selection process of the attackers. We then compare the available data about supply and 'survived' profiles over the following days (grey database) to decide what data should be included in the analysis, as the consequence of the incomplete data scraping ('shiny' green database). This procedure will inform us on the optimal trade-off between data completeness and sales model convergence. In the third step, we compute a sales prediction model and simulate the missing data to rebuild the entire dataset. Without this, we would only gather partial insights on the market. In the fourth and last step, we analyze the obtained data to infer attackers' preferences, characterize market supply, estimate revenues, and inform and extend a cyber risk model.

Data collection and enrichment

To conduct our study, we exploit the fact that Genesis only lists *still available* profiles on their listing and removes items only through a sale, or a reservation (as verified by us, the reservation mechanism allows a customer to reserve a profile for 30 minutes, temporarily removing the product from the listing). We exploit this mechanism to collect data on profile appearances and their persistence on the market. From the first data collected, we notice that the chances of sale for a profile sharply decrease after the first day and become negligible after the sixth day. Therefore, we establish six days as the time window of choice during which to monitor a profile after its appearance on Genesis Market. To scrape the market while addressing **Ch3**, we create six crawler instances: three *appearance* and three *persistence* crawlers. This choice was made after considerable trial and error (leading to account banning on the market) to strike a balance between the volume of data to collect, the level of stealth needed to remain 'under the radar', and the number of available accounts we could 'burn through' during the data collection. This last requirement is particularly critical as accounts were not easy to obtain across other platforms and communities. The *appearance* crawlers reach the market's listing section at midnight (Moscow time), tasked with obtaining a full description of appeared profiles in the previous 24 hours relative to the start of the crawling on day d. We decided to crawl during the eastern Europe night to reduce our impact on the platform's responsiveness during (likely) active hours, and therefore reduce possible alerts triggering further investigation from the market admins. In mitigation to **Ch2,3**, and to keep the architecture as simple as possible, the three crawlers split the workload independently by collecting the full list of appeared profiles and selecting only the $1/3$ that corresponds to their crawler id. Within their $1/3$, each crawler randomly selects 25% of the listed products. Initially, we attempted to download the whole offer of the day, but crawling sessions often exceeded six hours, which would in turn introduce large inconsistencies in the temporal dimensions of the measured sales. This procedure allows us to limit data crawling visibility (**Ch3**) while collecting a representative and valid sample of data on profile appearance and characteristics. For each appeared user profile we collect the full set of features that characterize it. The result is a data collection that fully represents what the market customer sees when viewing an item. Additionally, one *appearance* crawler is tasked with collecting a recap, offered by Genesis Market, of the number of appeared profiles during the last 24 hours. For each day d in the *observation* period up to day D, we aim at obtaining a data collection $L_0^d,\ \forall d \in [0..D]$ of all appeared user profiles on that day. In parallel, the three *persistence* crawlers monitor the market to collect the names of the profiles still available. The *persistence* crawlers monitor appeared profiles in each day d for six consecutive

days since d. Each *persistence* crawler is assigned a period of two days relative to the (midnight of) the day in which the crawler is run. The *persistence* crawlers collectively generate, for each day d, a dataset $L^d_{1..6}$ containing the IDs of the appeared user profiles on day d (and not yet sold) across each *monitoring* day $n \in [1..6]$ relative to d. To limit the impact of **Ch4** we probe the market for changes in product offering every 24 hrs.

Each run of the three appearance crawlers requires $30-90$ minutes depending on the products offered, market responsiveness (**Ch1**), and available bandwidth (**Ch2**). The three persistence crawlers take $15-45$ minutes on average. We consider this sufficiently fast to mitigate **Ch4**, while not aggravating **Ch1–3**. We further mitigate **Ch3** by throttling traffic.

We implement the crawlers using instrumented TOR Browser[251] instances via the Selenium[130]-based library *tbselenium*[5] to generate traffic from an instrumented browser without having to tinker with technical details that may raise a red-flag in crawler detection systems [Campobasso3]. Each crawler instance accesses a completely different TOR circuit to avoid using the same bastion host. Further, each of the crawler instances is assigned to a different user account under our control, limiting the activity of each account overall (**Ch3**). Finally, to assure an as-complete-as-possible data collection in the presence of **Ch1,2**, the crawlers are designed to automatically adjust timeouts to refresh pages when those cannot be fetched on the first attempt, by doubling the default fetch timeout of 15 seconds until the page is not successfully loaded, or retrying every 5 minutes if the market is not reachable. The crawler keeps attempting to connect to the market until 2am Moscow Time. This choice is to limit noise in the data collection whereby profiles disappear before they are collected by our crawler (ref. **Ch4**); this assures that comparisons across snapshots on different days remain meaningful.

We enrich obtained user profiles with data on the 2020 per capita GDP of the respective country of origin. To better reason about the characteristics of a profile we follow [Campobasso7], and aggregate and classify available resources (stolen credentials originating from a specific website) for that profile in six categories: `Services` (delivery of physical or digital goods, such as Netflix or Gmail); `MoneyTransfer` (traditional payment, like PayPal or American Express); `Crypto` (cryptocurrency exchanges, such as Crypto or Bitpanda); `Social` (user-generated content, like Facebook or Twitter); `Commerce` (purchase or book goods from one or multiple vendors, like Amazon); `Other` (for otherwise non-classified resources). The classification is done manually by an author and independently checked for a random sample of 100 resources in a blinded process by a second author until conflicts are resolved.

Feature extraction and orthogonalization

Due to the high uncertainty inherently involved in reconstructing purchase decisions, let alone criminal ones, we employ a set of techniques to maximize the amount of information available to our modelling. The objective is to transform the data to prevent correlated, high-variance variables [Campobasso7] to dominate the resulting analysis, while not losing information in the transformation. That is also challenging because profile characteristics are naturally 'nested' within groups of semantically related information on that profile [Campobasso7]: for example, both the available cookies and the available browser environments describe features of available browsers; as such, these features should *not* be treated

as independent entities. To accommodate for this we employ *Multiple Factor Analysis* (MFA) as the method of choice to derive linearly uncorrelated dimensions of the data for our analysis [2]. MFA integrates Principal Component Analysis (PCA) for the numerical variables and Multiple Correspondence Analysis (MCA) for the categorical variables while preserving effects at the group level. As a result, the original variables collected in our dataset are projected over several, orthogonal 'dimensions' with near-zero correlation, thus maximizing the explanatory power of each added dimension by removing overlap, helping in the identification of patterns in data and mitigating **Ch5**.

Data diagonalization and time window selection. The data diagonalization has the primary function of allowing us to make a well-informed decision on how wide the time window we consider, across the six day monitoring period, should be. A discussion of why this is necessary for modelling consistency and preserving the internal validity of this study is provided in Section 6.3.2. To inform this decision, we (a) estimate how many profiles are sold for each of the six monitoring days, and (b) evaluate whether sold profiles remain similar regardless of the day on which they are sold. To achieve (a), for each day, we mark a profile as sold if the product disappears after n days and does not appear on any subsequent day. To do this we only keep records of days that we have fully monitored up to a certain monitoring day n (i.e, $\bigcup_{d\in D} L^d_{0\ldots n} = \bigcup_{d\in D} L^d_0 \cap L^d_1 \cap \ldots \cap L^d_n$, with $n \in [1..6]$). For example, if we collect $L^{d'}_0$ and $L^{d'}_1$ but not $L^{d'}_2$, we will keep d' in $\bigcup_{d\in D} L^d_{0..1}$, but not in $\bigcup_{d\in D} L^d_{0..2}$ (i.e., day d' will result as a missing day in the diagonalized data for $n = 2$). We then achieve (b) by simply comparing profile characteristics (orthogonalized via MFA) across profiles sold on different days. To distinguish 'sold' from 'reserved' profiles (**Ch6**), we check for every collection L^d_0 if a profile disappeared in any of $L^d_{1..n}, n \in [1..6]$ reappears in any of $L^d_{1..n'}$, with $n' > n$ and label them accordingly.[1]

Sales prediction and listing reconstruction

In this step, we use the resulting dataset to derive a sales prediction model as a function of the profile's features and employ it to simulate data for which we have no observations.

Modelling profile sales. To build our sales model, an important consideration is that attacker decisions to purchase a profile may be affected by what alternatives are available for selection at the moment the decision is taken [19]. As we cannot fully reconstruct this (ref. **Ch4,5**), we model it at the level of the observation day d as a random effect (see [8, Ch.13, pp. 489, for a formal definition, and 13.2.3 pp. 495 for a discussion on coeff. interpretation for cluster-specific models]) that captures the (time-dependent) stochasticity introduced, on the customers' decision, by the alternative options available in that (those) day(s). We note that each monitoring day accounted for a sale requires considering the (random) effects caused by the availability of not-yet-sold profiles for all the previous days, increasing the overall uncertainty to model. This creates a trade-off, as it implies that for every additional monitoring day $n \in [1..6]$ included in the sample we necessarily remove observation days (i.e., those without complete monitoring up to day n), and therefore profiles, from $\bigcup_{\forall d\in D} L^d_{0..n}$ (as the chance that at least one data collection failed increases with n, due to **Ch1**). We therefore

[1] This leaves unchecked profiles reserved on monitoring day 6. In practice, this does not affect our data analysis and results, see Section 6.4.1.

prioritize keeping modeling complexity at a minimum while retaining the highest number of data points for our model.[2]

Data reconstruction and simulation. For each day d for which we have a data collection L_0^d but no subsequent observation in $L_{1..6}^d$, we use the estimated model to predict which profiles appeared in that day were likely to be sold. For every missing L_0^d, we (a) first estimate the number of products we should have collected for that day, and (b) run a simulation batch reconstructing which profiles could have appeared on that day. To have an estimate for (a), we consider the first available $L_{1..6}^d$ to make a lower bound figure of how many profiles appeared in L_0^d, and derive our estimation by scaling it up by the average rate of sale at that monitoring day; if this information is not available as well we use the overall market recap (provided daily by Genesis Market) with the number of appeared profiles for that day d. However, we find that this information is not always accurate as it reports fewer profiles than what we measure in $\approx 30\%$ of cases. Thus, we correct this figure by computing the average ratio between the measured offer and the numbers reported in the market recap. To perform (b) we build a set of simulations by sampling, with replacement, the number determined by (a) of profiles from the surrounding days.[3] On the simulated data we then apply the estimated model to predict sales and calculate central estimates and confidence intervals of market statistics and sale trends from the resulting data distributions. Due to computational constraints, we build two batches of simulations: one ($n = 100$) retaining detailed data on sampled profiles (e.g., geographic location, available resources, ..), is used in Section 6.4.2 and 6.4.2 to report on detailed profile descriptors. For the second batch of simulations ($n = 10'000$) we only retain chosen statistics from each simulation and use it to estimate the overall market value in Section 6.4.2. The high number of simulations here is chosen to provide as accurate an overall figure as possible of the sales data. In either case, simulated days are always clearly marked in the reported figures.

6.3.3. Ethical considerations

The details of available profiles advertised on Genesis Market before purchase do not contain any PII. Advertised profiles include a censored IP address (e.g., 14.25.xxx.xxx), country of origin, affected OS, and a list of the websites for which stolen credentials exist, with no details on said credentials. Similarly, available cookies are reported as a count per browser, alongside a list of the affected browsers. Because this study relies solely on information available in profile listings on Genesis, the collected data does not contain any PII. An ethical revision of this research was performed by the relevant board at our institution and approved under reference no. ERB2021MCS1.

[2]Due to the inherent uncertainty of the purchase decision process, we prioritize minimizing the True Negative Rate (TNR) of our estimator, and consider two different threshold values for sale prediction corresponding to $TNR = 95\%$ for the 'conservative' estimator, and $TNR = 80\%$ for the 'generous' one.

[3]This decision was taken after checking that profiles appearing on subsequent days have similar characteristics. We report results in the Appendix.

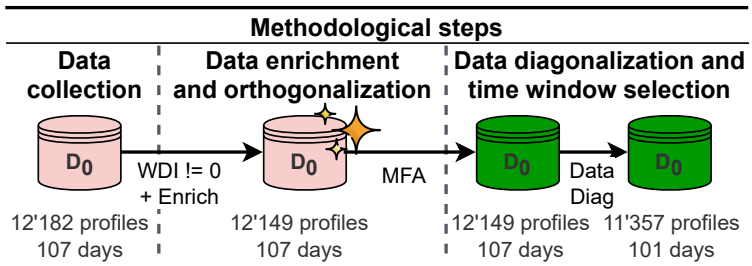


Figure 6.3: Data preprocessing pipeline.

6.4. Results

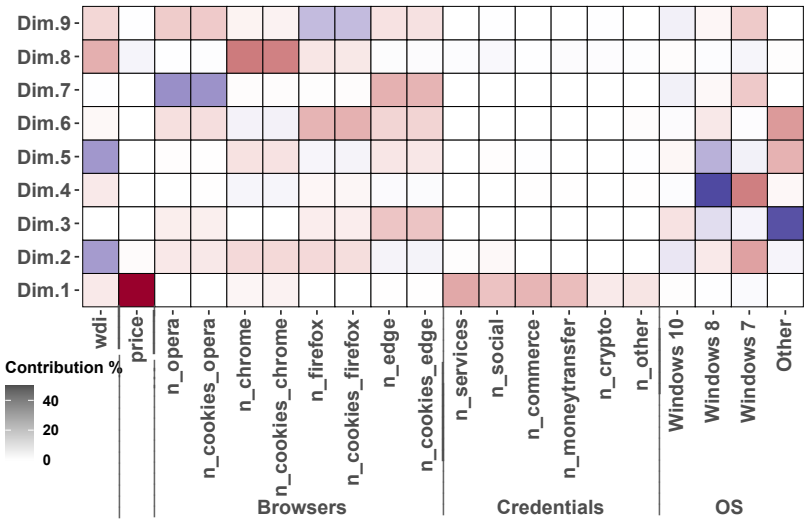
We first describe the data preprocessing and the resulting overview of the market data; the section then continues by analysing attacker profile acquisition trends, estimating sale volumes, and market size.

6.4.1. Data preprocessing

Data collection and enrichment

Figure 6.3 provides an overview of the data preprocessing pipeline. The data collection spans from Jan 21st 2021 to Jun 30th 2021[4] and counts a total of 107 complete L_0^d over an observation period of a 161 days, corresponding to a total of 12′182 profiles. From the country of origin of each user profile, we derive the 2020 per capita GDP (Worldbank NY.GDP.PCAP.CD) indicated as WDI[5]. We found 33 profiles originating from Reunion, Mayotte, French Guiana, Guadeloupe, and Taiwan, for which no information is available; we discarded them from the analysis, reducing the number of profiles to 12′149. Further, for each profile, we count the number of compromised browsers by family (e.g., Firefox, Opera), the available cookies by browser family, the available resources and related webplatforms, divided into six categories: `Services`, `Commerce`, `MoneyTransfer`, `Other`, `Social` and `Crypto`. Categories represent the purpose of the platform examined (e.g., `MoneyTransfer` contains websites of financial institutions enabling money transactions, `Commerce` includes e-commerce platforms, ...). For consistency and benchmark, we adopted the same categorization scheme reported from [Campobasso7]. We identified a total of 1′839 distinct platforms. 576 identifiers represent the same website or respective Android app (e.g., WellsFargo and `android://com.wf.wellsfargomobile/`); to avoid data duplication, we collapse those under the same identifier, reducing the number of distinct platforms to 1′297. We assign each platform to its corresponding category. This yields 475 platforms of type `Services`, 357 `Commerce`, 265 `MoneyTransfer`, 127 `Other`, 39 `Social`, and 34 of type `Crypto`. For each profile, we derive the number of resources in

[4]Crawling started in Nov 2020, but as Genesis Market went offline for an infrastructural upgrade from 11 Dec 2020 to 15 Jan 2021 we discard data from the previous period for consistency. Crawling resumed on the 21st Jan.
[5]If data from 2020 is not available for a country, we use the most recent estimation present in the same database.



Blue and red respectively indicate positive and negative contributions of each variable to a dimension. Color intensity is proportional to the magnitude of the contribution.

Figure 6.4: Variables' contributions top 9 MFA dimensions.

6

each category. Following the validation process outlined in Section 6.3.2, the final classification agreement was 97%. Table 6.2 reports the dimensions of the resulting data.

Feature extraction and orthogonalization

The MFA analysis comprises overall 18 variables (ref. Table 6.2). Variables are assigned to the groups `Price`, `Browsers`, `OS`, `WDI`, `Credentials`; `Sold` is considered only as a contrast variable and is not included in the MFA (as it represents an outcome and not a feature of the profile). We log-transform and scale every numeric variable to unit variance, to ensure each variable equally contributes to the definition of the factor space. As for our application, the main purpose of the MFA is to get rid of multicollinearity issues across variables (as opposed to dimensionality reduction), so we do not constrain the number of dimensions in output of the MFA. We employ the `FactoMineR` package's [161] `MFA` implementation in the statistical software package R. We run the MFA analysis on all the 12′149 enriched profiles. We obtain 20 orthogonal dimensions; for brevity, we report here the first 9, representing 89.25% of the overall variance (a full breakdown is available in the Appendix). Figure 6.4 offers a full breakdown of the contributions from each variable for the resulting top 9 dimensions. Each dimension is calculated as a linear combination of all variables; the coefficients assigned to each variable within a dimension (i.e., their 'loadings') are correlated to each variable's contribution to that dimension. The sign of that coefficient indicates whether the variable and the dimension are positively (red) or negatively (blue) correlated. Figure A6.1 (reported in the Appendix, together with an extended description of MFA interpretation) provides insight into the construction of the MFA dimensions and the contributions of each variable within their groups. To illustrate, we discuss the top 3 addressing the

most variance in the data. A closer look at Figure 6.4 shows that the three variables within the `Creds` group `n_moneytransfer`, `n_services` and `n_commerce` contribute the most, together with `Price`, to `Dim.1`. Therefore, `Dim.1` can be interpreted as representing high-resource, high-cost profiles within Genesis Market. That is to say, profiles similar in composition to the feature values captured by `Dim.1` (e.g., a high price) will score high on this dimension. Similarly, `Dim.2` is mostly influenced by profiles characterized by variables in the groups `Browsers`, `OS`, and `WDI`. `Dim.2` captures profiles from relatively poor countries according to the WDI index but rich in cookies. Interestingly, `Dim.2` also reveals that those profiles are more likely to exhibit older operating systems (Windows 8 and 7) and to feature browsers different from Edge. Profiles characterized by Edge running on Windows 10 instead seem largely captured by `Dim.3`. Similar considerations on the profiles' characteristics can be made by comparing the interaction patterns visible in Figure A6.1 across all dimensions and variables (groups).

Data diagonalization and time window selection. The diagonalization process offers insight into the available data that can be used to model customer purchases. We evaluate the fraction of data that remains available to our modelling when varying the size of the measurement window for $L^d_{0..n}$, $n \in [1..6]$. Results are reported in Table A6.1 in the Appendix, together with additional details on its construction.

Among sold profiles, more than half (58%) is sold within the first day. By contrast, the fraction of overall sales that can be accounted for by including subsequent days does not surpass 78% of sales overall (including up to L^d_3), but at the price of removing 19 observation days (as opposed to 6 with L^d_1) from the sample and $\approx 2'000$ profiles. These missing observations not only remove data for the model training but also create 'holes' in the data collection that will have to be 'filled back in' via model prediction, bringing in additional uncertainty. To identify whether profiles of specific types are more likely to be sold after a certain number of days since their listing on the market, we look (not reported here for brevity) at the features of sold and unsold profiles. A set of Wilcoxon Sign-ranked tests finds no overlap across observations, suggesting that looking at profiles sold on a given day is representative of looking at those sold on surrounding days.

For these reasons, we consider only looking at the first day of sales as an acceptable trade-off. This results in the final dataset comprising $11'357$ profiles (of the $12'128$ originally fetched), sampled across 101 (out of 107) observation days while capturing 58% of the overall sales.[6] This gives us a total of six L^d_0 days with missing L^d_1 and $161 - 107 = 54$ missing L^d_0 days to simulate, for which we predict sale outcomes as detailed in Section 6.3.2.

Overview of Genesis Market profiles

Table 6.2 provides descriptive statistics of the final dataset. Profiles are offered at an average price of 21.32 USD; the 5% most expensive profiles are priced at 59 USD or more. When looking at sold profiles, the average price reaches 25.96 USD, with the 5% most expensive exceeding 101 USD. Chrome appears to be the most popular browser among the affected

[6] This also excludes the data censoring our data diagonalization suffers from for profiles 'reserved' on the 6th monitoring day, discussed in Section 6.3.2.

Table 6.2: Descriptive stats for $L^d_{0,1}$ and related MFA groups.

	Grp	Variable	Min	Mean	Max	SD
Original Variables	Price	Price (USD)	1	21.32	350	24.91
	Browsers	# Opera	0	0.20	1	0.40
		# cookies	0	122.86	7332	501.58
		# Chrome	0	0.76	1	0.43
		# cookies	0	1165.81	9448	1215.45
		# Firefox	0	0.26	1	0.44
		# cookies	0	185.60	5911	601.31
		# Edge	0	0.10	1	0.30
		# cookies	0	43.22	4098	253.67
	OS	OS	–	–	–	–
	– ‡	Date infection	21-01-21	15-04-21	30-06-21	51.44
		Date update	21-01-21	15-04-21	30-06-21	51.45
		Country	–	–	–	–
Data Enrichment	WDI	WDI	126.90	26999.64	86601.56	18801.68
	Credentials	# Services	0	10.78	569	16.56
		# Social	0	4.09	263	7.36
		# Commerce	0	3.17	149	7.11
		# MoneyTransfer	0	1.38	248	4.96
		# Crypto	0	0.18	53	1.21
		# Other	0	0.25	38	1.08
	Sold†	Sold	–	–	–	–

† Supplementary variable of the MFA; ‡ Not part of the MFA.

victims, being on average 3 times more frequent than Firefox and Opera; Safari and Internet Explorer never appeared during the analyzed period. On average, profiles contain predominantly `Services` credentials, followed by `Social` and `Commerce`. Since the data collection happens at most 24 hours after a profile has been published, the date of the last update for each profile often matches the infection date; the former tells a customer whether a profile contains fresh credentials, and it is relevant when looking at older profiles.

By looking at the reported standard deviations, there are large variations in the number of stolen cookies across all browsers. This difference suggests that the target population shows diverse traits in terms of Internet usage: the 90% of the victims found in $L^d_{0,1}$ appear to use few services and few platforms only, counting 48 credentials or less, while the remaining 10% has 87 credentials on average and 883 at maximum. Similar considerations on the number of stolen credentials may shed light on some population characteristics. While a small number of credentials per profile may indicate a limited Internet activity of the victim, when paired with profiles presenting a large number of cookies it may indicate users not saving their passwords in the browser, or using a password manager.[7] Looking at the geographical distribution of profiles, the overwhelming majority of profiles originates from

[7]Widespread infostealer malware like AZORult and RedLine are incapable of stealing passwords from password managers, although an attacker could identify the master password by sniffing keystrokes[145, 23].

Europe (62.14%), followed by North America (11.97%), South America (11.83%), and Asia (11.01%)[8]; Africa and Oceania together account for the 3.04% of total profiles. Profile composition varies across regions; North American profiles are generally richer in credentials. These profiles offer, on average, 27 credentials, while Europe, South America, and Asia offer respectively 20, 19 and 13. The same trend is noticeable also with `Commerce` and, albeit less remarkably, with `Crypto` credentials. That is well reflected in the price of these profiles (respectively, on average 34.22, 20.29, 19.91, and 14.03 USD), following the intuition that wealthier countries have a more appealing resource composition for the market's customers, confirming the findings of [Campobasso7].

6.4.2. Attacker activity on Genesis Market

In this section, we provide an analysis of attackers' purchasing decisions and associated factors. Unless otherwise stated, reported significance statistics are produced via Wilcoxon Rank-Sum tests; we consider an α value of 5% as the threshold for statistical significance.

Analysis of attacker preferences

To evaluate customer preferences when selecting profiles to buy among those offered on Genesis, we define a set of nested generalized linear mixed models (GLMM) to estimate the relation between the obtained profile dimensions and a purchase decision. We build the final model including dimensions in output of the MFA in incremental steps, ordered by their relative contribution in explaining our dependent variable (i.e., sales; details on this process in the Appendix). The final model obtains an R^2 of 27.8%. The model construction assures that virtually all the information available in the market data is captured.[9] The model obtains a satisfactory AUC of 0.77, despite the high uncertainty inherent to the effect it models. For this discussion, the table below reports the dimensions that explain at least 1% of the total variance in the model (% of explained variance by each dimension reported below the coefficients).[10] Full details on the model are provided in the Appendix.

c	`Dim.8`	`Dim.2`	`Dim.13`	`Dim.9`	`Dim.4`	`Dim.6`	`Dim.5`
-2.51^{***}	0.62^{***}	-0.41^{***}	1.02^{***}	0.32^{***}	0.19^{***}	0.38^{***}	-0.17^{***}
–	(8.2%)	(5.7%)	(3.0%)	(2.7%)	(1.7%)	(1.6%)	(1.5%)

#obs = $11'357$, $R^2_m = 0.264$, $R^2_c = 0.278$, $std(c|day) = 0.25$, $^{***}p < 0.001$

[8] As it is often the case with Russian-based cybercriminal ventures, profiles in our collection do not concern countries in the Russian area of influence (Commonwealth of Independent States), except for Kyrgyzstan and Uzbekistan. A deeper look into the market shows that among the ten CIS countries, only Belarus, Kazakhstan, and Russia do not appear at all.

[9] To further enrich the model, one would have to look beyond the market data itself and, for example, interview the customers of the market when they make a purchase decision. That is an inherent limitation common to all studies of this type (see, for example, discussion in [238, 14]); the decision to provide conservative sale estimates derives from this observation, to avoid overshooting in the presence of structural uncertainty in the data.

[10] Because of the high uncertainty in the data, we employ all dimensions that an ANOVA test (not reported because of space constraints) evaluates as significant to be included in our sales prediction model. This is to maximize the model's power in predicting sales in our simulations for the missing data. However, for the purpose of interpreting results, we note that only dimensions that capture a large enough variance in the outcome are worth considering to add meaningful insights into attacker preferences.

Coefficients can be interpreted on the same scale; coefficients should be interpreted jointly with the dimension compositions reported in Figure 6.4. Positive (negative) regression coefficients mean that user profiles that score high on that dimension have a greater (lower) chance of being sold. The sign of the variable loadings for a given dimension (color-coded in Figure 6.4) indicates whether a variable is associated with a 'high score' on that dimension depending on its value in the original distribution (i.e., above or below the mean). For example, the positive coefficient of `Dim.8`, together with its variable compositions reported in Figure A6.1, suggests that profiles with high `WDI` with a large number of cookies originating from Chrome, and Firefox to a lesser extent, are preferred by the attackers. By contrast, the negative coefficient for `Dim.2` suggests that profiles from less wealthy countries featuring older operating systems (Win 7,8) are less likely to be sold even if they might be high in resources/cookies. `Dim.13` (dimension composition observable in Figure A6.2 in the Appendix) suggests that attackers are interested in profiles rich in `Social` and `Moneytransfer` credentials, as long as they are cheap; `Dim.9`, `Dim.4` and `Dim.5` further corroborate that attackers prefer profiles originating from wealthier countries and characterize different profile configurations; `Dim.9` and `Dim.4` identify a group of profiles originating from systems running Win 7 but with a different resource composition from that of `Dim.2`. In particular, `Dim.9` suggests that Win 7 profiles are more likely to be sold if they feature data for Opera or Edge. Interestingly `Dim.5` and `Dim.6` indicate that the presence of a specified OS loses importance ('OS=Other') provided that the profile is associated with a high WDI and has several resources for Chrome or Firefox.

Overall we find a positive association between the composition of profile characteristics and its likelihood of sale, with `WDI`, `Price`, and technical features such as the browser playing a predominant role in the purchase decision. Perhaps surprisingly, the type and number of included resources (e.g., `Social`, `Moneytransfer`, ..) seems to play only a limited role in the final decision. This may indicate that the average attacker does not necessarily prefer profiles rich in resources, as the victim's intrinsic value (i.e., their wealth, which may become available after a successful impersonation attack) is the same regardless of the type or number of associated resources (as long as the attacker has access to some of them). Indeed, the model coefficients indicate (across all dimensions, save for `Dim.13`) a strong attacker preference for profiles from high-WDI countries, suggesting that the perceived value of the victim's profiles is more relevant to the attacker than the number of ways in which that value can be accessed.

Trends in market supply and attacker demand

Figure 6.5 provides a bird's eye view of the volume and average prices for available and sold profiles, globally. The median price for offered profiles in Europe is 15 USD, while in North America reaches 18 USD, suggesting that profile composition is richer in the latter. The most expensive profiles originate from Oceania, with a median price of 19.5 USD, although they are a minority (0.93%). As per Section 6.4.2, and confirming results in [Campobasso7], profiles originating from wealthier countries show higher prices on average due to their per capita GDP (the only two countries attacked, Australia and New Zealand, have respectively 9th and the 21th highest per capita GDP in 2021). When looking at sold profiles, the demand sharply rises for North America, accounting for 34.54% of total sales, while Europe 'only' for

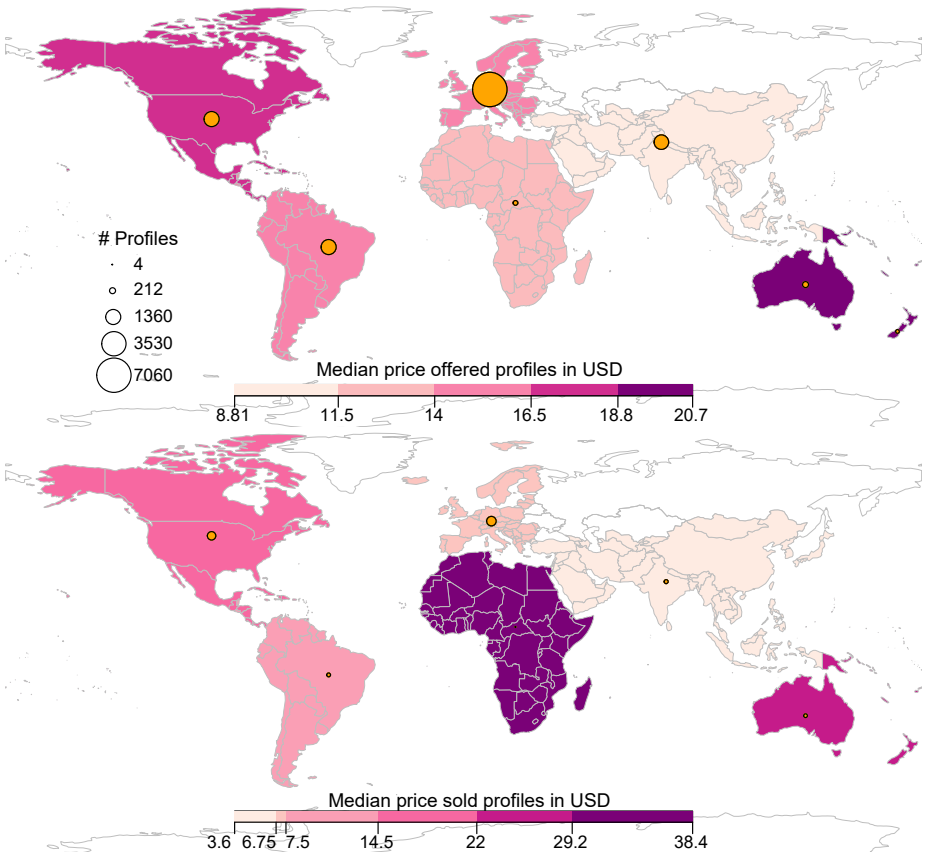


Figure 6.5: Overview of offered profiles (top) and acquired profiles (bottom) globally. Region colors represent median prices; superimposed dots represent volume of (either available or sold) profiles in the region. Comparing statistics from left to right provides an overview of profile offer and demand across regions.

43.67%, suggesting that attackers' relative demand for North America's profiles is four times higher than that for Europe. This difference is well reflected in the median prices of sold profiles across the two regions; if the gap in median prices between North America and Europe is $18 - 15 = 3$ USD, when looking at sales this significantly widens to $21 - 7 = 14$ USD. That suggests a clear preference for attackers in North American profiles over European ones, even if supply in the latter is almost six times larger than for the former. This seems in line with the discussion provided in Chapter 5, where less stringent authentication mechanisms implemented in the North American banking system could offer increased fraud opportunities to miscreants. Africa shows the highest profile median price (35.5 USD) but accounts only for 4 sales in our sample.

Figure 6.6 provides an overview of rates of offered profiles (yellow line) and sold profiles (blue line) across regions. Market supply is not constant overall and shows highs in late February, mid-March, and between May and June. From mid-March to mid-April, North

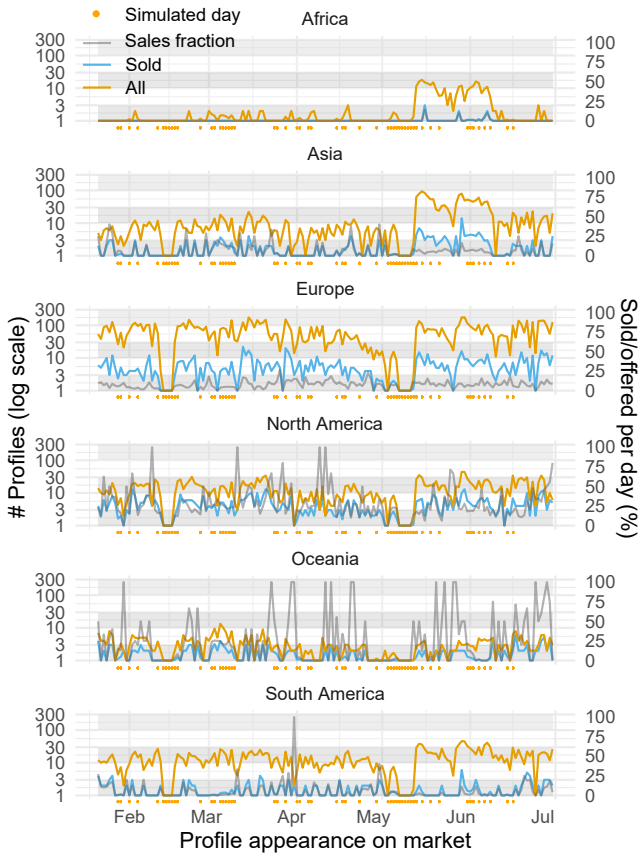


Figure 6.6: Timeline of (average) daily profile offering and sales by region.

America's offer is scarce, with an almost matching demand. Interestingly, although Europe shows the same decrease in supply, demand remains stable, moving the fraction of sold profiles roughly from 8% to 25%, suggesting that attackers could have bought some European profiles to make up for the shortage in North American ones. The same phenomenon is evident in Oceania, where the limited demand is often saturated in several days. In the second part of April, the trend partially reverts, with Europe's supply in strong decline (top early April $\approx$ 100 profiles/day, bottom late April $\approx$ 15 profiles/day, $p < 0.0001$) and North America still declining but at a slower pace (top early March $\approx$ 15 profiles/day, bottom late March $\approx$ 10 profiles/day, $p < 0.0001$). Overall, we observe a clear correlation between supply and demand for Europe (Pearson $cor = 0.74$, $p < 0.0001$) and North America (Pearson $cor = 0.76$, $p < 0.0001$); looking at the gap between the offered and sold curves, for North America we observe a higher fraction of sold profiles when compared to Europe and, in general, other regions. Among the latter, despite comparatively low volumes of provided profiles, Oceania appears to attract attackers. Looking at the fraction of sold profiles,

it appears that Asia gathers similar interest compared to Europe, despite being roughly underrepresented by a factor of 10 from the beginning of the observation period to the end of April. Asia shows a significant increase in supply after the market shut down in early May ($p < 0.0001$). Albeit South America provides similar amounts of profiles compared to North America, demand appears to be relatively low, with a few exceptions for some profiles in periods of particular shortage of profiles from other regions, such as mid-March to late April. Finally, Africa appears to be the most underrepresented region, with relative spikes in supply around mid-May and early June, leading to some of the only measured purchases we found in our observation period. No sale observation for Africa has been measured from February to the end of May 2021, although this may be partially explained as a byproduct of the adopted sampling mechanism whereby rare resources are likely not to be selected.

Attackers price sensitivity across profile types

We now investigate how price-sensitive buyers are when choosing profiles with certain characteristics. Figure 6.7, reports average prices for provided and sold profiles across regions. Supply and demand in North America exhibit modest prices in the offer compared to other regions from mid-May to the end of the observation period, while sales reach spikes of average price per sold profile as high as 300 USD in mid-May. The average sold profile in North America is oftentimes as expensive as the top 5% of provided profiles (blue dots in the region or above the dashed line indicating 95%CI in the figure). By comparing trends in sales reported in Figure 6.6 and prices in Figure 6.7, it emerges that the North American profiles from late March to late April result in attackers purchasing almost the entirety of the daily supply. Throughout the observation period, the average price for North American sold profiles is higher than the corresponding offer, despite the baseline price being higher than in other regions ($sold = 42.59$, $all = 34.77$). From February to April, European profiles gained some traction in sales, increasing the average sale price from $\approx$ 10 USD to $\approx$ 30 USD. By contrast, European profiles sold between April and early May are less on average ($m = 3.52$, $sd = 2.36$) than the previous period ($m = 6.46$, $sd = 5.08$, $p = 0.004$), while supply prices remain rather stable (April to early May $m = 19.20$, $sd = 2.27$, late February to March $m = 20.27$, $sd = 1.76$, $p < 0.001$). After a general decline up until May in Europe and North America, profiles originating in Asia soared in volumes both in terms of the offer and demand up until early June, together with a renewed interest in North American profiles until the end of May. After this period, sales volume in Asia started declining again, and North America and Europe became again the leading sources of attractive profiles. Interestingly, profiles originating from South America present similar prices to Europe and offer the same volumes as North America, but they rarely seem to interest attackers, resulting in generally low sale prices. That may reflect a general perception that profiles originating from that region are of low interest to attackers, who are willing to spend comparatively less to acquire those identities. When looking at Oceania, average prices for supply and demand move erratically, possibly due to the scarcity in the former; we witness a few notable sales reaching 100 USD on average per day during late May and June. Finally, Africa shows significantly lower prices in the supply until the early May shrink. From mid-May, prices grow to Europe levels for roughly a month, but sales do not gain traction.

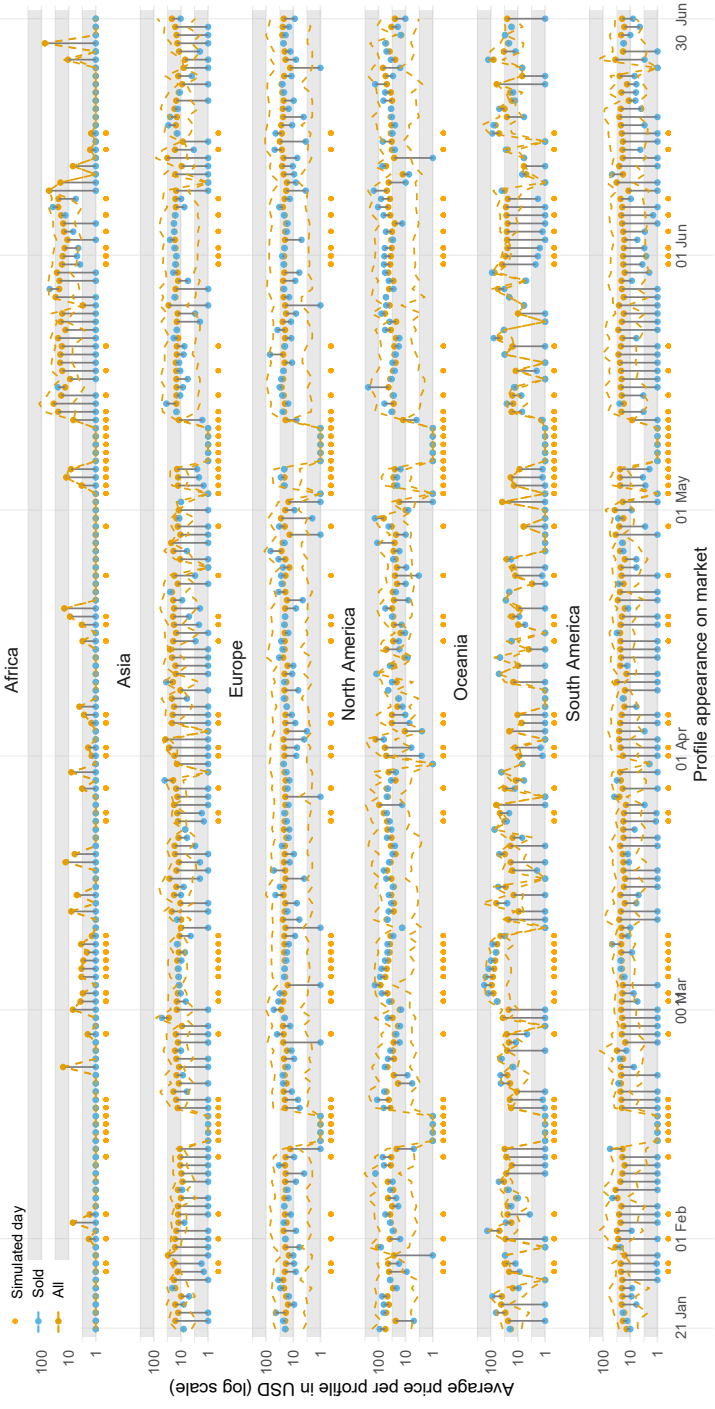
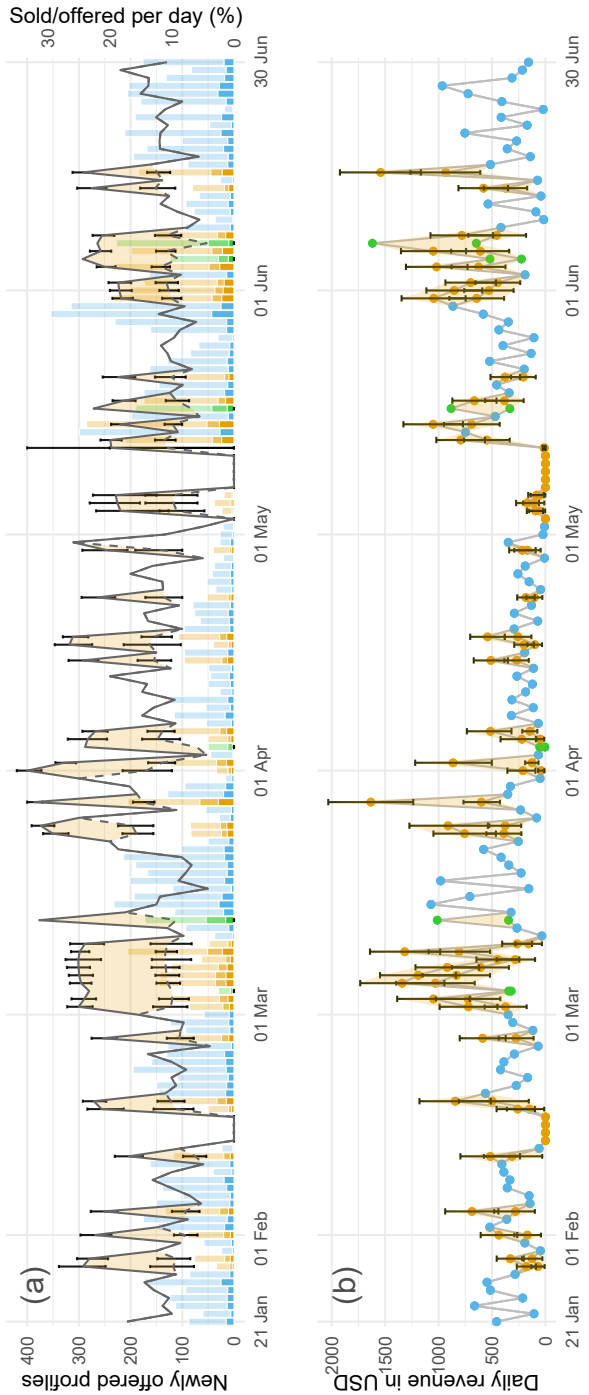


Figure 6.7: Average prices of offered (yellow dots) and sold (blue dots) profiles across geographical regions. Yellow dashed lines indicate 95% confidence intervals of available profile prices.

Overall profile acquisitions and market value

The analysis reports between $1'799$ ($95\% CI = [1'757, 1'843]$) and $2'518$ ($CI = [2'462, 2'575]$) sold profiles out of $17'171$ ($10.5\% - 14.7\%$ of offered profiles sold). Recall that we are collecting a random sample of the actual Genesis Market listings (ref. Section 6.3.2 and Section 9.3 in the Appendix for additional details), and measure only approximately half of the actual sales (ref. Section 6.4.1). To obtain a rough but realistic estimate of actual numbers, the reader can simply scale up reported figures by a factor of 10 (a more detailed review of this factor is provided in the Appendix). Figure 6.8 reports a daily breakdown of the overall sales. Figure 6.8(a) provides an aggregate overview of newly available and sold profiles per day (respectively light-shaded for 'generous' estimates and dark-shaded for conservative estimates for simulated/predicted days, and solid stacked bars for sold profiles) and the respective fraction of sold profiles (dashed line for conservative estimates, solid for 'generous'); Figure 6.8(b) reports daily revenues, reporting values for estimates for both simulated and predicted days. Here we report numbers from the analysis next to the scaled figures in parentheses. Looking at Figure 6.8(a), overall supply sharply varies across periods, with profile provisions ranging between 20(200) to a maximum of about 350(3500) in late May 2021, with sales peaking in the same period to 43(430) profiles per day. Periods of low/no supply are visible: next to market downtimes mid-February and early May, the profiles supply between April and May is very low overall, ranging from 120(1200) to less than 10(100) before completely terminating for 5 days. Interestingly, looking at daily sale patterns (trendline), the aggregate effects of sales during this period do not identify those effects of demand almost matching the offer as observed in the regional breakdown from Fig 6.6, but rather it is true the opposite: the fraction of sold profiles amounts to $\approx 25\%$ at the beginning of April and bottoms to $\approx 12\%$ by the beginning of May; this suggests that attackers still seek profiles with peculiar characteristics and in case of scarcity they are not tempted from less appealing profiles, suggesting they are strategic in victim selection. Some notable peaks are from mid-March to mid-April, and late May and June. Looking at daily revenues in Figure 6.8(b), we see an inflow averaging around $304(3'040)$ USD/day with peaks at approximately $720(7'200)$ USD/day from the conservative estimate, and $399(3'990)$ USD/day with peaks of $1'640$ ($16'400$) USD/day for the generous one. Daily revenues largely reflect sale volumes and appear to cycle between high-demand periods (March and June 2021), and lower-demand ones (Jan/Feb and April 2021).

Estimate of market size and revenue. From our data reconstruction, we estimate (conservatively) that, over the period of 161 days from Jan 21st 2021 (incl.) and Jun 30th 2021, Genesis published overall $\approx 97'655$ advertised profiles, $\approx 20'000$ of which sold within the first day ($95\%\ CI = [19'572, 20'530]$) at an average price of ≈ 27 USD ($[25.83, 28.64]$). Overall, we estimate that the total revenue for Genesis Market during the reported period is of $\approx 540'000$ USD ($[517'729, 574'118]$) (i.e., about $1.2m$ USD/yr, assuming that the observation period is representative of the unobserved one). A less conservative estimate ($TNR = 80\%$) results in $\approx 28'000$ profiles sold ($[27'426, 28'685]$) at an average price of 25.48 USD ($[24.25, 26.77]$), for a total revenue in the observation period of $\approx 715'000$ USD ($[680'312, 750'884]$).



Estimates relative the market sample. To obtain a rough but realistic figure of actual sale revenues and volumes, one can scale reported quantities by a factor of 10. Shaded areas indicate the range between conservative estimates and 'generous' estimates. Simulated sales (ratios) are reported with their respective standard deviations.

Figure 6.8: Measured, predicted and simulated sales volume (a) and daily revenues (b) on a sample of overall profiles.

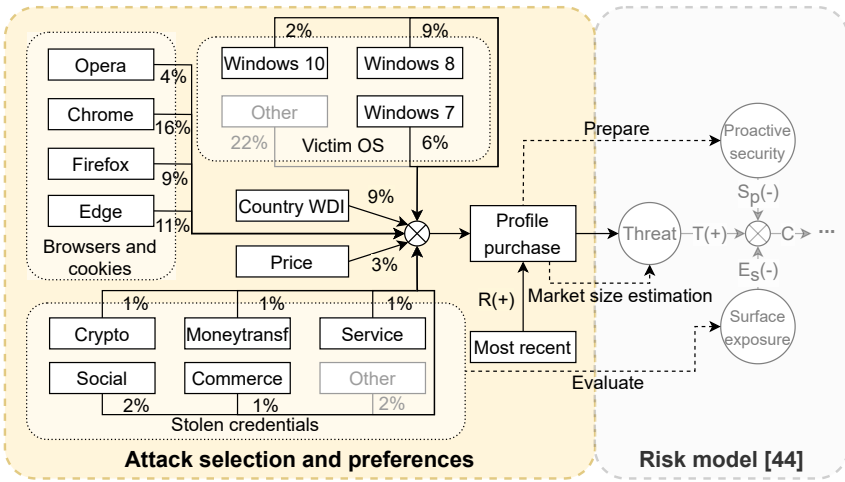


Figure 6.9: Attacker preferences within IMPaaS.

6.5. Discussion

Attack selection and preferences

The first observation emerging from our analysis is that attacker decisions and preferences within the IMPaaS threat model are complex: effects cannot be synthesized and quantified at the level of single factors. Rather, the attacker decision can be better modeled by accounting for the interactions across different profile characteristics. For example, we find that a profile low on resources may still be attractive if running on a recent OS and belonging to a profile from a wealthy country. This suggests that IMPaaS attackers may prefer high chances of success over having a wide attack surface (e.g., potentially targeting many online resources/websites within a profile). Because of this complexity, one cannot quantify and isolate the effect of a rise of one point in a variable alone (e.g., 'WDI') on the odds of purchase (differently, our model can be used directly to evaluate what is the probability of purchase of a specific profile configuration). However, by analysing the relative contribution of each factor across dimensions, and the importance of those dimensions in explaining the observed outcome (i.e., a sale, in terms of the change in R^2 for which that dimension is responsible), we can still derive a first indication of the *relative importance of each factor in the final decision*.[11]

Figure 6.9 reports the results within the overall framework of Woods and Böhme's risk model [272]. The figure offers a breakdown of the original variables involved in the attacker decision process and reports the variance in sales they explain across all dimensions. From

[11]This is different from assigning a given variable a signed coefficient quantifying its effect on odds of sales. Rather, this quantifies the variance in the final decision captured by a specific variable, across dimensions.

the analysis,[12] it emerges that the wealth of the country from which the profile originates is an important factor attackers consider when making a purchase decision (capturing ≈ 9% of its total variance). By contrast, the price of a profile only plays a minor role in the decision (3%), perhaps as a result of the profiles being overall relatively inexpensive. Interestingly, purchase decisions seem to be highly affected by the browser from which the stolen information and cookies originate. Google Chrome accounts by itself for 16% of the variance in the purchase decision, followed by Edge and Firefox (at approximately 10% each). Opera seems to be the least relevant browser in the decision. The high relevance of Chrome in the purchase decision may be confounded by Genesis Market providing their browser extension for Google Chrome itself (ref. Section 6.2), perhaps increasing an attacker's confidence that the purchased profile will work on their setup. Overall, the type of browser and the cookies they come with account for approximately 40% of the overall variance. The OS also plays an important role (≈ 17%), possibly indicating a selection mechanism that disregards older systems, as seems to be consistently (i.e., across all dimensions) suggested by the sales prediction model (Section 6.4.2). Surprisingly, the composition in credentials accounts for a minority of the total variance (≈ 6%), suggesting that these play a relatively minor role in the final decision. An explanation for this may be that most profiles are 'rich enough' in resources of different types, meaning that relative differences across profiles do not impact much the final decision. That is also in line with the notion of rational 'mass attackers' looking for any target for which their attacks will work, as opposed to specific targets [17], particularly when facing high costs to monetize the attack [119].

Perhaps unsurprisingly, we can also conclude that a key factor in the purchase decision is how recent the information within a profile is. An explanation is that attackers may believe that information within more recent profiles (e.g., a token within a cookie) is more likely to still be valid at purchase time. This however emerges only informally from the initial data analysis, as opposed to formally from the sales model (which only accounts for profiles sold on the first day).

Proactive security and surface exposure

Understanding attacker preferences provides awareness about the possible risks connected to `IMPaaS`. For example, an organization could monitor Genesis Market, or any other emergent `IMPaaS` service or provider, to gauge the level of exposure of their employees (e.g., through the presence of an employee-only login portal website amongst available resources) to possible attacks. We note that to do this, the organization needs not to buy specific profiles: Genesis provides the list of the (sub)domains for which credentials are available as part of the profile description. As sold accounts tend to be traded within a day, any preventative action should be taken swiftly and can be enforced only temporarily to minimize negative externalities on final users. For example, when observing the appearance of profiles for that organization (and/or predicting their sale), risk-based authentication mechanisms could be temporarily disabled or hardened to require second-factor authentication in all cases for the

[12] For completeness, in the figure we also report the categories 'Other' for both the `OS` and `Credentials` groups. However, as 'other' is a bin variable for which no clear classification emerges, we refrain from making conclusions. The high relevance of OS=other (i.e., OS is *not* specified) is due to almost all of the associated 122 profiles being sold. That suggests that this is an artifact of the data rather than a specific effect worth capturing.

upcoming period. Similarly, observing or predicting a sale for a profile with credentials for that organization may be communicated to central monitoring services (e.g., a Security Operation Center monitoring the infrastructure) to raise alert levels around suspicious login actions. Further, the geographical information of a profile could inform different branches, for example, to prioritize internal audits looking for affected employees. Further research may look at how to integrate 'live' indicators of compromise from underground markets in security processes. For example, sale predictions could be further used to prioritize specific responses or evaluate risk levels. Finally, an organization could consider investigating the risks posed to their specific RBA configuration. This may be achieved by acquiring profiles featuring compromised corporate (employee) accounts (barring any required legal checks) and identifying the corresponding infected devices in the organization.

Market size estimation

Findings related to the size of underground markets are oftentimes a precursor to law enforcement initiatives such as takedown actions. Evaluating the number of sales of a market is a rare opportunity requiring either market infiltration or usually only coming after the market has already suffered from some shock (e.g., a leak or hack). Our investigation reveals that in approximately six months, Genesis made available data on $\approx 100k$ Internet users; $20k$ have been sold, and therefore likely attacked, by Genesis Market customers in the same period. The sales activity indicates a remunerative business model, especially considering that Genesis is a single-vendor market (as opposed to a market platform [238]).

Lesson learned in measuring underground activities

Whereas our data collection methodology is tailored to Genesis Market it may also inform the design of other measurement methods addressing these or similar challenges. In particular, the community could attempt to address these challenges systematically to produce robust and reusable software for stealth underground monitoring, helping other researchers to tap data from the underground. In retrospect, features that could have eased the data collection process include a system to manage the unreachability of the market; among these, attempting to 'greedily' (i.e., as soon as possible) collect data could be viable in case the target is offline or supporting fallback navigation via Firefox (tunneled via an appropriate VPN service) in case of persistent congestion of the TOR network (Genesis is reachable from the surface web).

Limitations

Profiles sold immediately after appearing on the market may not be captured by our crawlers. We mitigate this problem by employing parallel crawlers, keeping the crawling time at a minimum. Running our crawling during the (east) European night further decreases the chances that profiles will both appear *and* disappear within our window, albeit the market is not reserved for East European customers only. Further, we cannot assure that a profile 'permanent' disappearance may not be due to causes other than a sale. However, the presence of many old, unsold profiles [Campobasso7] makes this unlikely. Additional commentary

on the phenomenon is reported in the Appendix. Similarly, we cannot verify that purchases are not performed by actors other than attackers, for example, researchers or LE. However, given the size of the market and the measured sale trends, it is unlikely that volumes of purchases for 'legitimate' purposes affect the overall analysis. Our simulations implicitly assume that the Genesis Market backend for the data harvesting is independent of the market frontend from which we fetch results; further, we cannot explicitly model the effect of market downtime on sales in our model.

Aftermath and accuracy of our estimates

On April 4, 2023, Genesis Market was shut down during the operation 'Cookie Monster' (more information on market take down are available in the following section); the FBI released the seizure warrant for the clear web domains of Genesis Market [86]. According to the released document, Genesis Market worked as advertised. In May 2022, the investigators obtained a forensic image of Genesis' backend servers; this copy covered Genesis Market's activity spanning from 2018 to May 18, 2022. From this data, it emerges that (1) approximately 1.5 million user profiles have been compromised and advertised for sale on Genesis Market, and (2) Genesis Market's users deposited more than 8 million USD into the market. These numbers appear quite in line with those presented in Section 6.4.2: our study based on 161 days of market activity monitoring report $\approx 100k$ user profiles advertised, and $540 - 720k$ USD in revenues; assuming these numbers as representative of the entire life of Genesis Market, scaling these numbers to 4.5 years of market activity indicate ≈ 1 million user profiles advertised, and $5.5 - 7.3$ million USD in revenues. These results suggest that, despite the limitations of our study and the assumptions on the representativeness of the observed period, our data analysis methodology accurately estimated the market activity and, whereas possible, it could be considered viable when estimating market activity based on censored data.

6.6. Chapter conclusion

In this chapter, we presented a unique data collection and rigorous analysis of data on attackers' profile acquisition on a prominent, still active, cybercrime market for user impersonation at scale. The proposed methodology identifies and addresses general challenges inherent to the problem of monitoring (prominent) criminal underground communities. We reconstruct attacker preferences, profile acquisition trends, and sale volumes, and estimate the overall market revenue. We discuss the implications of our work by integrating the risk model proposed by Woods and Böhme in [272].

Epitaph for Genesis Market: the tale of a takedown

On April 4, 2023, Genesis Market was shut down in a massive law enforcement operation, dubbed 'Cookie Monster', led by the U.S Federal Bureau of Investigation (FBI) and the Dutch National Police (Politie), with the support of EUROPOL and 17 countries. The landing page of Genesis Market's clear web domains was changed to the one reported in Figure 6.10.

Figure 6.10: Landing page of Genesis Market clear web domains, April 4th 2023.

According to the press release of EUROPOL released in April 5th 2023, the operation led to the seizure of the criminal infrastructure, 119 arrests, 208 property searches, and 97 knock and talk measures [88].

The same day, we independently verified the status of Genesis Market and confirmed the status of the clear web domains. However, the deepweb domain resulted to be still online, and the market interface appeared functional. The last stolen profile was advertised the previous day at 16:37 in MSK time, but some older profiles were still receiving updates, suggesting that part of the infrastructure was not seized during the operation. The following day, April 6th 2023, we observed another batch of 241 profiles flowing in the marketplace, although featuring only credentials.

On April 13, Genesis operators made an announcement on some communities, reported in Figure 6.11. In the announcement, they stated that the infrastructure was intact, but confirmed that their clear web domains were seized and that would cause the extension not to work regularly until further notice.

On April 22, Genesis operators posted once again on an affiliated community (Figure 6.12). In this post, they announce that the plugin enabling the impersonation attacks will start

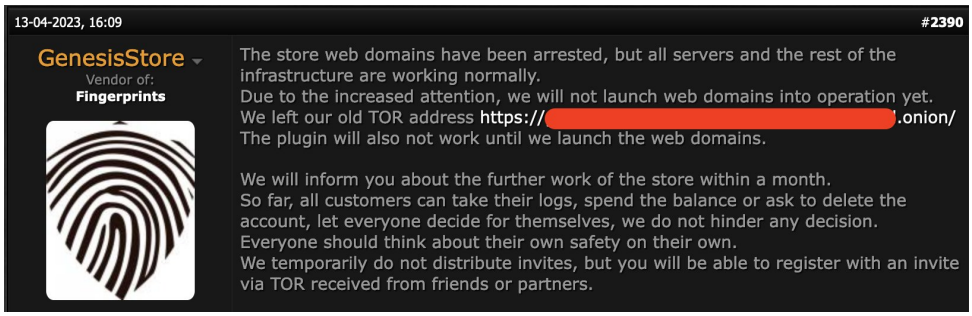


Figure 6.11: Genesis announcement on market takedown.

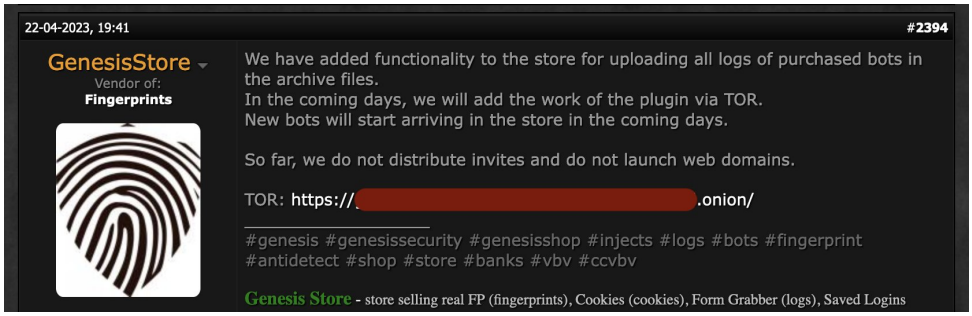


Figure 6.12: Genesis announcement on restored market operations.

working over TOR, circumventing the effects of the law enforcement operation and that new profiles will be offered in the market.

After two months, on June 28, Genesis admins published a new post (Figure 6.13) where they would retire. In the message, they express their intentions to 'sell the store, with all the developments, including a complete database (except for some details of the client base), source codes, scripts, with a certain agreement, as well as server infrastructure [sic.]'.

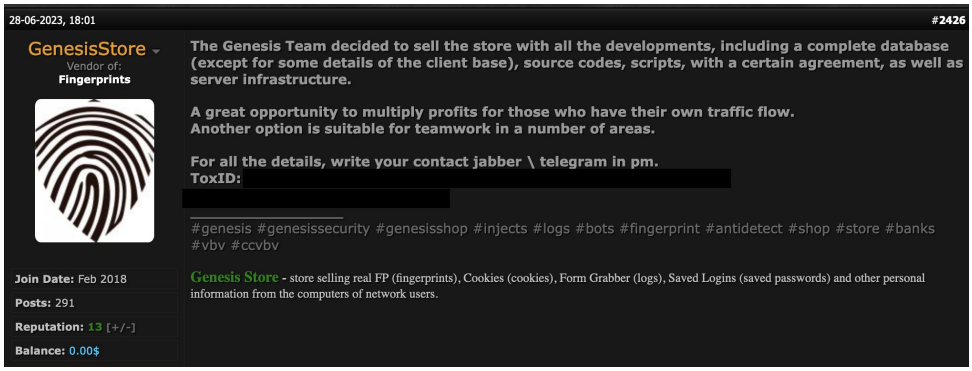


Figure 6.13: Genesis announce their retirement and sale of the platform.

Finally, on July 14, Genesis' operators made their last announcement (Figure 6.14): the market has been sold, and 'the store will be handed over to a new owner next month' (translated).

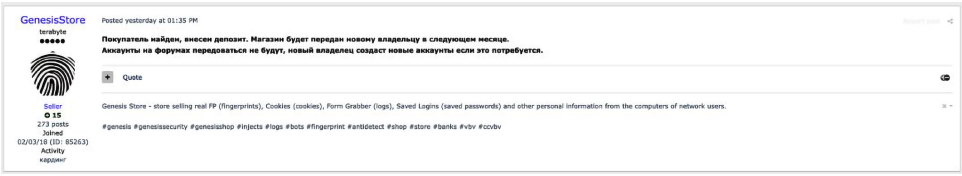


Figure 6.14: Genesis' administrators last post (source: `https://therecord.media/genesis-market-sold-despite-fbi-operation`).

Six months after the sale, we have not identified any sign of a new Genesis-like marketplace in the underground. It remains interesting to see what is going to be born from Genesis' ashes; undoubtedly, the brand is 'burned', but Genesis taught that it is a profitable business, and offers the opportunity 'multiply profits for those who have their own traffic flow' (i.e., reliable supply of infected end-user systems).

III

Characterizing the underground markets that ‘matter’ and research perspectives

7

A general framework to identify prominent underground forum markets

This chapter is based on [Campobasso2]:
M. Campobasso, R. Rădulescu, S. Brons, and L. Allodi
You Can Tell a Cybercriminal by the Company they Keep: A Framework to Infer the Relevance of Underground Communities to the Threat Landscape
22nd Workshop on the Economics of Information Security (WEIS 2023)

In the previous chapter, we characterized the attacker preferences and evaluated the threat levels posed by Genesis Market, an emerging (at the time) and innovative criminal service, capable of convincingly addressing the issues that cybercriminals face when performing credential stuffing attacks. Our analysis identifies Genesis Market as a mature and segregated criminal service, different from other illicit services. In the beginning, after finding an advertisement about Genesis Market, we explored multiple cybercriminal venues to gather additional information about this illicit venture; during this process, we realized that Genesis was advertised only on a handful of them, and we could not find information on other markets, despite those being popular, and having impressive membership size and interaction volumes. The lack of coverage across these markets we have access to appeared counterintuitive and led us to ask 'why those markets?'. Following the intuition that this could be more than a result of coincidence, we observed the characteristics of these markets. Over three markets, two of them require an invite or the payment of a fee to get access, in line with the preliminary findings presented in Chapter 2. Therefore, we decided to investigate their characteristics in more detail. From our analysis, it emerged that these forums are among the most longstanding ones, operate in the Russian cybercriminal scene, and generally appear more structured than others in terms of rules and access policies, with a tiered structure and featuring criminal services that consistently receive positive feedback over time. These

observations lead us to postulate the existence of a selection mechanism used by the operators of mature cybercriminal service providers, allowing them to discern the venues for their unlawful activities.

In light of that, in this chapter, we argue that experienced and motivated cybercriminals scrutinize underground market's characteristics to decide where to advertise their products. We rely on economic concepts from Agency Theory (Adverse Selection and Moral Hazard) to identify the key issues that affect markets and use those to identify the mechanisms implemented by criminal markets that aim at solving these issues. Once identified these mechanisms, we condense them into a general evaluation framework for underground forum markets, agnostic with regards to the five pillars indicated in Chapter 2, and proceed to describe 23 underground forum markets with the proposed framework. To discriminate between markets, we collect information on cybercriminal cases reported from the Department of Justice (DoJ) from 2011 to 2021 and use this information to identify the activity of the accused or arrested cybercriminals across the considered forums. Our findings indicate that markets featuring indicted or convicted criminals show more stable characteristics compared to those that do not, supporting the assumption that prominent cybercriminals may select the markets where to advertise and sell their illicit services.

Link to the created datasets (available under license):
`https://security1.win.tue.nl/doku.php?id=artefacts#data_sharing`

7.1. Introduction

The cybercrime underground is the subject of numerous scientific studies, and it is the source of a conspicuous volume of threat intelligence and threat information employed for a large set of objectives, including situational awareness [208, Campobasso7], criminal behavior [103, 239, 129], informing law enforcement actions and intervention [52, 246, 129, 18], and instrumenting security countermeasures [129, Campobasso7]. To the best of our knowledge, at least part of these applications implicitly assume that if cybercrime-related activities happen in a specific forum, then they must matter to the overall threat landscape. Yet, previous work already questioned whether this really is the case [122]: enabling effective trade in the underground communities is a hard, foundational problem that requires the right incentives and control mechanisms to be in place for 'good sellers and buyers' (as opposed to scammers) to have an opportunity to operate in the market in the first place. In fact, like in any other market, cybercriminals capable of generating actual value seek the opportunity to appropriately price their products and services, and to clearly differentiate themselves from their inept (or plainly fraudulent) competitors [11]. If 'good sellers and buyers' (i.e., those that can supply and consume effective, real attack technology and cybercrime products) are pushed out of the market, the remaining activity, made from mostly wanna-be criminals scamming each other with repackaged old malware or '0day exploits' downloaded from `exploit-db`, will hardly pose any plausible threat. We thus argue that the presence of appropriate mechanisms promoting and supporting the trade of technologies and services with high intrinsic value is a key element of a credible underground market.

In this chapter, we develop an evaluation framework identifying market features addressing *moral hazard* and *adverse selection* issues in underground markets. We then evaluate the composition of these features across a set of 23 underground markets to assess whether differences emerge between 'successful' and 'unsuccessful' markets. As no clear and objective classification currently exists of what a 'successful' market is, we consider whether we can find evidence that trade activity happened in that market in the form of advertisements or interactions from sellers that led to the arrest by law enforcement of at least a cybercriminal for selling criminal products or services. The rationale for this decision is that a law enforcement arrest is a testament to the real-world relevance of the specific crime for which the warrant was issued. Therefore, we consider finding evidence linking that specific crime to trade in a specific forum as evidence that that forum is at least in principle capable of supporting the trade of effective criminal technologies, data, or services. Obviously, *not* finding evidence of that trade does not mean that the market cannot support it. On the other hand, it does highlight the choice of the arrested criminals to trade in a set of markets, but not others[1]. Therefore, we can infer whether convicted criminals providing effective attack technology or services were more likely to sell their goods in market forums with certain features than others. Following this criterion, we find that, in general, successful markets tend to be more similar to each other, over the feature set we define in our framework than to other forum markets. Interestingly, we also find this to be true within groups: that is, successful forum markets are more similar to each other than unsuccessful forum markets are among each other, suggesting that specific features should be present to support effective trade. We proceed with investigating these features in detail and find that, among others, the lack of involvement in trade from market operators in the market they administer plays a positive role in the odds of identifying one or more convicted criminals in the same market. In the discussion, we offer an interpretation of the implications that these features have, and how these could identify different business models for the markets, in accordance with the goals of the market administrators.

Section 7.2 provides a background on the hurdles of conducting trade of criminal goods in marketplaces. Section 7.3 describes the methodology used to select a set of underground communities, label them according to the framework, and then assess whether a community is successful or not. Section 7.4 offers an overview of the framework used to classify communities based on the derived set of features. Section 7.5 presents the results of the applied methodology, in particular the results of the validation process of the framework, and provides a qualitative analysis of the gathered data, and Section 7.6 provides an interpretation of the results. Section 7.7 offers a thorough overview of the limitations of this study. Finally, Section 7.8 discusses the relevant related work, and Section 7.9 concludes the chapter.

7.2. Background

Cybercriminals meet online in forums to trade the byproducts of their attacks, such as leaks and initial access, stolen credentials and credit card numbers, as well as the attack technol-

[1]In this work, we will call these markets 'successful' or 'unsuccessful' for simplicity. However, the reader should keep in mind these considerations when interpreting these terms.

ogy itself and to exchange knowledge and meet new partners [274, 16, 246, 239]. Dozens of communities are scattered across the 'surface' and 'deep' web, and the majority of those are easy to find and get access to. However, prior research suggests a potentially large portion of these markets may not support the trade of valuable products, featuring obsolete malware, expired credit cards, old dumps of leaked passwords, and generally publicly available information repackaged as 'hacking tools' or 'password leaks' [82, 81, 122, 239]. Arguably, the reason for the unconvincing quality of their offer is an effect of the lack of mechanisms that establish trust and regulate trade, which have a key role in any transaction, let alone those between mutually distrustful parties [103, Chap.2], [188, 16, 81, 274, 275, 239]. As initially underlined by Herley and Florêncio [122], in the absence of trade regulation, these markets assume the typical characteristics of 'markets for lemons' [11]: these markets are fraught with quality uncertainty and operate under constraints (and related risks) of strong information asymmetry [16, 275, 122, 15, 239]. These issues materialize both for market participants and administrators, and each group identifies different problem dimensions.

7.2.1. Market participation

The *principal-agent problem* captures the dynamics between two parties (namely, the *agent*, who is the person who acts on behalf of someone else, the *principal*), where each party acts both in function of an agreement (contract) with their counterpart and in the pursuit of self-interests [100, 87]. In general, Agency Theory distinguishes *ex-ante* from *ex-post* risks, depending on whether they materialize before or after the stipulated contract between agent and principal is enforced [51]. Among ex-ante risks, *adverse selection* affects the ability of a buyer to correctly choose the service or product they need. Among ex-post risks, *moral hazard* risks materialize if either party in the contract unilaterally changes their behavior, *de facto* exploiting some weaknesses of the contract, or outright dismissing their obligations toward the other party. This risk is generally present when one of the parties does not bear the full costs of their actions, i.e., there is no expectation of punishment or penalty ('shadow of the future') for the misbehavior.

Adverse selection

Information asymmetry is a common issue in cybercriminal markets; the problem is exacerbated by the (often) impossibility of sellers to provide detailed information about the quality of their products, leading to *adverse selection* (i.e., buyers can not distinguish between products of different quality, or to choose the correct one for their needs). In fact, providing details on the product's effectiveness from the seller (e.g., how a 0day exploit works, validity of credit cards) could spoil the value of the product itself by revealing critical information to reproduce the artifact without purchasing it. In other words, it is difficult for a buyer to assess the quality of an offered product, and buyers can often fall victim to scams. That creates an incentive for scammers, who offer their low-quality products (or even no product at all) for competitive prices. As Akerlof noted in his seminal work on the 'market for lemons' [11], since buyers have no way to evaluate the qualities of the offered products, these markets suffer two major pitfalls: first, cheap (and low-quality) products become indistinguishable from better (and therefore likely more expensive) products; this pushes 'honest' sellers out

of the market as their profit margins quickly become unsustainable (or negative) as the 'competition' can set prices arbitrarily low. Second, the exit of 'honest' sellers from the market spreads awareness among market participants that all products are of bad quality, leading to a failed market.

Moral hazard

Cybercrime benefits from a lower risk of exposure and arrests thanks to the use of privacy-enhancing technologies and the relatively slow response of law enforcement prosecution [170, 82]. However, this creates a challenge for cybercrooks to correctly identify their fellow criminals. In fact, in the lack of social control and signaling mechanisms typical of traditional crime [82, 170, 103] (which may resort to the use of violence to ensure contract compliance [210]), establishing and maintaining trust relationships with co-offenders becomes a hurdle. As an effect, dishonest cybercriminals could see the opportunity to easily join communities to advertise products that will never be delivered, without exposing their identity. When scammers have little to no fear of punishment, potential victims may incur *moral hazard*. Moral hazard is the risk that one of the two parties during a transaction violates the agreed terms during its execution (e.g., the product is not delivered after payment).

7.2.2. Market administration

A different set of dimensions related to how the market is administered, and in particular on what incentives are present to sustain the market.

Cost structure and risks

Maintaining an underground community with (sometimes hundreds of) thousands of members is a costly venture [110, 18]. Furthermore, such marketplaces often are victims of DDoS attacks from competitor platforms [61], and the nature of their content is an appealing target for law enforcement. These problems further increase costs, as it calls for the use of bullet-proof hosting and DDoS mitigation services from the gray markets, which may apply higher fees than commercial solutions due to the shortage of competent providers [62]. Therefore, market administrators need to find the funds to run and scale up their infrastructure as the community grows. Common strategies adopted by market administrators are to ask their members taxes for conducting business, running ad campaigns, and imposing registration fees [107]. Arguably, community contributions (e.g., in the form of account verification fees) to the maintenance of the community can be seen as a positive incentive in making fraudulent options (e.g., exit scams) less attractive.

Further, underground communities face the problem of attracting the interest of threat intelligence companies, researchers, and law enforcement, which monitor their activity to gauge the threat levels they pose, study their functioning, aim at undermining their operations by disrupting the correct functioning of their reputation systems [122, 98], or perform market take-downs. In addition, competitor communities or disgruntled members may attack the market to exfiltrate private data from the forum (e.g., market 'leaks') [188, 205] and damage their reputation. To face these threats respectively CAPTCHAs and anti-DDoS services

are employed to discourage attackers and limit the effectiveness of vulnerability scanners, granting some more privacy to the community members.

Underground markets as business endeavours

In this sense, forum cybercriminal markets can be thought of as businesses. To create a functional marketplace, administrators need to account for the needs of their members (both customers and sellers) to effectively support trade, as well as the cost structure and the operative risks they face. As discussed, reliably running and maintaining a marketplace comes at a cost; hence, market operators need to carefully model their business in function of two stakeholders: themselves, to obtain revenue (and profits), and market participants, to create transaction volume. The Business Model Canvas (BMC) introduced by Osterwalder [204] allows for an overview of an existing business model or supports the creation of a new one. The BMC could be used to decompose the value chain of the market in its key components, to understand how the value added from the market is supported by a careful selection of their suppliers and activities that produce value to customers that need to be reached and taken care of while making a profit. Thus, exploring underground communities from this angle could potentially indicate flaws in the business model of some markets, indicating structural problems that must be accounted for when a new criminal community is established.

7.3. Approach

In this section, we detail our approach to the definition of the framework and its evaluation.

7.3.1. Selection of underground communities

The selection of communities involved in this study is the result of two years of market infiltration and monitoring. Forum markets are included in our selection opportunistically, with network effects (e.g., interactions within the communities pointing at other forums not yet on our list) accounting for the majority of the market identification mechanism. The resulting list includes markets oftentimes studied in the scientific literature, such as `Forum 1`, `Forum 6`, `Forum 23`, `Forum 13`, `Forum 11`, and `Forum 14` [35, 208, 36, 7], as well as less frequently studied (to the best of our knowledge) markets such as `Forum 9`, `Forum 5`, `Forum 3`, and `Forum 12`.

Overall, this study involves the analysis of 23 underground communities, 19 of which are still active and 4 of which are deceased for which a data collection is available; of these, 3 are obtained from data shared by Web-IQ, an industry player active in the domain of cyber threat intelligence sharing; the fourth is a market we had previous access to for which we have a dump. Table 7.1 provides an overview of the forum communities included in this study. A community is indicated as 'alive' if at the date of the study (Feb 2023) the community is active and reachable on the Internet. We report as 'segregated' communities that require vouching, screening, or upfront payment before access to the community is granted. Among these, we had long-standing access to `Forum 9`, via invitation. To gain access to `Forum`

Table 7.1: Overview of analyzed forum communities

Market name	Members	Threads	Posts	Description	Segregated	Alive	First activity
Forum 1[1]	748′348	121′271	3′821′014	Leaked databases, hosting, proxies, combolists, stolen accounts	✗	✗	Apr 2015[†]
Forum 2	224′213	97′308	429′608	Carding, counterfeit documents, traffic	✗	✓	Jun 2013[‡]
Forum 3	52′813	62′170	484′452	IAB, Exploits, malware, malware development, spam, carding, traffic, hosting, counterfeit documents, dating scams	✗	✓	May 2005[†]
Forum 4[2]	43′638	29′137	119′874	Malware development, obfuscation, and AV evasion, hosting, proxies	✗	✗	Aug 2009[†]
Forum 5	117′850	264′564	1′246′883	Wire fraud, cashing out, carding, malware, botnets, hosting, dating scams	✓	✓	May 2005
Forum 6	5′326′374	6′234′742	61′891′796	Malware, botnets, stolen accounts, gaming hacks, fraud schemes	✗	✓	May 2007[†]
Forum 7	85′303	46′101	376′396	Carding, counterfeit documents, stolen accounts	✓	✓	Dec 2010
Forum 8	19′772	4′173	46′288	Carding, hosting, cashing out, stolen accounts	✗	✓	Sep 2016
Forum 9	81′558	201′412	1′275′035	Exploits, malware, malware development, IAB, hosting, spam, traffic, financial fraud	✓	✓	Feb 2005
Forum 10	63′711	12′880	140′440	Carding, stolen accounts, database leaks, porn, financial fraud	✗	✓	Jul 2015
Forum 11	1′717′605	1′479′148	16′155′076	Traffic, hosting, financial fraud, malvertisement, proxies, SEO fraud, PPI	✗	✓	Oct 2005
Forum 12[3]	54′829	?	71′466	Carding, malware	✗	✗	May 2017[††]
Forum 13	147′809	201′189	2′432′625	SEO fraud, proxies, hosting, traffic, malware development	✗	✓	Aug 2001[‡]
Forum 14	4′801′334	1′273′687	38′118′954	Combolists, proxies, stolen accounts, leaked databases, financial fraud, dating scams	✗	✓	Jan 2015
Forum 15	333′147	186′458	2′028′241	Stolen accounts, porn, proxies, combolists	✗	✓	Dec 2016
Forum 16	65′911	12′664	40′877	Malware, financial fraud, hosting	✓	✓	Jul 2017
Forum 17	1′245′284	1′117′023	8′086′074	Financial fraud, porn, carding, proxies	✗	✓	May 2018[‡]
Forum 18	320′964	39′722	4′360′559	Cashing out, drugs, financial fraud, information gathering, weapons, counterfeit documents, money laundering	✗	✓	Oct 2014
Forum 19	45′262	17′017	63′937	Data leaks, drugs, banking accounts, counterfeit documents, weapons	✗	✓	Feb 2016
Forum 20	?	?	?	Carding, travel fraud, hosting, financial fraud	✓	✗	Jan 2021 (?)[††]
Forum 21	7′205	1′449	28′119	reverse engineering, hacking tools, game hacks, malware, malware obfuscation	✗	✓	Apr 2010[†]
Forum 22	311′015	555′208	5′641′640	DDoS, proxy, hosting, game items, carding, PPI, malware, leaks	✗	✓	Nov 2016
Forum 23	3′594′708	769′668	23′176′695	Financial fraud, porn, dating scams, stolen accounts, counterfeit documents, proxies	✗	✓	Mar 2018

All the reported figures have been fetched in Mar 1st, 2023, or otherwise specified with the latest information available; [1]: Feb 23rd, 2022, due to market closure (source `archive.org`); [2]: Aug 31st, 2021, due to market closure (source Web-IQ); [3]: Dec 21st, 2019, due to market closure (source Web-IQ). In the column first activity, we report the registration date of forum administrators, unless otherwise specified; [†]: oldest `archive.org` copy available; [‡]: whois domain registration date; [††]: first available information found online

16 we were screened for knowledge of penetration testing and for our motivation to join the community (for which we gave generic answers indicating interest in potential commercial opportunities). We registered to Forum 20 (currently a deceased market) during a period when free registration to the market was available. To access markets imposing a paywall at registration (Forum 5 and Forum 7), we paid the registration fee[2]. We evaluated the benefit of accessing these communities based on the overall feedback and 'reputation' those had within the communities we already had access to. We purchased premium subscriptions for five communities to get full access to their content, namely Forum 10, Forum 15, Forum 14, Forum 17, and Forum 23. To collect data from the markets, we developed four simple crawlers, configured to interact with the 20 (19 still alive markets now, plus Forum 20 which at the time of our first investigation was still active) marketplaces to identify users of interest (see Section 7.3.3) and related activity. To minimize crawler exposure during the activity, we follow [Campobasso3] and employ the browser instrumentation library Selenium [130] and tbselenium [5]. We provide a discussion on ethical considerations for the data collection in Section 7.3.4.

7.3.2. Framework derivation and instantiation

We structure our framework over the dimensions identified in Section 7.2 (*Market participants:* moral hazard, adverse selection; *Market administration:* cost structure and risks). To finalize our framework, we first identify from the literature mechanisms commonly employed by underground forum markets that map to the identified dimensions (e.g., a reputation system is a mechanism addressing both adverse selection and moral hazard). To identify the relevant literature, we queried Google Scholar using combinations of the following keywords: cybercrime, underground, forums, markets, moral hazard, adverse selection, trust. We limited our research to the top 5 pages of results for each query. We read the papers and included those discussing the operational and economic factors of underground markets, and analyzed the relevant related works of each paper. This led us to identify 23 studies discussing underground market forum characteristics to different degrees. We report the identified mechanisms and map them to the literature in Section 7.4.1. Table 7.2 provides an overview of the mapping.

However, these 'dimensions' are generally high-level features that can be implemented differently and with different strategies by each forum (e.g., some markets may restrict reputation changes to happen only after trade, and others may have no restrictions on who can assign 'reputation points' to a user). To capture these differences, we adopt a two-iterations 'bottom-up' approach whereby for each market feature we first identify within the set of 23 markets under analysis what concrete features are implemented in support of the identified mechanisms, and refine the framework accordingly. During this procedure, we pay particular attention to identifying mechanisms adopted by different marketplaces, and evaluate them in relation to the problem dimensions (Section 7.2) for inclusion in the framework.

Once all features enumerated from the market observations are included in the framework, we re-iterate across all 23 market forums to evaluate the presence of a certain feature in any

[2]We discuss the ethical implications in Section 7.3.4

given market. We follow this process (rather than assigning features to markets as we go by in the first iteration) as some features may only be 'implicitly' present in a market, and be ignored at the first pass. This way we ensure that all marketplaces are evaluated over the same set of features. Each feature can be labeled as `present`, `non-present`, or `unknown`.

To aid the analysis of the identified features from the perspective of the market's 'business proposition', we further map the identified features to the BMC [204]. For details on the mapping, we refer the reader to Figure 7.1 and associated discussion. We report on all the identified features, and their mappings, in Section 7.4.

7.3.3. Framework validation

Ground truth definition

To evaluate whether the presence of the identified market features correlates with or signals the ability of a market to support successful trade of effective criminal technology, we first need to build a ground truth of which markets in our collection can be considered 'successful'. As there is no commonly accepted definition of a 'successful market', we define it as a market for which there is evidence that it is capable of supporting the trade of real, effective criminal technology or services. To do this we want to identify classification criteria that are (a) objective; (b) independent from our analysis; (c) unambiguous; and (d) credibly measurable in our data. Following these criteria, we consider whether (at least one) convicted cybercriminal relied on a specific market to sell technology or services for which they were arrested and convicted. This, of course, limits our definition of success to those markets whose users have been eventually convicted, and risks overlooking 'successful' (perhaps, 'even more successful') markets no member of which has been arrested yet (or for which an arrest cannot be linked to that market). On the other hand, it does speak about what choices those convicted cybercriminals made when deciding in which market to trade their goods. Very importantly, this also *explicitly defines the scope within which the claims of this study should be interpreted*. The adopted criterion is (a) objective, because it considers the real-world impact of the crimes committed and enabled through that marketplace; (b) independent from our analysis, as the investigation leading to a warrant, arrest, and ultimately the conviction of a cybercriminal are run by law enforcement and played no role in the forum selection considered for this study; (c) unambiguous, because the employment of a technology or service leading to a conviction is clearly working and effective; (d) credibly measurable in the data, because trial and warrant documents come rich in information about the reasons for the arrest and any indication of the online presence of the criminal, which can be at least partially mapped back to the criminal's activity in the underground market space. Details on how we process this information follow.

Ground truth measurement

To obtain a list of convicted cybercriminals, we rely on the `arresttracker.com` [65] cybercrime online database[3]. The dataset contains 2775 individual cyber-related crime in-

[3]As of today, the website is down. We tried to get in touch with its administration but without success. We share the dataset at `https://security1.win.tue.nl`.

cidents spread from 1970 to 2021, documented from warrants and indictments from the USA Department of Justice, Europol press releases, and media outlets[4]. An individual entry in the database corresponds to a convicted criminal and contains information fields about their online presence (i.e., their online *alias*) and technical details on the specific crime (e.g., stealing/selling credit card data, offering specific malware or criminal services). For this research, we consider only criminals arrested from 2011, and for which an alias is indicated. We disregard cases before 2011 to limit the biases caused by not accounting for the restricted (and different) forum choices available to cybercriminals in 2011 ($\approx$ 35% forums created before 2011) in our sample. That resulted in 836 aliases of 480 cybercriminals indicted in the 2011-2021 interval. The data available on `arresttracker.com` terminates in December 2021 and the project is abandoned. We define three classes of possible matches:

- *clear hit*: a forum user in any of the considered forum markets whose username corresponds exactly to an alias reported in a conviction report, *and* that presents activity in that forum that closely resembles the crime for which the related criminal has been arrested and convicted. For example, if a username in a forum corresponds to the alias of a criminal convicted for running and renting a botnet, we only consider a *clear hit* if the forum activity of that username is directly related to running and renting a botnet. Furthermore, we check whether the timelines of a user activity on a forum and a conviction are compatible. This is to avoid misclassifying possible copycats.
- *inconclusive hit*: a forum user whose username exactly matches the alias of a convicted criminal but whose activity on the forum (whereas related to cybercrime, e.g., selling stolen data) does not match the reason for the arrest (e.g., running and renting a botnet). Furthermore, we include those cases where the activity on the forum matches the reason for the arrest, but there is no convincing evidence of trade.
- *no hit*: a forum username matches an alias, but no activity related to cybercrime is found.

Each market may have none, only one, or multiple matches for each class. We consider 'clear hits' as the strongest evidence that a convicted criminal chose a specific forum market (or multiple ones) to trade effective criminal technology.

Example of 'hits'

To exemplify, we consider the case of Alexey Klimenko, accused of being part of the Infraud Organization as the provider of bulletproof hosting to create, operate, maintain, and protect their own online contraband stores [74]. According to the superseding indictment, dated Oct 31st, 2017, Alexey Klimenko was operating under the alias of 'Grandhost'. We found account usernames that match the alias 'Grandhost' on six marketplaces in our data: `Forum 4`, `Forum 9`, `Forum 7`, `Forum 5`, `Forum 3`, and `Forum 13`; for `Forum 7`, the registered user exists but has no messages nor activity in the forum. Therefore, we classify this alias in that market as a 'no hit'. In `Forum 5` and `Forum 13`, we find evidence that the user 'Grandhost' interacted with members of the community on topics related to bulletproof hosting, but we could not find any evidence of trade of services or products on those

[4]We discuss the limitations of this dataset in Section 7.7

forums linked to bulletproof hosting for this user. Therefore, we classify the occurrences of 'Grandhost' in `Forum 5` and `Forum 13` as 'inconclusive hits'. On the three remaining marketplaces (`Forum 4`, `Forum 9` and `Forum 3`) we found three threads created by the user 'Grandhost' titled (with little variations of) 'Bulletproof Hosting Service' (translated from Russian). In those, the author advertises services related to dedicated servers, VPS/VDS, bulletproof domains, and SSL certification, providing a range of locations for their servers, and exemplifying some of the possible (unlawful) uses they allow. The first advertisements dated February 2010 in both `Forum 9` and `Forum 3`, and a similar one follows the same year in July on `Forum 4`. Up until February 2016, on all three markets, product advertisement is updated by 'Grandhost' when new offers or features are available to customers; this suggests a long-standing commercial activity for that user in all three forums. The last recorded activity in our data for the user 'Grandhost' is in June 2016, on `Forum 9`. The most recent evidence of trade for Grandhost's offerings we find on `Forum 9` in the form of user (positive) feedback left in the forum in July 2017. Currently, the account for 'Grandhost' on `Forum 9` appears as 'deactivated'. We note that the timeline of the forum activity is also compatible with the timeline of the conviction. Hence, we classify the alias 'Grandhost' as a 'clear hit' for `Forum 4`, `Forum 9`, and `Forum 3`.

We repeat this process for all username-alias matches we find in our data. Section 7.5.1 provides an overview of results.

Evaluation

To evaluate whether the identified feature set correlates to the ground truth, we compare whether (a) 'successful' forums are more similar to each other over the defined feature set than they are to 'unsuccessful' forums; this test allows us to evaluate whether specific feature compositions are proper of 'successful' markets but not of 'unsuccessful' markets. Further, we test whether (b) 'successful' forums are more similar to each other than they are with 'unsuccessful' ones; this test allows us to evaluate whether the feature set of 'successful' forums tends to be more stable than the feature set of 'unsuccessful' forums. To evaluate similarity (for both tests), we compute the Jaccard distance for all forum pairings and employ two one-sided Wilcoxon rank sum tests to evaluate the hypotheses above.

Further, we explore whether we can find differences at the level of specific groups of features (as defined by the BMC), as well as at the level of the specific features between the two groups of forums. To do this, we constrain all features we identify to have the same 'direction' (i.e., all should contribute positively to the mitigation of either adverse selection or moral hazard). Under this constraint, we perform a set of Fisher exact tests evaluating the count of forums of each type ('successful', 'unsuccessful') for which the tested set of features is present.

For all tests, we consider a statistical significance level $\alpha = 0.1$ as the threshold for significance. We note that, given the inherently qualitative nature of this study, we consider statistical tests as merely a guide to drive the discussion around our findings, rather than a quantitative assessment of their statistical robustness.

7.3.4. Ethical aspects

To conduct this study, we had to gain complete access to a number of underground communities and crawl them, which required us to infiltrate, sometimes interact with community members, or even pay a fee for registering or to get full access.

Crawling

Our crawling activity was performed in a way that: (1) our crawlers were solely verifying the existence of usernames in the forums via the appropriate search function, (2) we relied on the Tor network only when the target community was only reachable via it, (3) we limited the bandwidth usage of the crawler to be more similar to a human when fetching new content (especially when the use of Tor was necessary), in accordance to the considerations of [Campobasso3], helping us maintaining stealth and avoiding compromising the stability of the target [52]. Even though collecting data from underground marketplaces could be considered an activity not in agreement with their terms of service [209], studies highlight that the societal benefit of studying cybercriminal ventures outweighs these concerns [176, 34].

Paying access

During our investigation, we identified a number of forums requiring to pay a fee to register or to access portions of the forum. Respectively, from discussions in affiliated communities or within the same ones, we made considerations on whether the benefit from obtaining access could justify the expense. As pointed out by Benjamin [34], when the information that could be made accessible could be of great value for the research (in our case, to avoid undermining the internal validity of our study), this should be considered and discussed with the relevant ERB. For paywalls at registration, we only considered two markets for which subscription fees did not exceed 100 USD; in the case of private sections, we often had access to a 1-month premium subscription for $\approx$ 10 USD. In cases where fees were higher or came with additional problems (e.g., `Forum 11` requires a fee of 147 USD and 100 posts to obtain the upgrade), we refrained from proceeding with the purchase. We discussed our project with our institutional ERB, and we received approval with reference number ERB2021MCS1.

7.4. The Framework

7.4.1. Identification of market mechanisms addressing moral hazard and adverse selection

In this section, we identify from the literature the key mechanisms forum market (administrators) put in place and link them to foundational dimensions identified in Section 7.2. Table 7.2 outlines the identified mechanisms and the rationale for the mapping.

Table 7.2: Mapping market mechanisms to adverse selection, moral hazard, cost structure and risks.

Foundational aspect	Key issue	Market mechanism	Mapping rationale
Market participation	Adverse selection	Reputation systems	The presence of a reputation system is oftentimes enabled by a feedback mechanism that allows potential buyers (and possibly sellers) to assess the nature of the other party, as well as of their products (e.g., through product reviews or open feedback from previous buyers).
		Interaction model	The interaction between community members enables communication between buyers and sellers, helping customers in the product assessment process and sharing information with future customers (in case of public interaction).
	Moral hazard	Reputation systems	A reputation system may serve as a 'punishment mechanism' for misbehaving users, and provide an indication of future (expected) behavior.
		Restricted access model	Limiting access to the forum market to selected users that have either been vouched by existing members and/or the forum administrators or that committed to their participation in the forum via payment of an entry fee.
		Escrow for payments	The presence of escrow services may help to establish trust after one of the two parties commits to his part of the agreement, thus serving as a fallback mechanism to avoid scams caused by, for example, undelivered or non-functional goods for which a sum of money has already been transferred (in this case, to the escrow) by the buyer.
		Dispute resolution system	A dispute resolution system may allow market participants to voice their claims and seek justice for unfair behavior. Importantly, it enables punishment mechanisms (such as labeling a user as a scammer and/or banning them from the community), which may deter unfair behavior.
		Rule system	The presence of rules on trade and punishments in case of misconduct, together with their enforcement, may act as a deterrent for misconduct.
Market administration	Cost structure and risks	Market business plan	Clear revenue strategies for market administrators should guarantee that they have the right incentives to continue running the market as a source of revenue, rather than exit-scamming.
		CAPTCHAs and DDoS protection	The presence of custom-created and commercial CAPTCHAs poses a threat to the execution of a crawler, limiting the possibility of exfiltrating data from the market, while DDoS protection could avoid costly downtimes.

Restricted access model

Different access mechanisms regulate the influx of members to a community or parts of it. Some studies show that markets present different levels of segregation, ranging from open access markets where registration only requires a valid email address [122, 15, 82, 107], up to access mechanisms that require the payment of a fee and perform 'background checks' on the applicant [15, 257, 107, 157, 129, 275, 246]; these access mechanisms may be enforced to grant access to the forum as a whole [81, 170] or only to portions of it [188]. While stricter access mechanisms pose a barrier for new members, these allow screening new applicants by proving their intentions either economically (i.e., by paying a registration fee), by receiving 'vouches' from members of the community who grant on the intentions or identity of the new member, or via interview. These barriers potentially discourage 'dishonest' merchants from joining the community by increasing the costs associated with the planned fraudulent activities [129], thus potentially creating a more selective and trustworthy community [122, 15, 16, 275].

Dispute resolution system

The lack of accountability for misbehavior in the underground calls for strategies to take action against miscreants and to settle disputes [81, 170, 107]. Dispute resolution systems aim at settling a disagreement between two (or more) community users, generally caused by one of the parties not adhering to the conditions agreed upon in trade. Typically, the offended part of the claim (the plaintiff) can open a request for arbitration and has the burden of proof to support their claims; the defendant will be called to respond and to prove that they acted according to the agreement. The dispute is generally mediated, if at all, by community administrators or moderators. Eventually, the parties could settle on a new agreement or, in case of no deal, the party found guilty by the arbitrator is punished with a loss in reputation or with a ban from the community [107]. It is worth noting that the mere presence of an arbitration section in a market may not be enough; the mediators should intervene to take action swiftly and should be impartial, as a biased mediator could render an unfair verdict [16].

Reputation systems

Reputation systems aim at creating conditions in which users with a higher reputation are perceived by other users as being more trustworthy, resulting in a greater willingness to trade [164]. In other words, reputation systems help mitigate information asymmetry problems [81, 82, 243, 275] by indicating the quality of the offered products, thus easing customers in assessing the products' properties ahead of purchase [274, 129]. Similarly, reputation systems can serve as a mechanism to 'punish' users who misbehave [238], potentially discouraging dishonest members from participating in trade. That is particularly important as it promotes predictable, 'honest' behavior in the community (as opposed to capitalizing on misbehavior) [107, 170], and may enable a more aggressive pricing strategy from reputable members [274], and hence greater economic returns. Similarly to restricted access models, obtaining a privileged status in a community may be subject to the payment of a fee or the scrutiny of the community administrators [107, 246]. For example, administrators may require the provision of product samples to verify their quality [132, 107, 246] and ap-

ply different fees according to the type of product to sell. On the other hand, the literature shows that reputation systems can also be abused by market participants [239, 246]. For example, in Sybil attacks, a user creates multiple identities in a forum to artificially increase the number of positive reviews associated with their main identity. Forum mechanisms can be put in place to alleviate this risk, for example, by allowing the assigning of positive/negative feedback on a user reputation only after trade [246].

Mitigation of perverse incentives

In underground communities, the problems of trust are not only limited to the interaction among peers. Administrators are oftentimes in the powerful position of controlling user funds (e.g., through escrow services provided by the market) while having to deal with the cost of running a marketplace, both in terms of efforts and economic expenses [129]. The presence of clear and transparent revenue streams for market administrators can play a role in discouraging actions such as exit scams. Revenue may be generated, for example, from the payment of fees to earn a status for both regular users and sellers [107], transaction costs or fees [107, 170], and by providing advertisement space on the platform [107, 129].

Rule system

Next to regulatory mechanisms to establish trust in trade, markets can provide a list of enforceable rules to guide trade on the market [81, 170] and present a hierarchical structure with members in charge of enforcing those by moderating content and taking actions against offenders [275]. Rules can specify how contracts and agreements are enforced and who (and which users in a forum) are the authorities acting in case of non-compliance [170, 227]. Rules can be relatively specific, for example stating specific procedures to advertise products and engage in trade, or very loose (e.g., 'scamming is not allowed').

Strong authentication and anti-bot features

Law enforcement agencies, threat intelligence operators, and researchers closely observe the activity of underground communities, often with the support of automated crawlers to extract and aggregate data on ongoing criminal operations. It is not uncommon to find messages from community administrators warning about the presence of law enforcement and bots whose purpose is to monitor activities [264]. Sometimes, law enforcement operators clearly state their presence on the market as a deterrent [4]. Market operators employ a range of techniques to limit these undesired activities [257, 69, Campobasso3]; for example, they may employ CAPTCHAs to hinder automated monitoring and may offer their members strong authentication mechanisms limiting unwanted access to user accounts.

Escrow systems

Numerous marketplaces employ trade protection mechanisms, such as escrow services [238, 170, 157, 107, 275], to prevent users from falling victim to fraud. Generally, a party external to the trade (e.g., a market admin) acts as the escrow service. This party may hold payments until both parties confirm the other has fulfilled their part of the contract. Community users

are generally free to use their preferred escrow services; however, some communities actively promote or provide escrow services on their platform to keep complete control in case of dispute (and potentially to earn a commission).

Interaction model

How users interact by writing and reading feedback or comments on products can significantly affect the availability of related information [157, 71]. Most forums allow unconstrained interactions in the public space of the forum, allowing buyers to ask for additional information from sellers, and to send positive or negative signals to their peers regarding specific products or users [275]. In some systems, parts of this information are hidden and only available after payment of a fee (paid with a forum currency or with the account balance). Further, some markets encourage or provide private channels for users to interact outside of the public eye [16, 274].

7.4.2. Framework construction

An enumeration of the identified features is provided in Figure 7.1, showing their mapping to both the BMC and to the foundational dimensions identified in Section 7.2 (*Market participants:* moral hazard, adverse selection; *Market administration:* cost structure and risks. As a visual guideline, the left-hand side of the BMC (Key Partners, Key Resources, and Key Activities) identifies the 'needs' of the business model examined (that is, the suppliers, the physical or intellectual resources, and the activities that support the business). The right-hand side of the canvas identifies business aspects for Customer Relationships, Channels, and Customer Segments, which identify the target of the business examined, how to reach new customers, and how to maintain relationships with existing ones. In the middle, the Value Proposition represents the added value of the enterprise, in the form of a product or service created and delivered to customers. Finally, at the bottom of the canvas, the Cost Structure and Revenue Streams provide a breakdown of the cash flows of the business. For the sake of brevity, and because most of the identified features are self-explanatory, definitions for each feature are reported in the Appendix. Following is a brief explanation of the mapping rationale of the market mechanisms to the BMC. To build the association, we follow the guidelines reported in Section 7.3.2.

Key Partners, Key Activities, Key Resources. This area mostly covers the mechanisms employed by markets to select their Key Partners (sellers) by employing a *restricted access model*, punish them in case of misbehavior via a dispute resolution system, in accordance with trade rules (*rule system*) and give its participants a *reputation system* to identify honest sellers selling quality products. To further limit moral hazard, a market could provide an *escrow system*. To improve their Value Proposition, markets could include in their Key Resources means to mitigate the risks of account hijacking attacks and to provide more privacy to their participants, such as *strong authentication* and the implementation of *anti-bot features*. It is worth noting that guaranteeing market administrations a constant influx of money via the payment of recurrent fees from market participants is beneficial to mitigate *perverse incentives*.

Customer Relationships, Customer Segments, Channels. Features related to buyers and

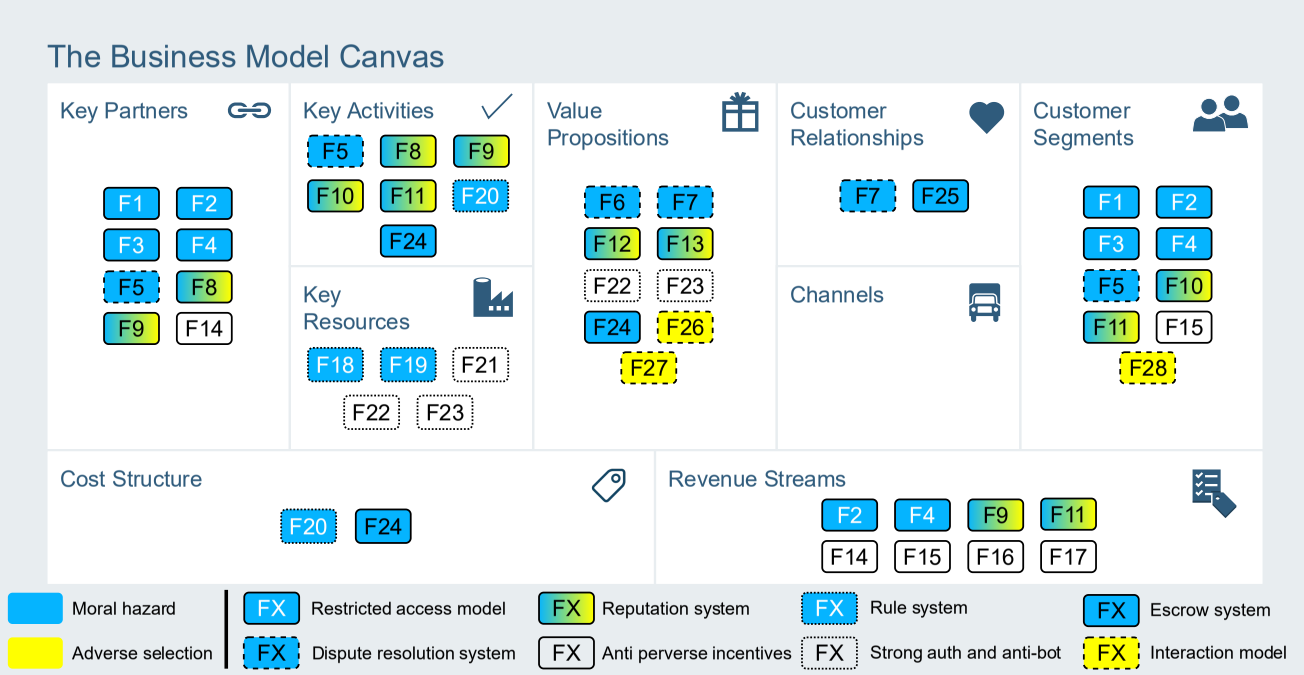


Figure 7.1: Business Model Canvas populated with the identified market features.

Restricted access model
F1. Restricted sections - pull-in
F2. Restricted sections - payment
F3. Restricted registration - pull-in
F4. Restricted registration - payment

Dispute resolution system
F5. Scammers banned
F6. Working dispute resolution system
F7. Neutral mediator

Reputation system
F8. Seller status - verification
F9. Seller status - payment
F10. User status - verification
F11. User status - payment
F12. Reputation change on trade
F13. Reputation change by VIP

Anti perverse incentives
F14. Seller status - recurrent fee
F15. User status - recurrent fee
F16. Escrow fee
F17. Sponsored ads

Rule system
F18. Clear trade rules
F19. Moderators roles
F20. Active moderation

Strong authentication and anti-bot
F21. 2FA available
F22. CAPTCHA on authentication
F23. CAPTCHA on access

Escrow system
F24. Escrow available
F25. Escrow recommended

Interaction model
F26. Public interaction
F27. Private interaction
F28. Pay to show content

the establishment of trust relationships are mapped to the right side of the BMC. Similarly to the previous paragraph, markets can identify their Customer Segments by screening and selecting market participants with the use of strictly *restricted access models*. Honest participants could benefit from a higher *reputation*; markets could decide whether this status should be earned by generic *interactions* with other members or as a result of trade thus mitigating the problems related to adverse selection, or with the payment of a (*recurrent*) fee, thus supporting the market over time and mitigating *perverse incentives*. Markets commit to their customers and the general health of the community by offering a *dispute resolution system* that impartially punishes dishonest members and provides guidelines and recommendations for the use of *escrow systems*.

Value Proposition. Depending on how a market screened its participants, the resources allocated and the activities performed to support the business, and the relationship it has with its customers, a different *Value Proposition* may emerge. A market could offer its participants a working and impartial *dispute resolution system*, together with an *escrow system* that guarantees that the honest party obtains their money in case of disagreement. Also, a market could implement a transparent *reputation system* to promote quality *interactions* to create signals among peers and could commit to keeping unwanted visitors out of it with the implementation of *anti-bot mechanisms*.

Cost Structure, Revenue Streams. Features related to revenue streams and administration risks are grouped under Cost Structure and Revenue Streams. Intuitively, a market's revenue depends on the presence of (recurrent) fees for both sellers and buyers. Some fees could be imposed during the screening of new market participants (*restricted access model*), while others in exchange for privileged statuses (*reputation system*). Finally, the implementation of smaller, periodic fees for services and advertisement provides a market's administration with a constant influx of money, creating revenue and thus offering incentives to keep the market functional and attractive for new members, mitigating *perverse incentives*.

7.5. Results

7.5.1. 'Hits' within our market selection

From a total of 836 aliases associated with convicted cybercriminals, 380 had at least one account on one of the analyzed forums with an identical name. Thus, 456 aliases are automatically labeled as no hits. The remaining 380 aliases allowed us to identify, across all forums, 32 instances that resulted in clear hits (25 distinct aliases), 113 instances of accounts that resulted in inconclusive hits (88 distinct aliases), while the remaining 225 aliases resulted in accounts with unrelated activity or no activity whatsoever (no hits). Figure 7.2 reports an overview of the distribution of the hits across the examined forums. Among the clear hits, three aliases were identified as belonging to the same person in three or more forums. `Forum 7`, `Forum 9` and `Forum 2` account for $\approx$ 60% of clear hits we could identify. An interesting aspect shared across the marketplaces with the highest number of clear hits is longevity: `Forum 7`, `Forum 9`, `Forum 2` and `Forum 5` are longstanding underground communities that have been operating for 9 years or more. This could possibly indicate that

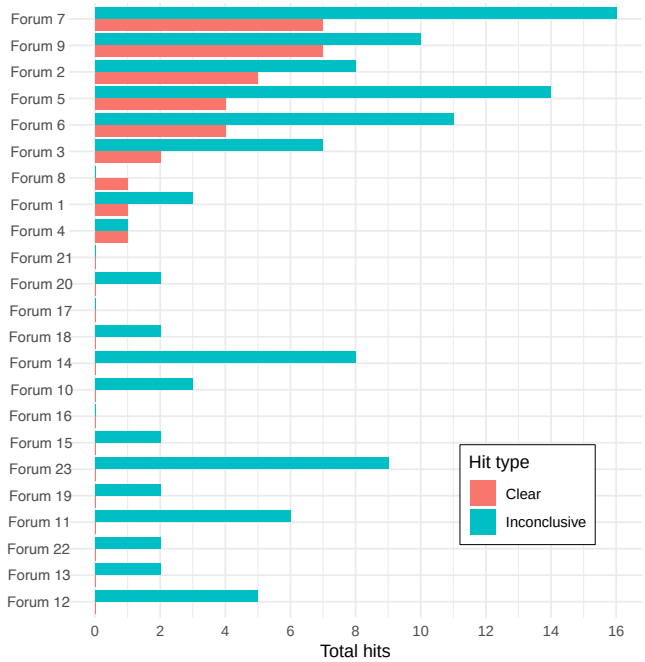


Figure 7.2: The number of clear and inconclusive hits over criminal communities

their success as criminal venues comes from having convincingly addressed trust issues over multiple years of experience. Observing the top four markets for inconclusive hits (`Forum 7`, `Forum 5`, `Forum 6`, and `Forum 9`), accounting for $\approx$ 45% of them, we note that they also feature a rather high number of clear hits ($\approx$ 70%), possibly indicating that some content from those members labeled as inconclusive hits could have been deleted. In fact, interestingly, the forums `Forum 2`, `Forum 5` and `Forum 7` seem to have performed a shadow-banning of some of their members (respectively 9, 9, and 16 aliases); when searched for them, we were shown a different error page than the one shown when the username or associated content could not be found. For all three forums, we could identify the URL that would have displayed the page associated with the shadow-banned alias; we accessed it and labeled them accordingly. It is worth noting that the three markets shadow-banning aliases are among those that scored the highest number of hits. That suggests that the administration is aware of their exposure and take actions to mitigate collateral damage following law enforcement operations. Similarly to shadow-banning, `Forum 2`, `Forum 8`, `Forum 7`, `Forum 5` and `Forum 15` seem to have purged content from the searched accounts in 7, 6, 3, 1, and 1 cases, respectively. In fact, for the investigated forums, simply banning an account does not remove the content posted by that author. We noted the absence of content in some profiles from the mismatch in the count of posts associated with each account and the posts we could find, and we verified that the discrepancy could not be caused by the removal of content by the authors themselves, as it is not allowed in any of the platforms. In this case, however, in the absence of content, these aliases were labeled as no hits.

7.5.2. Market features

We proceed to label the selected markets according to the framework. In Table 7.3, we report an overview of all communities with their corresponding features. The upper part of the table reports successful markets. The bottom of the table reports the aggregated fractions of features available per group (as identified from the BMC), calculated both for successful and unsuccessful markets. For the purposes of the analysis and to discuss their relevance, we do not include in the said fraction non-binary features (F16 - escrow fee, F19 - moderator roles) and variables with constant values (F26 - public interaction). Furthermore, we compute the correlation matrices to identify potential multi-collinearity across features within each group and to remove the problematic ones. Among the 23 selected markets, 4 (`Forum 1`, `Forum 4`, `Forum 12` and `Forum 20`) are no longer reachable. For `Forum 20`, we could rely on partial labeling of the market we performed at a prior version of the framework, and for which we already investigated the available aliases; for `Forum 12`, `Forum 4` and `Forum 6` we had access to historical data provided from an industrial partner, Web-IQ, which allowed us to perform partial labeling and investigate on the presence of criminals on the market; in addition, for `Forum 4` and `Forum 6`, we relied on `archive.org` to explore some portions of the market (e.g., community rules) that were not collected from Web-IQ. For the remaining 19 markets, we could confidently verify the presence of the investigated features for the majority of cases. Nonetheless, sometimes, even when having access to said markets, it was not possible to clearly decide on the presence of some features. For example, `Forum 10` features sections where 'Trusted & Verified Sellers' are allowed to post. However, we could not identify any clear mention of how to become a trusted seller in the market.

From the framework, we identify some substantial differences across groups between successful and unsuccessful markets. A set of one-tailed Fisher tests for the alternative hypothesis that successful markets have a greater count of features within a group than unsuccessful markets finds at least marginally statistically significant differences for the groups Key Partners ($p - value = 0.032$), Key Resources ($p = 0.074$), and Customer Relationships ($p = 0.083$). In particular, the presence of features in Key Partners and Customer Relationships is more frequent for successful markets while, perhaps surprisingly, they are less frequent for Key Resources. The features collected in Key Partners indicate a greater tendency for successful markets to be more segregated, screening or vetting their community base, creating tiers of access to portions of their markets than unsuccessful markets. The importance of Customer Relationships is mostly attributable to F7: neutral mediator (which turns out to be the only statistically significant variable when considered alone; $p = 0.017$), indicating that criminals may value impartiality in the administration of markets[5]. A closer look at the considered variables in Key Resources shows that CAPTCHAs are more frequent for unsuccessful markets and that they feature more often the possibility to enable 2FA, perhaps indicating a greater risk perceived from those markets to be victims of DDoS attacks and other abuses (as it is the case of markets promoting booter services, suffering attacks from their competition [223, 144]).

[5]We note that, for `Forum 18`, we conservatively assign the ✓for F7 as the size of the forum made unfeasible the verification of administrators not being involved in trade activities.

	Key Partners									Key Resources					Key Activities									Value Proposition												Cst. Rel.		Customer Segments									Revenue Streams										Cost Str.	
	F1	F2	F3	F4	F5	F8	F9	F14		F18	F19	F21	F22	F23	F5	F8	F9	F10	F11	F20	F24		F6	F7	F12	F13	F22	F23	F24	F26	F27		F7	F25		F1	F2	F3	F4	F5	F10	F11	F15	F28		F2	F4	F9	F11	F14	F15	F16	F17		F20	F24
Forum 1	-	✓	-	-	✓	-	-	-		-	-	?	✓	✓	✓	-	-	-	✓	✓	✓		✓	✓	✓	-	✓	✓	✓	✓	-		✓	✓		-	✓	-	-	✓	-	✓	-	✓		✓	-	✓	-	✓	-	-	-		✓	✓
Forum 2	✓	✓	-	-	✓	✓	✓	✓		-	-	-	-	✓	✓	✓	✓	-	✓	✓	✓		✓	✓	-	✓	-	-	✓	✓	✓		✓	-		✓	✓	-	-	✓	✓	-	-	-		✓	-	✓	-	✓	-	-	✓		✓	✓
Forum 3	✓	✓	-	-	✓	-	✓	✓		✓	-	✓	-	-	✓	-	✓	-	-	✓	✓		✓	✓	-	-	-	-	✓	✓	✓		✓	✓		✓	✓	-	-	✓	-	-	-	-		✓	-	✓	-	✓	-	-	✓		✓	✓
Forum 4	-	-	-	-	✓	-	?	?		?	-	?	-	-	✓	-	?	-	?	✓	?		✓	✓	-	-	-	-	?	✓	✓		✓	-		-	-	-	-	✓	-	?	?	?		-	-	?	?	?	?	-	✓		✓	?
Forum 5	✓	✓	-	✓	✓	✓	✓	-		✓	-	✓	-	✓	✓	✓	✓	-	✓	✓	✓		✓	✓	✓	-	-	✓	✓	✓	✓		✓	✓		✓	✓	-	✓	✓	-	✓	✓	-		✓	✓	✓	✓	-	✓	-	✓		✓	✓
Forum 6	-	-	-	-	✓	✓	✓	-		✓	-	✓	-	-	✓	✓	✓	-	✓	✓	-		✓	-	✓	-	-	-	-	✓	✓		-	-		-	-	-	-	✓	-	✓	✓	-		-	-	✓	✓	-	✓	-	✓		✓	-
Forum 7	?	?	-	✓	-	-	✓	-		-	-	-	-	-	-	-	✓	-	?	-	✓		-	-	-	-	-	-	✓	✓	✓		-	✓		?	?	-	✓	-	?	-	?		?	✓	✓	?	-	?	-	✓		-	✓
Forum 8	-	-	-	-	-	-	✓	?		-	-	-	-	-	-	-	✓	-	-	-	-		-	-	-	-	-	-	-	✓	✓		-	-		-	-	-	-	-	-	-	-	-		-	-	✓	-	?	-	-	✓		-	-
Forum 9	✓	-	✓	✓	✓	-	?	?		✓	-	✓	-	-	✓	-	?	-	-	✓	✓		✓	✓	-	-	-	-	✓	✓	✓		✓	✓		✓	-	✓	✓	✓	-	-	-	-		-	✓	?	-	?	-	-	✓		✓	✓
Forum 10	-	✓	-	-	-	?	?	?		-	-	✓	-	-	-	?	?	-	?	-	✓		-	-	-	-	-	-	✓	✓	-		-	-		-	✓	-	-	-	?	?	-		✓	-	?	?	?	?	-	✓		-	✓
Forum 11	✓	✓	-	-	✓	-	✓	✓		✓	-	✓	-	✓	✓	-	✓	-	✓	✓	-		✓	✓	✓	-	-	✓	-	✓	✓		✓	-		✓	✓	-	-	✓	-	✓	✓	-		✓	-	✓	✓	✓	✓	-	✓		✓	-
Forum 12	-	-	-	-	-	-	-	-		?	-	?	?	?	-	-	-	-	?	-	-		-	-	-	?	?	-	✓	✓	-		-	-		-	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	-		-	?	-
Forum 13	✓	-	-	-	✓	-	-	-		✓	-	✓	-	✓	✓	-	-	✓	-	✓	✓		-	-	✓	✓	-	✓	✓	✓	✓		-	✓		✓	-	-	-	✓	✓	-	-	-		-	-	-	-	-	-	-	-		✓	✓	✓
Forum 14	-	✓	-	-	✓	-	?	?		-	-	✓	✓	✓	✓	-	?	-	✓	✓	✓		✓	-	✓	✓	✓	✓	✓	✓	✓		-	✓		-	✓	-	-	✓	-	✓	✓	-		✓	-	?	✓	?	✓	-	✓		✓	✓
Forum 15	-	✓	-	-	-	-	✓	✓		-	-	-	-	-	-	✓	-	✓	-	-	-		-	-	-	-	-	-	-	✓	✓		-	✓		-	✓	-	-	-	-	✓	✓	-		✓	-	✓	✓	✓	✓	-	✓		✓	-	-
Forum 16	-	-	✓	-	-	✓	-	-		-	-	-	✓	-	-	✓	-	-	-	-	✓		-	-	-	-	✓	-	✓	✓	✓		-	✓		-	-	✓	-	-	-	-	-	-		-	-	-	-	-	-	-	-		✓	-	✓
Forum 17	-	✓	-	-	✓	-	✓	-		-	-	✓	✓	-	✓	-	✓	-	✓	✓	✓		✓	-	✓	-	✓	-	✓	✓	✓		-	✓		-	✓	-	-	✓	✓	✓	-	-		✓	-	✓	✓	-	✓	-	✓		✓	✓
Forum 18	-	✓	-	-	✓	-	✓	-		✓	-	✓	-	✓	✓	-	✓	-	✓	✓	✓		✓	✓	-	-	-	✓	✓	✓	-		✓	✓		-	✓	-	-	✓	-	✓	-	-		✓	-	✓	✓	-	-	-	✓		✓	✓
Forum 19	✓	-	-	-	✓	✓	✓	-		-	-	✓	-	✓	✓	✓	✓	✓	✓	-	✓		✓	-	✓	✓	-	✓	✓	✓	✓		-	✓		✓	-	-	-	✓	✓	✓	-	-		-	-	✓	✓	-	-	-	-		-	-	✓
Forum 20	-	-	-	✓	✓	?	?	?		?	-	?	?	?	✓	?	?	?	-	?	✓		✓	-	?	?	?	?	✓	✓	✓		-	✓		-	-	-	✓	✓	?	-	?	?		-	✓	?	-	?	?	-	?		?	✓
Forum 21	-	-	-	-	-	-	-	-		-	-	✓	-	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	✓	✓		-	-		-	-	-	-	-	-	-	-	-		-	-	-	-	-	-	-	-		-	-	-
Forum 22	-	-	-	-	✓	-	✓	✓		✓	-	✓	✓	-	✓	-	✓	-	✓	✓	✓		✓	-	-	✓	-	✓	✓	✓	✓		-	-		-	-	-	-	✓	✓	✓	✓	-		-	-	✓	✓	✓	✓	-	✓		✓	✓
Forum 23	-	✓	-	-	✓	-	-	-		-	-	✓	✓	✓	✓	-	-	-	✓	✓	✓		✓	-	✓	✓	✓	✓	✓	✓	✓		-	✓		-	✓	-	-	✓	-	✓	✓	-		✓	-	-	✓	-	✓	-	✓		✓	✓
Successful markets	**46.15%**									33.33%					53.06%								47.89%										**61.11%**		37.31%								58.70%								76.47%	
Unsuccessful markets	30.77%									**52.08%**					46.67%								50.00%										**35.71%**		23.16%								46.05%								65.38%	

Table 7.3: Overview of the framework applied to the selected markets.

The upper part of the table contains the communities that have at least one clear hit, while the lower part contains the communities that have no clear hit. We use the following encoding: - if the feature does not exist in the respective community, ✓if the community exhibits it, and **?** if we were unable to find whether the feature exists or not in the respective community. We highlight the features excluded from the computed fraction of available features per group: in gray, the non-binary and constant dimensions; in yellow, the dimensions causing multi-collinearity.

Key Partners
F1. Restricted sections - pull-in
F2. Restricted sections - payment
F3. Restricted registration - pull-in
F4. Restricted registration - payment
F5. Scammers banned
F8. Seller status - verification
F9. Seller status - payment
F14. Seller status - recurrent fee

Key Resources
F18. Clear trade rules
F19. Moderators roles
F21. 2FA available
F22. CAPTCHA on authentication
F23. CAPTCHA on access

Key Activities
F5. Scammers banned
F8. Seller status - verification
F9. Seller status - payment
F10. User status - verification
F11. User status - payment
F20. Active moderation
F24. Escrow available

Value Proposition
F6. Working dispute resolution system
F7. Neutral mediator
F12. Reputation change on trade
F13. Reputation change by VIP
F22. CAPTCHA on authentication
F23. CAPTCHA on access
F24. Escrow available
F26. Public interaction
F27. Private interaction

Customer Relationships
F7. Neutral mediator
F25. Escrow recommended

Customer Segments
F1. Restricted sections - pull-in
F2. Restricted sections - payment
F3. Restricted registration - pull-in
F4. Restricted registration - payment
F5. Scammers banned
F10. User status - verification
F11. User status - payment
F15. User status - recurrent fee
F28. Pay to show content

Revenue Streams
F2. Restricted sections - payment
F4. Restricted registration - payment
F9. Seller status - payment
F11. User status - payment
F14. Seller status - recurrent fee
F15. User status - recurrent fee
F16. Escrow fee
F17. Sponsored ads

Cost Structure
F20. Active moderation
F24. Escrow available

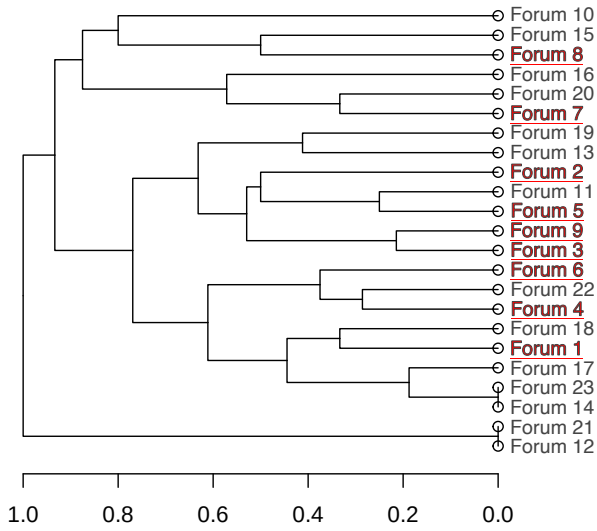


Figure 7.3: Hierarchical clustering of all communities based on their features. 'Successful' communities are reported in red, underlined.

7.5.3. Market similarity

After analyzing the features at a group level and individually, we further investigate the similarities among the communities. We perform a hierarchical clustering of the markets with complete linkage dissimilarity and compute the Jaccard distance on their feature set. The resulting dendrogram is depicted in Figure 7.3. The level at which forums are connected in the dendrogram indicates how similar the analyzed communities are when compared through their labeled features (i.e., the lower the level, the higher the similarity). Communities featuring at least a clear hit (i.e., 'successful') are reported in red. Features set to unknown for a given market (represented with **?**) are ignored for comparison.

We first notice that most successful markets are relatively close in the dendrogram, with two clusters in the lower part of the figure linking at a dissimilarity level of 0.6, except for `Forum 7` and `Forum 8`. Interestingly, according to some discussion in affiliated communities, we learn that `Forum 7` (a still-active, large marketplace) experienced its pinnacle of fame in the carding scene during the mid/late 2010s and now is in a descending phase. We could speculate that the classification of `Forum 7` as a successful market and its current descending phase could indicate that its characteristics may have been comparatively attractive for its time. However, the lack of evolution towards other mechanisms employed by competitors moved customers and sellers to other attractive alternatives, leading us to consider it more as an outlier. `Forum 8` features only one hit and seems a low-active market with little activity from the administrators. We contacted the administrators to obtain information on the benefits we would obtain by paying the 'activation fee'. After more than two months, we did not receive any answer yet; when we asked another user who activated their account, they confirmed our suspicions that the payment does not grant access to any additional market

section. That suggests that this could be considered more of an outlier, rather than a successful market. These considerations may justify the reason why these two markets seem less correlated to the other sub-tree.

Among the successful markets, `Forum 9` and `Forum 3` are the most similar; the two markets differ only in terms of access mechanisms: `Forum 9` vets registrations manually or via the payment of a fee, while `Forum 3` employs those mechanisms to limit access to a portion of the market. These two markets link at around 0.5 with `Forum 2` and `Forum 5`. Comparatively, `Forum 2` appears to be leaner on the regulatory side, with no explicit rules or recommendations on trade, the lack of a trade-based reputation system, although sporadically offers privileged statuses to vetted members. Both `Forum 2` and `Forum 5` look more concerned than `Forum 9` and `Forum 3` in vetting sellers; `Forum 5` offers a score associated with trade activity, and is more concerned about the revenue streams overall.

On the other hand, unsuccessful markets appear more spread out over the dendrogram and only connect at higher levels closer to 1. To further corroborate this observation, we compare the distances between all successful and unsuccessful markets. A Wilcoxon rank-sum test indicates that successful markets are on average more similar to each other than unsuccessful markets are ($p = 0.016$). That suggests that the feature configuration of successful markets has less variability (i.e., is more stable) than it has for unsuccessful markets in our set. Similarly, we find strong evidence that successful markets are more similar to each other than they are to unsuccessful markets ($p < 0.0001$); in other words, there seems to be a combination of features over which successful markets are characterized, but unsuccessful markets are not.

7.5.4. Qualitative observations on market features

We also observe some differences in the features between the two types of markets. From the framework, it emerges that the fraction of markets implementing stricter access control policies seems higher for successful markets ($F1_s$: 50%, $F4_s$: 33.33%) than for unsuccessful ones ($F1_u$: 21.43%, $F4_u$: 7.14%). An interesting relation emerges between the possibility to obtain privileged statuses for sellers and regular users: these mechanisms are in place for sellers more often in successful markets ($F8_s$: 33.33% $F9_s$: 85.71%) than in unsuccessful ones ($F8_u$: 16.67%, $F9_u$: 54.55%), while for users it is true the opposite ($F10_s$: 11.11%, $F11_s$: 42.86% vs $F10_u$: 15.38%, $F11_u$: 61.54%). Another notable aspect is that successful markets more often feature working dispute resolution systems ($F6_s$: 77.78%, $F6_u$: 57.14%), and rely more on advertisement as their source of revenue ($F17_s$: 100.00%; $F17_u$: 76.92%). Rather surprisingly, the presence of a reputation score based on trade seems to be a more recurrent characteristic for bad markets ($F12_s$: 33.33%; $F12_u$: 46.15%) [6]. Finally, marginal positive effects could be induced by the presence of clear trade rules ($F18_s$: 64.71%; $F18_u$: 45.83%) and the activity of moderators ($F20_s$: 77.78%, $F20_u$: 58.33%).

[6]This does not appear to be an isolated case however; Dupont et al. already reported that, despite the existence of a working reputation system in the infamous forum Darkode, trust remained elusive between members of the market, even in case of reputable members of the platform [81]

7.6. Discussion

Our analysis suggests that motivated cybercriminals may tend to choose markets with specific feature sets where they engage with the illicit trade of goods and that these criteria are rather stable across the forums they choose.

7.6.1. Impartiality in trade and seller verification

According to our findings in Section 7.5.2, it emerges that motivated cybercriminals prefer to conduct trade in marketplaces that have greater attention when selecting their *Key Partners* (Fisher test, $p = 0.032$). That implies that high-profile and motivated sellers are willing to undergo screening procedures and payment of fees to get access to segregated marketplaces to conduct trade that punishes dishonest traders. Another business aspect that seems to play a role in determining what markets are going to be the trade venues for motivated sellers is *Customer Relationships* (Fisher test, $p = 0.083$). Customer Relationships are mostly influenced by the impartiality of the administration when dealing with disputes. In other words, the administration is not directly involved in trade; when considering the opposite scenario, the administration may give greater visibility to their products at the expense of the competition. Whereas there is a process to earn the seller status on the platform, this could cause a conflict of interests in the administration, which could be tempted to refuse some applications to not lose their market position. Therefore, it appears that these characteristics lay the foundations for flawed marketplaces, where their operators lack the incentive to provide a fair and functional trading environment for other sellers, eventually resulting in markets not populated by prominent actors.

7.6.2. Revenue streams and admin incentives matter

The characteristics discussed in Section 7.5.4 seem to identify, on average, two different market categories which show some different traits. On the one hand, *successful markets* generally appear to be longstanding, with not extremely large communities, performing background checks on the sellers that want to operate on the market, and their administration is not involved with trade. *Unsuccessful markets*, on the other hand, appear to focus more on offering subscriptions to their members and their administration often offers products of various natures. That somewhat suggests that there are two types of markets, based on their mission and revenue strategy. On one hand, we find unsuccessful marketplaces where the administrators advertise their own products, focusing on earning as much as possible from their customers with the use of subscription services which give aesthetic perks and other miscellaneous privileges on the platform, but neglect care when selecting their partners (sellers) and keeping the entrance as unrestricted as possible. The high attractiveness of such markets and the lack of interest in prosecuting scammers makes those markets more similar to the markets for lemons; some of them eventually fail, while others still thrive. A potential explanation for their apparent success could be related to the high intake of new (inexperienced) members, and to the capability of the administrators of providing products of sufficient quality, even for free. Differently, administrators of successful markets aim at

obtaining value by providing a neutral, and competitive market, whose revenues rely on the payment of fees from sellers and promote fairness as an incentive to increase trade volume.

7.6.3. Smaller, less exposed markets tend to be more successful

The apparently irrelevant aspects identified in *Key Resources* presented in Section 7.5.2, like the presence of CAPTCHAs and the possibility of enabling 2FA, seem to play a negative role in the selection process. However, these could be interpreted as the effects of other causes; looking at Table 7.1, it emerges that the 9 forums featuring at least one hit (the first 9 forums of the table) have on average fewer posts and are on the underground scene for more time than their counterpart. That may indicate that the latter are more exposed and easily accessible markets (that is, well-known markets, often populated by script-kiddies) than more segregated and longstanding markets. Considering the target population of easily accessible forums (as they can be found easily by googling 'hacking forums'), these are more subjected to DDoS attacks and other abuses, rather than more segregated ones. That could be interpreted as a signal of less attractive markets for prominent cybercriminals.

7.6.4. Markets actively try to remove evidence of criminal activity

As noted in Section 7.5.1, we identified five markets, four of which successful, performing content removal. We noted that the post count reported in a profile of some usernames was higher than the actual number of posts available for that user; we verified across all the forums that it is not possible to delete your own content, suggesting then that this is an action performed from the market administrators. Furthermore, three markets performed a shadow-banning of some profiles on their market. For example, in `Forum 2`, we find that five of the nine removed profiles were associated with members of the Infraud Organization [73], a case connected to a large-scale carding organization; another two were connected to the Kostyukov case [72], where the 39 alleged cybercriminals were accused for being part of a racketeering scheme and counterfeit documents production; another user was the admin of the infamous marketplace Darkode, taken down from the police. That indicates that evidence of major criminal endeavors may disappear from related markets with time, suggesting that retrospective studies may systematically *not* include major criminal service or technological solutions in their analyses.

7.6.5. User expectations may signal 'virtuous' market forums

During our investigation, we noted a peculiar mechanism emerging from the discussion in the dispute resolution sections of two 'successful' markets, `Forum 3` and `Forum 9`. In July 2022, a user advertised on `Forum 3` a 1-click 0day trojan for Android and iOS for an undisclosed price of '8 digit price (USD)' (sic.). The alleged software is a stolen copy of the complete source code of governmental use spyware from a cyberwarfare company. Whereas not specified in the forum rules, several members pointed out that to execute deals of this proportion, depositing a conspicuous amount of money is necessary as proof of commitment; a longstanding member with high status in the community reports: " *You seller what kind of guarantees can you offer? You don't even have a 500k/1M deposit on forum (Which is required to sell products here). [...] I honestly believe that even the most honest escrow would*

be tempted once he sees 50.000.000$ in Crypto deposited in his wallet." The seller brings additional evidence supporting the idea that they own the product (e.g., screenshots of the manual, a quotation from the company with the list of features included, ...), but the community remained unconvinced. The same seller was banned from `Forum 9` because the seller and a potential buyer failed to converge on the execution modality of the transaction, raising the suspicion that the seller was acting in bad faith. In particular, the seller proposed that part of the funds would have been directly transferred to them, while the remaining part would have been transferred to an escrow service. This episode highlights how user expectations and community effects can make up for the lack of rules for specific edge cases. In turn, this may signal a virtuous market where these expectations can thrive, or at least develop.

7.6.6. Is this the full picture?

To discuss the features that support successful trade in markets, we rely on the principles of moral hazard and adverse selection, and consider how market features inform the business model of an underground community. From our analysis, we do not see a 'clear cut' distinction between the properties of successful and unsuccessful markets, and we find that, for example, less regulated markets can still thrive. We are aware that agency theory and the framing of underground marketplaces as businesses may still be insufficient to grasp the full extent of issues that characterize trade in mutually distrusted relations. Hence, we ask: *how can this picture be improved?* Several scientists from different disciplines have made efforts to study how markets can thrive from different angles; from a criminologist perspective, Soudijn and Zegers study the (private) communications among participants to a carding market obtained from a leak and discuss how the properties of a forum ease the execution of a criminal script [239]. In particular, they note the analogies of forums with the concept of 'offender convergence setting', first formulated by Felson [93]. With this concept, Felson describes locations where criminals converge for relaxation, exchange of information, and trade. Lusthaus analyzes the problem of trust in cybercriminal communities from a sociological point of view [170] and identifies three major problems: establishing cybercriminal identities, assessing cybercriminal attributes, and extra-legal governance. Lusthaus notes that marketplaces use several mechanisms to mitigate these problems and to keep scammers and police out of their communities. Gambetta and Lusthaus note that many of those mechanisms are based on the production of signals, which can be trusted especially when they are costly to produce [103, Chap.1], [170]. Finally, aspects related to the offline social ties of offenders could play a role in creating 'trending' markets or products traded therein. These phenomena, endogenous from our observation point, should be investigated more in detail. In particular, recalling what discussed in Section 1.1.3, this framework could be a sufficient starting point to further proceed in the research with a more human-centered approach. For example, our assumptions could be validated with the use of surveys and interviews with market participants, with particular attention to long-lasting members in said communities. In addition, questioning them regarding their choices, perceived risks, and motivation could shed light on additional factors playing a role in the success of a community.

The question remains of which additional theoretical angles and empirical measures should be accounted for to integrate (or re-invent) the current framework. We consider this part of future work and we lay the foundations for potential next steps in Chapter 8.

7.7. Limitations

Data representativeness

We cannot claim that our data is fully representative of existing underground market forums, or arrested cybercriminals. To build our ground truth, we relied on `arresttracker.com`, a website (now offline) that collects a list of cases related to cybercrime spanning from 1970 to 2021. The website is run on a voluntary basis and provides no warranties on the completeness and correctness of the collected and reported cybercriminal cases. The reported data comes from public statements from the USA Department of Justice, Europol press releases, or cybersecurity media outlets, and sources are reported for the majority of cases. Therefore, our study is limited only to criminals that have been observed by the USA authorities, including those that have been prosecuted during collaborations with international law enforcement agencies. Further, the inclusion of additional forums may provide different insights into the data. We share the dataset at `https://security1.win.tue.nl/doku.php?id=artefacts#data_sharing`.

Missing 'hits'

To minimize the risk of including false positives in our findings, we define a stringent criterion to correlate market usernames, activity, aliases, accusations, and the temporal correspondence between the criminal activity in the forum and the start of the prosecution. As a result, we discard little variations of usernames that could be related to the activity of copycats. However, this assumes that the aliases we collected are correct and accurate, and that their users have used them consistently across all forums. For example, in `Forum 4` we identified a user registered with the alias '† Voland †' and, despite having in our list the alias 'Voland', we refrained from including it in the analysis. Further, during our investigation, we noted that some accounts in some markets reported a message count greater than the actual messages that we could retrieve. That leads us to believe that said markets remove content from certain profiles as a consequence of an indictment to limit the risk of collateral damage to the community or upon account owners' request. Similarly, in some cases, we found evidence of accounts that have been renamed following an indictment. We witnessed two cases in `Forum 6` where these users wrote a 'farewell letter' to their customers, explaining the reasons that led them to end business, and warning them of the risks they may incur. To that extent, these profiles remain hard to identify and even manual investigation may be insufficient when the content of their advertisements has been redacted or removed entirely. Finally, we do not account for the actual existence of the forums in our selection before the time of arrest of each convicted cybercriminal. That may bias our results as it may not reflect the options available to each cybercriminal on which forums to trade on before they were convicted. However, as we consider convicted cybercriminals over a period within which all forums have been active (and that all forums but two have at least an 'inconclusive hit'), we consider any bias generated by this to be unlikely.

Markets evolution

In this study, we evaluate market features in accordance with the proposed framework. We inspected the markets over the period that spans from October 2022 and February 2023. That allows us to draw considerations on the *current state* of the markets in relation to the identified hits. However, this does not account for any evolution in the markets' settings, and indicted/convicted criminals could have based their choice on a different set of features. As such, the proposed framework should be considered as an evaluation instrument about the current state of markets to believably support the trade of cybercriminal technology. In fact, a market presenting the aspects of a 'successful' one does *not* necessarily imply that trade of effective criminal technology must happen, but rather that it could represent a viable alternative in case other marketplaces do not convincingly establish trust among participants and in the *bona fide* of their administration (anymore).

7.8. Related Work

7.8.1. Threat intelligence limitations

Currently, the cybercriminal landscape is dotted with communities serving the purpose of encountering like-minded people to exchange information on how to perform unlawful activities and trade the products of their misconduct. From a threat intelligence (TI) perspective, the possibility of joining these communities offers valuable insights into attacker behavior and emerging threats to produce early indicators of malicious activity. However, in the last few years, some studies have highlighted how TI suffers from coverage and accuracy problems [245, 167]. Bouwman et al. empirically evaluated the threat indicators overlap for 22 threat actors produced from two leading commercial TI vendors, finding marginal to no overlap among them [43]. A study reported how defenders, on average, rely on 7.7 TI providers to ensure coverage in their threat feeds [216], which is not always a viable solution due to the high fees those services impose. The adoption of TI feeds from multiple providers further exacerbates the problem of responding to real threats; the amount of Indicators of Compromise used in Intrusion Detection Systems produces an overwhelming volume of false positives, hindering the capability of performing timely actions in a Security Operation Center [233, Campobasso4, Campobasso6]. The problem with the volume of generated indicators also lies in the difficulty of extracting the most relevant signals from the threat landscape, rather than the most obvious ones [43]. That suggests that there could be a problem in the assessment of what are the sources of valuable threat intelligence and that a remarkable amount of noise could be collected in the process [129, 246].

7.8.2. Underground ecosystem characterization

By analogy, research faces similar problems too. There are multiple studies analyzing new markets, their population, and how they interact, but it remains unclear whether these findings capture the full picture of the underground economy [129, 246]. In fact, some studies suggest how questionable the quality of certain offered goods may be in relation to their price when no market regulatory mechanism is in place [122, 275]. Previous research studying

high-profile marketplaces for 0day exploits spotlighted how these communities implement different mechanisms to establish trust among peers [15]. In criminal settings, where interaction happens among mutually distrusted parties, several mechanisms must be in place to mitigate the insurgence of dishonest behavior, facilitated by the information asymmetry and ultimately leading to adverse selection (i.e., buyers can not distinguish between good and bad products) [122, 16, 239, 275]. Some studies highlighted how illicit marketplaces may limit the influx of new members in a community via the adoption registration fees, or vetting registrations via interviews, or using pull-in mechanisms where members of the community guarantee ('vouch') for the intentions of the new applicant [15, 107, 129]. As an effect, these restrictions severely limit the chances of security researchers infiltrating these communities [246], which face the ethical concerns of (economically) supporting cybercriminal ventures. In other scenarios, constructing a credible online identity to obtain clearance is deemed too (economically and timely) expensive to pursue [225]. As a result, some illicit marketplaces may remain out of reach. In research, studies on underground markets for which data leaks are available are not uncommon [274, 81, 188, 205, 7], or studies including specific markets for being 'famous' [107], 'popular' [36], or 'large' [208, 82, 170]. Notwithstanding the valuable contributions these studies made to our comprehension of cybercrime, it remains unclear if these selection mechanisms are shared by motivated cybercriminals too, and whether these assumptions introduce biases in our understanding of cybercrime at large. In fact, 'fame' could be attributed to a large user base or abundance of content, which in turn could be a proxy of their ease of access, indicate the presence of spam/low-quality content, or even how they are ranked on search engines. If this is the case, their selection as study subjects may be representative only of specific segments of the underground economy, while more segregated markets may remain unexplored. However, closed access is not the only aspect fostering the creation of flourishing markets fueled by high-quality products. Hence, understanding what are the features of successful markets is critical to grasping the criminal selection process when choosing the venue to conduct their criminal operations, helping researchers to make a more informed decision on the markets of choice for their research.

7.9. Chapter conclusion

In light of that, we acknowledge that this work cannot fully picture the complex phenomenons that characterize the cybercriminal activity of underground forums; indeed, the endogenous and exogenous phenomenons that impact the rise and fall of markets are too multifaceted to be captured from a single framework. Nonetheless, we hope that our contribution may help to improve our current understanding of the complexity of cybercrime, and will inspire further research from multiple disciplines. We believe that the answer to this complex question is yet to be found[7], and the key lies in the joint efforts from multiple disciplines studying the dynamics driving this fast-paced threatscape.

Acknowledgments

We thank our industrial partner Web-IQ for providing us with valuable historical information from underground forums.

[7] One may argue that the answer has already been found, but it is unclear to which *question* '`42`' is the answer [6, 270].

8

What next? Research perspectives on a fast-paced cybercriminal landscape

In the previous chapter, we discussed the characteristics of underground forum markets featuring 'successful' cybercriminals. Albeit our findings indicate that notorious/convicted criminals select the communities where to advertise their products based on the presence of trust- and accountability-enabling mechanisms, which can be considered agreeable and rational choices, our results cannot be considered representative of the cybercriminal forum ecosystem as a whole, and they do not necessarily reflect the decision-making of the majority of attackers because of the aggregate effects accounted by our observations. In fact, it is known that cyberattacks originate from a wide range of attackers optimizing for different goals [116]. Multiple studies analyze the offender's decision-making process, and frame attackers as rational players in a game theoretical sense, although with some limitations, optimizing different goals based on their heuristics and biases; some studies describe attacker choices as a function of their experiences and the effects that deterrence has on them [215], while others look at their decision model using an expected utility model [63], including potential biases that may affect the offender's evaluation of the real risks involved with the offense [255, 256]. Some studies describe attackers as 'work-averse', performing opportunistic attacks to maximize their returns with minimal efforts by reducing complexity and operational costs in exchange for low but safe returns [17, 61]. Other studies indicate that attackers perform their attacks in relation to the resources they have access to; Herley points out that scalable attacks reach orders of magnitude more users [121], and as such they represent a viable alternative for the majority of resource-bound attackers. However, there is literature that proposes prospect theory as an offender decision model, which indicates that attackers should not be considered perfectly rational, due to the interplay of biased perceptions, heuristics, and shallow planning before engaging in unlawful actions in their decisions [143]. Without taking sides on a single, specific view of the domain space of the offender decision-making model, the literature seems to agree on the existence of multiple cognitive aspects involved in the decision of attackers, and that attackers should not be considered necessarily

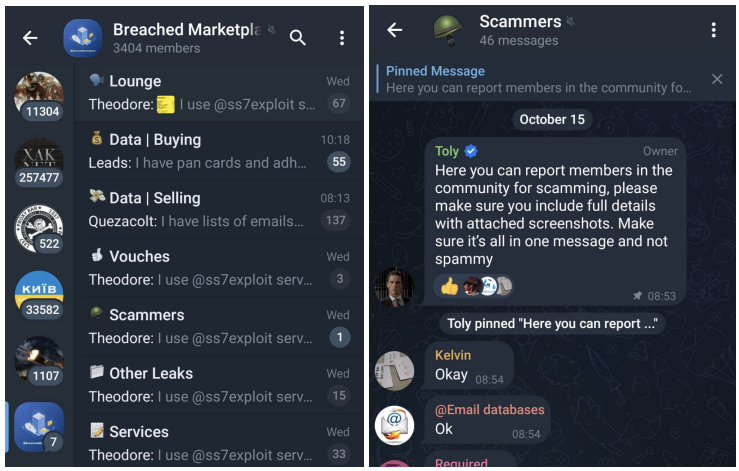


Figure 8.1: Telegram groups. To the left, a Telegram group featuring different topics ('subgroups', 'sections'). To the right, a topic dedicated to report scammers within the same Telegram group.

sophisticated [60]. Therefore, it seems that there is a significant portion of unsophisticated attackers able to thrive in sufficiently functional communities accessible to them, which enable them to mount damaging attacks at scale.

With regards to the different choices of attackers, during the research performed in the context of this thesis, we have assisted in multiple cases to the advertisements of criminal products on platforms different from forums, such as Telegram and ICQ, and there is increasing evidence that some criminal marketplaces are being hosted on Discord servers too [236]. However, it is currently unclear what is their role and relevance to the overall threatscape, despite the presence of anecdotal evidence that may indicate that this ecosystem could be sufficiently mature to support trade and cause harm, at least to some extent. For example, starting in November 2022, Telegram implemented the 'topics' functionality, which allows to create subgroups within a group, similar to forum sections (Figure 8.1, left); furthermore, some of them started to feature topics (or sections) within a group where users can report scammers, allegedly leading to an action from the administrators (Figure 8.1, right). Nonetheless, this information alone is far from indicating that this ecosystem can mitigate the foundational problem of quality uncertainty in trade; as such, these aspects deserve a proper investigation.

In light of these considerations, and from the findings on the attackers' preferences on underground communities drawn from the preliminary study presented in Chapter 7, it emerges that we lack a deeper understanding of the role of new technologies/platforms to conduct illicit business in the cybercriminal landscape, and what are the incentives for attackers towards their adoption. In general, literature studied the advent and establishment of forums [274, 122, 103, 16, 15, Campobasso1] from an economic point of view, and qualitative studies are based from the measurement of the aggregate effects of a diverse population of offenders [82, 81, 15, Campobasso2]. To explore the role of novel platforms in the cybercriminal ecosystem, we formulate the following future main research question:

FMRQ
What is the role of new platforms like Telegram and Discord in the overall threat landscape, and what are the traits of today's attackers populating these new venues?

The rest of this chapter proposes a number of future research questions as ways forward to investigate into this matter. In Section 8.1, we ask what are the characteristics of the population of today's attackers in terms of technical expertise, and how they relate to the emergence of new products and platforms. In Section 8.2, we argue that it would be interesting to investigate whether there are specific traits in the population of attackers joining messaging platforms like Telegram for cybercriminal purposes, while in Section 8.3 we consider whether the incentives in the selection of these new venues should be traced back to aspects dictated by (practical) convenience or nature of the traded product. Finally, we still consider the economic perspective as a meaningful view of attackers' decisions; therefore, in Section 8.4 we propose to investigate what are the characteristics of new venues in relation to the existing corpus of literature on trust and accountability establishment in self-regulated illicit marketplaces, with a focus on the mechanisms that mitigate quality uncertainty.

8.1. Wannabes or innovators?

In our work, we studied attacker preferences and choices from the observation of the aggregate effects of their decisions (e.g., in our case, the decision to purchase stolen browser profiles). Our observational standpoint does not allow us to discern different classes of threat actors populating these communities, which may act for different reasons and goals, thus complicating the task of isolating effects in our aggregate measurements. Even when considering a single marketplace, it may be populated by threat actors with different levels of expertise and goals in mind. Our study on Genesis Market presented in Chapter 6 offers an interpretation of attackers' preferences that compounds a potentially diverse population. What can we *really* say about its population? As discussed in Section 5.6.2, we obtained anecdotal evidence about attackers maximizing different goals, with potentially different skill sets: for example, some attackers prefer well-known and 'safe' scripts to commit fraud (e.g., cash-out bank accounts from a specific bank using a defined chain of service providers, ...), which do not necessarily require advanced technical capabilities except for some operational hygiene, while other may see Genesis as an IAB for obtaining a cheap foothold within an organization and stage more damaging attacks like ransomware, which in turn threat actors with more mature technical expertise. Research on the technical expertise of offenders is not new [133, 266, 165] but, to the best of our knowledge, there are no studies that systematically analyze the characteristics of attackers in relation to the communities they participate in, especially for relatively recent platforms like Telegram and Discord. In addition, one could consider that the rapid change of offered products (and the complexity connected with their creation and use) may affect the expertise distribution within the offender population, creating new opportunities for new attackers, while potentially solidifying the leading position of the more experienced ones, and changing the characteristics of cybercriminal organizations. There is evidence that the ransomware economy relies on organizations that include figures with different skills and seniority; for example, the Conti leak unveiled the existence of com-

plex business-like organization structures within the Conti ransomware group, where each member has a precise role [94, 111]. The organization includes strategic figures like human resources and carries out business processes that include quality control and market analysis, roles probably executed by experienced members; on the other end of the pyramid, the offensive teams are doing the actual heavy lifting, earning a fraction of the enterprise's revenues [186]. We hypothesize that the growing interest in Cybercrime-as-a-Service products (with the associated complexity in producing, maintaining, and retailing them) is creating new 'job opportunities' for a diverse set of attackers with specific skill sets that would not be able to take advantage of the proceeds of illicit activities otherwise (similar to offline figures like money mules [163]). These considerations, together with the apparent emergence of new platforms supporting cybercriminal activities, lead us to ask:

FRQ1: *What is the technical expertise of today's attackers, both as providers and consumers of criminal technology?*

A possible approach would be to survey market participants, including both sellers providing offensive technology and consumers, which could shed light on aspects that tie online with offline crime, offering a valuable connection with existing literature on traditional crime to validate or reject our hypotheses. Some of the studies in the field, among other insights, demonstrate that some attackers are willing to talk about their activities and, in some cases, even to share demographic information about them [131, 133]. Alternatively, other studies rely on recruiting participants among indicted criminals in collaboration with local law enforcement agencies [133, 266, 267]. Obtaining these answers from new platforms like Telegram, or from a wider corpus of forums than the one considered in Chapter 7, including forums from under-represented countries too, could provide a more robust understanding of this ecosystem, supporting the development of deterrence and educational measures in the fight against cybercrime.

8

8.2. A different target population for a different market?

On the same line drawn in the previous section, with the democratization of technology [269] and commodification of attack techniques, even less tech-savvy people have been granted access to information on topics regarding cybersecurity, penetration testing, and hacking, allowing more people to engage both in a cybersecurity and cybercriminal career. To that extent, the apparent adoption of novel messaging platforms like Telegram and Discord for the trade of illicit products may seem in contrast with the concerns on anonymity, trust, and accountability that traditional forum-based marketplaces over the dark web have (mostly) addressed. Therefore, one may speculate that participants of these new marketplaces could be less experienced and aware of the risks overall; nonetheless, this should be considered only as a factor contributing to the same effect, as there are already studies that look into social, demographic, and cultural traits of offenders [267]. To this extent, there are some excellent examples of studies that look into the historical reasons and social fabric within which cybercrime flourished in ex-Soviet countries [142], Nigeria [10], and Roma-

nia [172], and how these are bond with different levels of technical expertise. In light of this, we deem it plausible that these differences could play a role in the offenders' decision-making process for the adoption of novel messaging platforms for illicit purposes.

However, expertise and relation with risk may not be the only meaningful facets of this phenomenon, but it could also possibly be linked to the new ways today's offenders communicate. It is known that a fraction of cybercriminal activities like DDoS is often perpetrated by teenagers [108, 9, 185, 131]; hence, using apps and a communication style more akin to the target population, with captivating animated ads (as opposed to technical, purely textual jargon) could be a conscious choice that (not necessarily particularly skilled) threat actors take to cover a market segment with different tastes, giving an interesting insight on pathways into cybercrime from young offenders. We speculate that these new platforms could be an additional venue to reach new customers, but moving in parallel to the more traditional cybercriminal communities. In fact, our current understanding of the factors driving the trade of innovative offensive criminal products indicates that cybercriminals prefer regulated, forum-based marketplaces for trade, which offer means to establish trust and accountability among peers. However, at first glance, messaging platforms like Telegram do not seem to offer support to functionalities that help prove the identity and intentions of offenders and deter scammers in a meaningful way. This observation supports the intuition that there may be a number of aspects, different from those investigated in the forum-based criminal ecosystem, that offer a sense of trust in criminals and make them attractive, potentially enabling the trade of (at least some types of) cybercriminal products.

To further support our intuition, we recall that from Chapter 2 we observe that different (quality of) products exist in relation to the threat actors that offer and buy them. Chapter 7 sheds light on the characteristics of 'successful' underground communities that support trade, and shows that not all venues are populated by professional criminals capable of introducing innovation. On the contrary, it is very commonplace to identify marketplaces populated by script-kiddies, wannabes, or scammers. Considering this, we speculate that the cybercriminal population in novel messaging platforms may feature different social and cultural traits from those populating more traditional venues, in reaction to more accessible and appealing, new economic opportunities. Considering the changes in the available products, and the possibility that new venues attract attackers with different characteristics, we ask the following future research question:

FRQ2: *What are the demographics, and the social and cultural traits of threat actors populating new platforms like Telegram for conducting illicit activities?*

In relation to the features specific to Telegram, we conducted a preliminary exploration of alleged illicit marketplaces (including drugs, weapons, and illicit pornography) via the use of the 'Nearby People' feature in Telegram. From an initial assessment, it emerges that differences in the advertisements from groups localized in different areas of the globe exist, suggesting that at least some differences could be tied to the socio-cultural characteristics of the offenders: risk perception, law enforcement response, and features of offline crime. To investigate this matter, we propose to survey market participants about the reasons behind the choice to adopt such a platform, not limited to what is their awareness regarding the conducted activity, how concerned are they about the potential consequences, alongside their

demographics. Once again, conducting an investigation from this angle is crucial, as it could help establish bridges with existing literature from different domains and offer orthogonal views on the evolution of cybercrime.

8.3. Technical convenience or 'the right product'?

With the emergence of new and easily accessible messaging platforms that claim to offer privacy and support automation (e.g., Telegram offers well-documented APIs that allow the creation of bots and autoshops) completely for free (as opposed to the costs and complexity that come when running and maintaining a website that hosts illegal content) it is an increasingly convenient option to at least consider these services as additional venues for advertisement and sales. This convenience could offer a platform to the plethora of ever-increasing commodified offer of cybercriminal products; it may be the case that these platforms could allow offering a cheap service (on a subscription basis) via a messaging client with the use of autoshops/bots, allowing service providers to reach a larger market share, grant greater gains over time for a fraction of the infrastructure management cost, and requiring as little as an anonymous SIM card. On one hand, if all this sounds appealing and extremely convenient, currently it remains unclear to what extent marketplaces over such platforms mitigate the foundational problem of quality uncertainty in lack of authority. We argue that, if the apparent interest of attackers to populate platforms like Telegram is non-negligible and it would turn not to be fully explained from an ethnographic perspective, there should be characteristics of these marketplaces playing an incentive that currently we do not fully comprehend. To investigate this matter, we need two orthogonal views on the phenomenon: one that looks into the incentives of attackers to participate in these communities, and another aiming to produce a taxonomy of products offered therein. In light of this, we ask:

FRQ3.1: *What are the incentives that push attackers to adopt new platforms to conduct illicit business?*

FRQ3.2: *What products are traded in modern platforms like Telegram, and how do they differ from those traded in traditional cybercrime forums?*

To investigate FRQ3.1, we propose again a survey-based study covering both providers and consumers of offensive technologies traded in modern platforms, with a similar approach to the one proposed in Section 8.1. The survey's design could include questions on the respondent's ethnographic traits while deepening aspects that frame the motivation of an attacker to participate in such markets.

To address FRQ3.2, we should obtain access to groups, channels, and servers (from now onward, we refer to them as groups). Similarly to what was discussed in Chapter 7, it is necessary to select a number of groups. This could be performed opportunistically, using the group search features by keywords offered from messaging platforms like Telegram, or collecting this information from industry reports and collections offered online. Likewise, forums could be sampled with a similar strategy, using sources available online from threat intelligence firms and independent researchers as a starting point. From this first collection, it is possible to find more links to new groups via snowball sampling by parsing their con-

tent. During a preliminary analysis of the Telegram ecosystem, we encountered groups that required a (paid) invite to get access. As discussed in Section 7.3.4, in principle this does not pose an ethical concern, but we encourage researchers to discuss it with their relevant ERB. The obtained groups could be initially inspected manually to obtain an overview of how the content is organized and what is deemed relevant to answer the formulated research question. After this initial assessment, the content could be scraped using the offered APIs (e.g., Telegram, Discord) or with the use of tools like `THREAT/crawl` (Chapter 4). This information could be used to offer an overview of the current ecosystem based on those platforms, and it could be compared with the large corpus of existing literature on forum-based marketplaces to gain more insights into their potential differences.

8.4. New venues, same problems?

Finally, we believe that the approach used in Chapter 7 and the relation between product sophistication and platform maturity may still hold even for the domain of new platforms. We argue that irrespective of what are the characteristics of the threat actors populating modern platforms and the offered products thereof, the foundational problem of quality uncertainty remains. Therefore, if the activities within markets were to pose a believable threat, the marketplaces themselves should be sufficiently mature to foster trade of noteworthy offensive cyber capabilities. Therefore, we ask:

FRQ4: *How do market participants tell apart good products from bad products in modern platforms?*

To this extent, before devising a rigorous research plan, it is necessary first to obtain an overview of the different alleged marketplaces in these new platforms and to obtain initial evidence of trade. We believe that a preliminary 'in-field' investigation is necessary to mitigate the risk of conducting a biased analysis of trade mechanics that could be remarkably different from those of more traditional cybercriminal venues. Following this assessment, a mixed approach using data collection to identify advertisements and in-group follow-up conversations and evidence of trade, and the use of surveys as proposed in the previous sections could offer valuable insights into the decision-making process of these attackers. We suspect that a correlation may exist between the products offered, the expertise of threat actors, and the (in)capability of such markets to convincingly address the problems of trust and accountability. We hypothesize that messaging platforms could indeed be capable of supporting the trade of *some* unsophisticated products (e.g., database leaks, stolen credit cards, documents, and credentials, ...) or more complex ones offered on a subscription basis (e.g., commodity malware rented with a per-customer build to obfuscate it, ...), and these products could be of interest to some *specific* market segments.

In conclusion, we believe that answering these questions could greatly help to determine whether new platforms *today* represent a mature source of threats 'worth investigating', or if they only represent additional noise that could otherwise be discarded.

8.5. Chapter conclusion

In conclusion, the pace at which the cybercriminal ecosystem evolves still offers many research opportunities. Our current understanding of the foundational problems that forum-based underground markets address makes it unlikely that their demise is in sight. We suspect that, considering that new platforms could offer conditions facilitating trade (and the sheer volume of interactions seems to support this idea), a coexistence of the two ecosystems is plausible, where a part of the economy finding forums' population or characteristics unsuitable for their purposes could find fertile soil in the new ecosystem. We believe that times are ripe for a rigorous investigation, and we hope and encourage brilliant scholars to take this challenge to help better understand the evolution of cybercrime and defend our societies.

9

Conclusion

In this final chapter, we summarize the results of our works and how they contribute to answering the research questions presented in Chapter 1. From our work, we draw a number of conclusions that help supporting research in network monitoring-hostile environments like segregated underground communities, we make a step forward in the understanding of *what* are the attacker preferences when it comes to target selection, and study the characteristics of underground marketplaces that convincingly address foundational problems of trade to promote commerce and foster innovation.

9.1. Summary of Contributions

9.1.1. Part I: the underground ecosystem and how to measure it

Criminals proliferate in online communities, often in the shape of forums. In forums, cybercriminals meet like-minded people, share information and trade products. However, at first glance, it appears that these communities are diverse and might play a different role in the overall picture. Hence, our first research question is:

RQ1: *How can we preliminarily characterize the space of underground marketplaces supporting the provisioning of offensive cyber capabilities?*

In Chapter 2, we conduct a preliminary investigation of the characteristics of marketplaces in relation with their provisioning of offensive cyber capabilities (OCC) supporting the Access-as-a-Service (AaaS) threat model, and group the OCC in 5 pillars that identify a specific stage of an offensive operation. From our investigation based on first-hand observation, literature review, technical and governmental reports it emerges that the ability of a market to provision OCC supporting AaaS operations largely varies in relation with the maturity of the market and the required level of clearance from the customer. We identify two major families of OCO providers in the cyberspace: self- and semi-regulated marketplaces.

In the self-regulated space of black markets, we broadly identify three categories of markets, based on how restricted their access is: free access, pull-in, and segregated, while in the semi-

regulated space we identify companies specialized in the production of espionnage tools and governments with large funds capable of researching and producing offensive tools in-house. From our overview, we identify a positive correlation between product quality and market segregation. In the self-regulated space, free access marketplaces are mostly or completely unable to support any of the pillars of AaaS operations, while segregated markets can even meet quality standards of the private sector in some cases. The semi-regulated space instead provides resourceful, often state-sponsored threat actors with state-of-the-art tools, adopted to conduct targeted espionage operations covering a spectrum of goals ranging from counter-terrorism to human rights activists haunting.

Our findings allowed us to answer to **RQ1**, which indicate that there are measurable differences among marketplaces of different maturity and that their segregation level is positively linked to their ability to offer effective offensive cyber capabilities. Therefore, this gives us the motivation to further study those marketplaces capable of offering advanced offensive technology, as it could potentially allow to monitor a handful marketplaces whose operations account for the majority of the real world threats. However, we empirically find that these differences are not only limited to their access policy and product provisioning; these marketplaces tend to perform network monitoring activities to thwart crawlers, as they are seen as unwanted investigation activities from law enforcement and researchers. Therefore, the second research question that emerges is:

RQ2: *How can we stealthily extract information on market activity from underground forum markets while circumventing crawler detection mechanisms, and scaling up monitoring to multiple forums?*

In Chapter 3, we develop `CARONTE`, a stealthy crawler that models human behavior to extract data from high-profile communities while staying under the radar. We compare the performances of `CARONTE` against state-of-the-art tools and human participants, and the results indicate that `CARONTE` shows similar browser activity with humans. These results indicate that stealth crawling is possible with the use of human behavior modeling, while reproducing a network fingerprint indistinguishable from the one of a regular user using a fully-fledged browser to navigate. To scale up crawling operations, in the same chapter we experimented with a guided procedure that allows human operators of the software to create tailored instances of the crawler to scrape the content of a specific underground forum with positive results. The overall results proved to be encouraging, and in Chapter 4 we attempt to address in more rigorously the (new) challenges of extracting data from monitored underground communities, while providing the scientific community with a tool that allows to perform research in highly-monitored underground communities. We test the resulting tool, `THREAT/crawl` against seven live cybercriminal underground forums and obtain partially positive results. Our benchmark shows that the proposed tool successfully handles the training in multiple scenarios, including a number of edge cases; however, the diversity of the considered forum sample shows some aspects in terms of robustness that require further improvement before `THREAT/crawl` could be considered a completely reliable solution, and we discuss how some of these improvements could be achieved. Despite these limitations, to the best of our knowledge, `THREAT/crawl` represents one of the most advanced attempts to offer a flexible and reusable software to support academic research on highly-guarded cybercriminal communities. Findings in these two chapters, and the devel-

oped tools allowed us to answer to **RQ2**, identifying novel strategies to operate a crawler in disguise while ensuring the reusability of the implemented solutions.

9.1.2. Part II: Investigation and evaluation of a prominent, emerging threat from underground markets

At this stage of our research, we gathered evidence suggesting that differences among marketplaces and the related provisioning of offensive capabilities exist (Chapter 2), and that extracting data from more segregated and guarded communities is possible with the adoption of stealthy tools in line with the findings reported in Chapters 3 and 4. Therefore, the research question now is:

RQ3: *How can these monitoring capabilities and market characteristics be used to identify and evaluate high-relevance cyber-threats?*

Following the considerations made in Chapter 2, we explored marketplaces with different levels of segregation, applying for memberships and pretending to be interested commercial partners to gain access to their closed access areas. During our investigation, we discovered the existence of Genesis Market, an underground marketplace offering an innovative (at the time) service for user impersonation at scale. The criminal community's momentum around this product, the invite-only access model, and the thorough description of the capabilities of the offered service made it an interesting subject of study as a potential innovative threat. Once our application to access Genesis Market was approved, a preliminary inspection of the platform further corroborated the intuition that Genesis represented a stark example of cybercriminal innovation.

To test in a real-case scenario the tools presented in Chapters 3 and 4, in Chapter 5 we design and develop crawlers in accordance to the findings from the aforementioned chapters to extract data about Genesis Market's offer. From the established foothold into Genesis Market and extracted data, we obtain a clear picture of the market operations and derive the new threat model it operates on, named Impersonation-as-a-Service (IMPaaS). We analyze the scraped data to derive the market pricing model, which results stable and accounts for the expected value of the offered profiles. These findings, together with the professional outlook, detailed documentation, and continuous flow of new victim profiles into the platform lead us to conclude that IMPaaS is a mature threat model.

In Chapter 6, we further deepen our studies on Genesis Market to characterize the impact of the IMPaaS model on the real world. With the use of a large crawling infrastructure using multiple accounts, we extensively collected data on Genesis' offer and demand throughout 161 days of its activity. We devise a robust methodology to account for the missing data in our dataset and obtain reliable figures of the market activity. Our findings indicate that Genesis customers (i.e., attackers) show nuanced preferences, although they present a preference for profiles originating from Northern America and Oceania. Furthermore, our study shows that the profiles offered within the last 24 hours have higher chances of being sold than those lasting longer, suggesting that threat actors prefer newly attacked victims. During this period, almost 100k new victims have been affected and 20-28k profiles have been

sold, corresponding to as many potential attacked users, for a total revenue for Genesis of 540-720k USD. These rare insights offer a clearer understanding of the attacker's motivations. We employ these findings to extend a cyber risk model, and we highlight how it can be used to inform organizations and the public sector about the risks that IMPaaS poses to them.

The efforts condensed in these two chapters provide a solid positive answer to **RQ3**: indeed, Genesis Market was advertised only on two of the 30+ marketplaces that we had access to, and they share characteristics of segregation and member verification similar to those described in the context of pull-in marketplaces in Chapter 2. Furthermore, the data extraction from Genesis Market was only possible with the use of stealthy crawlers; during the beginning of our investigation in the market, we partially implemented a `THREAT/crawl`-like solution to scrape data neglecting the user-behavior modelling, resulting in our account banned within 48 hours from the start of our data collection.

In conclusion, the study of Genesis Market indicates that innovative threats can emerge from the underground criminal panorama, capable of causing damage to society at large, and it shows that it is possible to study them for a timely preparation of defenses. We argue that being able to predict 'where' the next innovative products are likely to emerge could greatly help both the public and private sectors to protect our societies at large.

9.1.3. Part III: Characterizing the underground markets that 'matter' and research perspectives

Despite having obtained some indications from the answer to **RQ1**, it is still unclear which factors foster criminal innovation and guide experienced threat actors to choose one venue rather than another to promote products like those from Genesis Market. With that considered, we ask:

RQ4: *What characteristics differentiate underground forum markets capable of supporting high-relevance cyber-threats from those that cannot, and how can this difference be evaluated through market observation?*

To answer to that question, in Chapter 7 we identify a number of problems affecting trade in unregulated environments. We consider moral hazard, adverse selection, cost structure, and (legal) risks as the primary problems affecting the success of underground marketplaces. We consider these problems as a compass to identify market features that attempt to mitigate them, and condense these features into an evaluation framework. Subsequently, we employ our framework to evaluate 23 underground forums; among these, 9 forums feature criminals who have been indicted or convicted from the US Department of Justice between 2011 and 2021. From the comparison of their characteristics, it emerges that ('successful') forums featuring prosecuted ('successful') criminals, on average, tend to scrutinize sellers, promote impartiality in trade by not being directly involved with it (thus administrators not abusing of their privileged position), have clear revenue streams to ensure sustainability, and in general are more segregated, requiring registration fees or scrutinizing even regular customers. To that extent, our findings suggest that some marketplaces implement a number of strategies to establish trust and accountability among the members of their platform; by mea-

suring their presence or absence one could preliminarily estimate whether a marketplace's activity could pose a significant threat or not, helping law enforcement and researchers to direct their efforts to investigate the marketplaces that 'matter'. Albeit the results could not be considered conclusive due to the size of the sampled forums and to the likely presence of additional, unexplored factors (e.g., demographic on market's participants could indicate global differences in the prevalence of one type of attack rather than another, some products – and customers – appear more frequently in marketplaces with different characteristics, noise introduced by aggregate measurements, ...), we argue that our findings provide an answer to **RQ4**, but the matter should be further investigated.

For this reason, in Chapter 8, we retrospectively evaluate the results from Chapter 7 to discuss the possible ways forward in the research. We argue that the criminal ecosystem is in rapid evolution, and this should be reflected in the characteristics of the attackers that populate it. At the root of this statement there is our strong belief that rigorously measuring the 'symptoms' of cybercrime is an effective tool for estimation and inference, but it fails to grasp some subtleties. Furthermore, we observe the growing adoption of Telegram and other instant messaging apps as venues to create new marketplaces; however, our current understanding of the dynamics within them is marginal. We speculate that this ecosystem could be mature enough to support the trade of *some* technology and that market participants may show different traits from those from more traditional forum-based marketplaces. We propose possible ways forward to their analysis, especially in relation to our current understanding of the traditional forum-based ecosystem.

9.2. Answering the Main Research Question

Our research on the illicit marketplaces helps up to better understand what are the characteristics of platforms that foster innovation and pose a believable threat to the real world. The results of our studies combined enable us to answer the main research question that we formulated at the beginning of this dissertation:

MRQ
How can we identify which cybercriminal marketplaces can support the trade of innovative offensive products and services, and how can we effectively monitor their activity to evaluate the threat they pose?

In the context of this study, we obtained evidence that not all marketplaces are made equal, and the sophistication of the offensive capabilities they can provide to a large spectrum of threat actors is linked to their maturity in ensuring a trustworthy trade environment. Threat actors rely on service providers to craft offensive criminal operations targeting large segments of the population with scalable and unsophisticated attacks with the ultimate goal of pursuing financial gains, but also critical infrastructure and prominent political figures to achieve strategic and political goals.

To learn more on their characteristics, we employ a bottom-up approach, starting from the stealth monitoring of underground marketplaces to gather data, and analyzing it to esti-

mate their threat levels. We propose and successfully validate methods and tools to covertly extract information from highly-guarded underground communities, and investigate their ability to support the trade of effective products in contrast with those who cannot. The gathered information allows us to provide an initial evaluation framework to identify 'successful' marketplaces that can pose a significant threat to the real world. We argue that by pursuing the path of characterizing 'successful' marketplaces, it would be possible to produce knowledge that could inform policy makers, researchers and law enforcement. To support this statement, we studied Genesis Market, an (at the time) emerging threat advertised in pull-in marketplaces that commodifies and weaponizes stolen user profiles to conduct impersonation attacks at scale. With the use of our data extraction tools, we perform a large data collection, characterize the underlying threat model that we called Impersonation-as-a-Service, and conclude that Genesis Market represent(ed) a mature threat globally. Furthermore, we study the market's activity to draw conclusions on the attackers' (customers') preferences, identifying the features playing a positive role in the victimization process and condensing our results in an extended cyber-risk model we instantiated for IMPaaS.

These results indicate that, although the process is fraught with context-dependant considerations that shape diverse data analysis strategies, it is possible to monitor underground markets to extract valuable threat intelligence that could be used to study threats before they become mature enough to pose a threat at scale to our society.

9.3. Final words

This work wants to be an encouragement to scholars, industry professionals, and law enforcement officials to pursue the fight against cybercrime with a particular attention to the characteristics of cybercriminal venues. During our work, we witnessed how virtuous underground communities tackle a number of problems to remain functional, and how some of them foster innovation and understand market needs. Hence, being able to select and monitor those 'that matter' could greatly help researchers and law enforcement to focus on the most damaging venues of the underworld, enhancing our comprehension of this phenomenon and helping to allocate resources more effectively to counteract them. Furthermore, innovation is not only a process that involves the product, but it includes the characteristics of criminal venues too. Historically, the first cybercriminal markets were hosted on IRC chatrooms, but eventually they moved to forums, and today's often benefit from the additional anonymity introduced by TOR to keep their operations running. We believe that cybercrime has a strong social connotation, where participants shape their functionality and aspect. Considering the rapid evolution of our societies caused by the pervasiveness of technology, we suppose that new platform and ways to communicate could be on the rise, and the time to study them as sources of significant threat may be ripe.

Appendices

Chapter 3 appendix

The proposed method: CARONTE

From the literature analysis in the previous section, we derive a set of characteristics for our method for stealthy crawling. We use these characteristics as desiderata to define the architecture for CARONTE. In this appendix, we provide a more detailed overview of the technical architecture of CARONTE.

Method definition

Functional and behavioral characteristics. First and foremost, a hypothetical tool must have the ability to diverge from crawler behavior and, where possible, to mimic human behavior. In this regard, as emerged from the time patterns paragraph in Section 3.2.3, keeping in mind that one significant aspect of crawlers is their greed for resources, the tool should not exhibit high fetch rates and mimic as much as possible human time to browse and read resources, whether the content is appealing for it or not. Therefore, the crawler must be able to mimic interest in a specified subset of the forum, exploring only certain sections of it, according to the hypothetical goals of our modeled actor. To achieve this, the tool must be able to semi-automatically learn forum structures without the need for extensive pre-collected datasets on which to train automated models [217]. This should be a one-time-only process, employed for each new forum structure that has not already been learned. In addition to this, a tool should be able to receive instructions about which areas are valuable to crawl and which to skip in terms of sections and topics. Thanks to the acquired knowledge, the crawler will be able to explore forum content through *navigational* and *informational* queries; in particular, it will access quickly resources, like posts in threads already read and the resources related to path traversing that occur from the landing page to the section of interest, while it will take more time and produce less frequent clicks while staying on pages with new content from the section of interest. To improve its stealthiness on this aspect, we design a navigation schedule on a forum like an actual human being having in mind variables such as time of the day and stochastic interruptions.

Technical characteristics. To avoid detection at the network level, the hypothetical tool will have to act indistinguishably from a regular browser in terms of generated traffic and differ from regular crawlers. The primary aspect is to produce not suspicious HTTP requests against the webserver; crawlers' traffic is characterized by the adoption of HEAD HTTP requests to determine whether the resource to download is of interest or not, non-filling of

Table A3.1: Summary of identified characteristics for CARONTE

Characteristic	Description	Implementation
Learning forum structures	Understanding forum structure, how to browse it and where valuable information is	Creation of a human supervised learning module that identifies needed resources
Regular browser behavior	Realistic user agents, caching behavior, referral handling	Exploration of required sections only, throttling requests accordingly to text volume of the page, mimicking reading time. Confining crawling activity in semi-random time slots during the day and suspending it for random amounts of time during the day
Realistic browser configuration	Addons install and pages download feature	Install NoScript and Page Save WE, preparation of the browser to support shortcuts for downloading a page
Anonymity	Browsing session needs to be anonymous	TOR Browser adoption, JavaScript disabled at browser level and changing default to refuse JavaScript and active content

referral link field in requests, and by the usage of a bogus user agent [224, 79, 240, 139, 28]. Also, depending on the goal of the crawler, they might be interested in scraping content without rendering and providing the opportunity to browse it, thus missing the support for a proper browser engine that will allow them to consistently handle cache, manage cookies, and execute JavaScript. Further, crawlers might be interested in fetching only text content, refusing to download styles, images [240] and JavaScript (e.g., to minimize network footprint), or will not actively execute client-side code such as JavaScript, handle sessions and cache as a 'regular' web browser would do. In our study case, we assume it to be legitimate to have JavaScript completely disabled to increase the anonymity of our tool; this countermeasure inside the Dark Web is quite common and should not raise any suspicion.

With this in mind, we opt for a fully functional browser that by design covers all these aspects coherently with a legitimate one, while offering the possibility to be maneuvered programmatically. Table A3.1 provides an overview of the identified characteristics of the tool.

Proof-of-concept architecture and implementation

We design and implement CARONTE a proof-of-concept to benchmark the proposed method. CARONTE adopts a two-tier architecture for the *training* and *crawling* operations.

Base mechanism. The trainer module has the task of building a knowledge base for traversing the forum structure (Figure A3.1). For each page where relevant content or fields are present, the trainer will load, save, and render a modified copy of it to the user. For each of them, the operator will be asked to click on the desired resources inside of the rendered page. Before being rendered, pages are preprocessed; in particular, we inject JavaScript scripts to

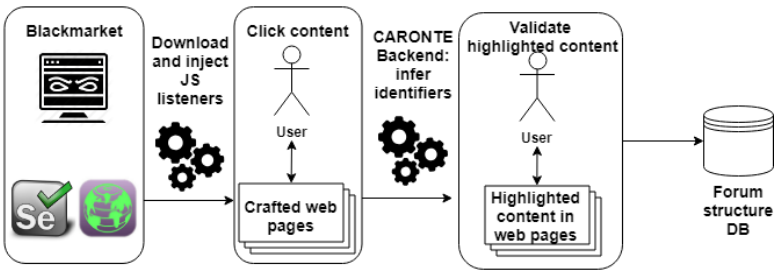


Figure A3.1: CARONTE trainer module structure.

allow CARONTE to gather the events triggered by the human operator. With different combinations of `onclick()` and `addEventListener()`, we control these interactions and generate AJAX requests against CARONTE's backend. The payload of these requests is a resource identifier (see 'Resource identifiers' paragraph in this section) that will allow the crawler module to access to the required information or interact with it, where necessary. Subsequently, it then proceeds to render again the saved page, but highlighting the previously identified content, allowing the user to confirm if the identifiers for the resources have been inferred correctly by the tool or not (Figure A3.2). In some cases, user-generated clicks are not possible or we aim to identify a group of resources. For example, this is the case for identifying multiple posts inside of a thread; for this kind of resource, our goal is to infer a resource identifier that can operate like a regular expression, enabling the tool to resolve all the required elements on the page. Our strategy here is based on the collection of multiple snippets of text contained in each of these resources (Figure A3.3). For each of the received fragments, CARONTE will query the browser's JavaScript engine via Selenium in order to resolve their identifiers and, through syntactical similarity, generate a matching one. Text content will be gathered with the help of the human operator on a special page (here referred to as *content collector page*) that is presented to the user together with the original page.

Resource identifiers. The desired resources can be identified through two different approaches: **XPath** or **HTML common classes**. XPath is a standardized query language that identifies elements inside of an XML-like document; it supports regular expressions for matching several elements. HTML classes instead are attributes assigned to nodes of an HTML file for which different styles are assigned. Even though XPath is an *ad-hoc* technique for identifying elements in an HTML page, sometimes inferring HTML classes is easier than XPaths. During the training phase, when a resource is clicked, the loaded page will identify the associated identifier through a series of heuristics, and send it to the backend. If the resources are multiple, the *content-collector page* will be rendered with the downloaded page and the user will fill the fields with the required data. The process to identify the most likely resource identifiers depends on the data structure and the number of classes associated with that resource. CARONTE supports the following four cases:

- **Technique 1**. Extract the XPath of the exact resource. If the resources are multiple, the most frequent XPath will be the candidate;

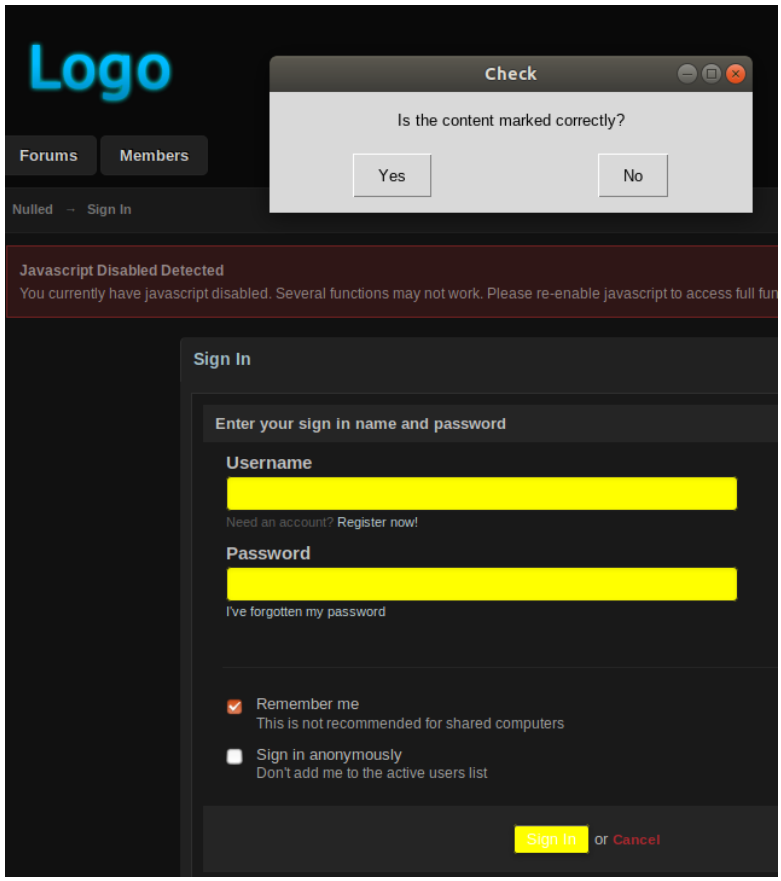


Figure A3.2: Validation of identifiers inferred.

- **Technique 2**. Extract the XPath of the exact resource, but the last node is truncated. The XPath approach may fail due to the presence of extra HTML tags (e.g., due to text formatting), that can then be disregarded. If the resources are multiple, the most frequent XPath will be the candidate after removing the last node;

- **Technique 3**. The class of the exact resource. If the resources are multiple, the most frequent class will be the candidate. This approach solves the problem of calculating an XPath on a page where the content is dynamic, resulting in a non-predictable XPath for a certain resource, depending on the loaded content on the page. If the resources are not assigned to a class, the element will be replaced with its parent, which will act as a wrapper;

- **Technique 4**. Two classes of the exact resource. If the resources are multiple, the two most frequent classes will be the candidates. This approach is adopted to handle elements on a page that exhibit the same class of the desired content, resulting in a misclassification. Therefore, this strategy allows to have a stricter condition on the

search criteria for the required resource. If the resource(s) has no class, the element will be replaced with its parent, which will act as a wrapper.

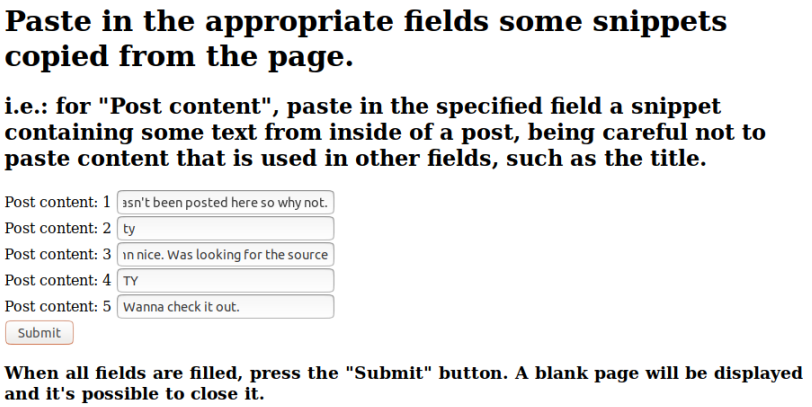


Figure A3.3: Gathering of text snippets from the saved page (in next tab).

Crawler module. Based on the structural details collected with the trainer module, the crawler module will traverse the forum to reach the required resources, explore threads, and collect all the required data. The crawler will generate traffic from a regular browser while camouflaging its nature adopting low fetch rates for pages. How time is calculated before accessing the next resource is deepened in Section 9.3.

CARONTE further keeps track of updated threads and selects those opportunistically for visiting. Threads that have not been updated are not traversed a second time.

Behavioral aspects

Legitimate browser traffic - Browser. To implement CARONTE's browser functionalities we adopt **Tor Browser Selenium**, or *tbselenium* for short. *tbselenium* accesses *geckodriver*, the browser engine branded Mozilla that allows to maneuver the browser's behavior and UI. Moreover, *tbselenium* exposes an interface for customizing the environment and, finally, produces traffic identical to Tor Browser.

Mimicking legitimate human traffic. *Work schedule.* CARONTE can be configured to work within pre-defined timeslots during the week or on the weekends, late afternoons and evenings during the week, and all three sessions on weekends. Between each session, a randomized time of inactivity simulates short pauses (between 5 minutes and half an hour) and longer ones at pre-defined times (e.g. 2 hours around dinner time). These can be configured. Each session has a start time and an end time; each of them can vary up to 25% of the total duration of the crawling session randomly. Each session has a 20% of chance being skipped. Nonetheless, we would avoid having 24 hours of inactivity, so if there are no sessions scheduled in the next 24 hours, a compatible one with the default schedule will be executed. Start

and end times are shifted according to the timezone of the geographical location of our forum user profile.

Reading time. The time spent between two requests is calculated using two main criteria:

- If the current page does not show significant content to be read (e.g. pressing the login button, reaching the section of interest of a forum, moving to page 2 of a forum section, ...) or the content has been already read (a thread may contain new messages, therefore old will be skipped), the time spent before going to the next page is a random number of seconds between 3 and 7. This decision is based on the fact that the information on the page is more essential and visual. This enables our fake actor to skim rapidly and choose what to read, resulting in fulfilling the expectation of having a *navigational queries* pattern;
- If the current page is the body of a thread, the tool will wait, for each unread post, an arbitrary amount of seconds calculated as the time to read the post at a speed in the range of 120-180 WPM. This behavior validates the expectation of producing *informational* queries.

User event generation. CARONTE's modeled user goal is to reach the threads of interest and iterate them to extract their content. When starting the crawling process, CARONTE loads the forum homepage, as it was typed on the address bar, then reaches the login page. Once logged in, it reaches one of the sections of interest expressed during the training and opens a thread at a time (if it has been never read or has new replies). For each thread, it browses each page until the thread has been read in the whole. The click patterns generated match the purpose of our fake user, which considers relevant the content of pages with a significant quantity of text like a thread instead of a login page.

Chapter 5 appendix

Market Features

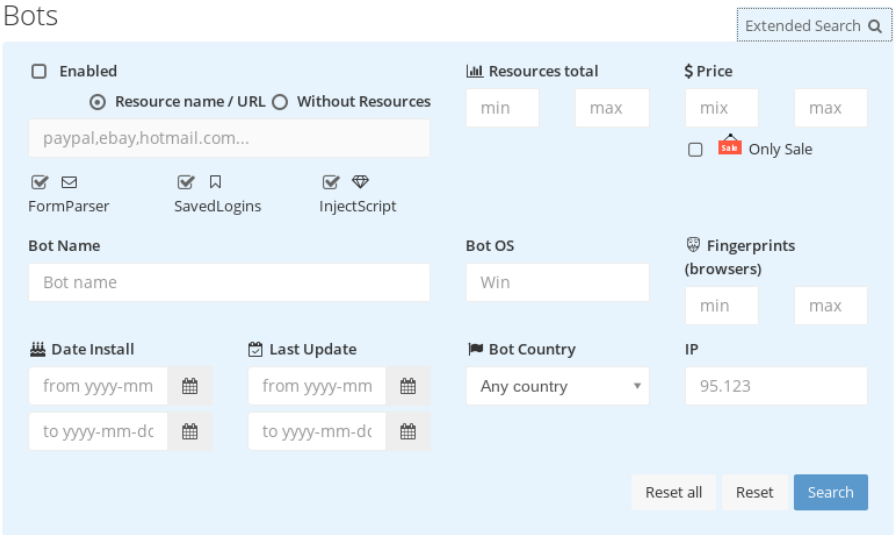


Figure A5.1: View on the advanced search functionality.

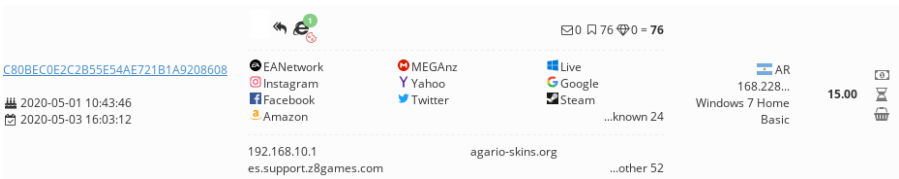


Figure A5.2: Overview of a listed profile on Genesis Market.

Figure A5.1 and Figure A5.2 report screenshots from the market. Figure A5.1 depicts the search function of the market. Attackers can access a fine-grained research tool that enables them to search for profiles with specific resource composition, number of available browser fingerprints, and other information. In Figure A5.2, an overview of the details for each profile is provided. On the left, from top to bottom, the name of the profile and the installation and update dates of the profile are listed. On top, in the center, is the list of the available browsers (here only Microsoft Edge). The superimposed number indicates the number of fingerprints available for that specific browser and the superimposed icon whether cookies are available (green) or not (red). On the top-right is reported the number of resources available (here 76). In the center, an overview of the websites for which resources are present. On the right, details about the country, IP prefix, and operating system are provided. Finally, to the right-most, there is the price expressed in USD and buttons to respectively buy, reserve, or add the profile to the cart.

Table A5.1: Logistic regression for discounted profiles

$y = sale$	Model
β_0	-0.09^{*}
	(0.05)
`Resources`	0.00^{***}
	(0.00)
`Year 2020`	-0.92^{***}
	(0.04)
`Cookies`	0.00
	(0.00)
`Browsers`	0.14^{***}
	(0.02)
R^2	0.04
Num. obs.	11683

$^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05$

Table A5.2: Autocorrelation matrix among categories of resources available for each bot.

	Crypto	Social	Services	Other	MoneyTransfer
Crypto					
Social	0.08				
Services	0.04	0.09			
Other	0.05	0.11	0.05		
MoneyTransfer	0.16	0.26	0.10	0.17	
Commerce	0.12	0.28	0.09	0.17	0.42

Further data insights

Table A5.1 shows correlation coefficients for the logistic model $sale = \beta_0 + \beta_1 Resources + \beta_2 Year + \beta_3 Cookies + \beta_4 Browsers$ (binary response variable).

Whereas Resources, Year, and Browsers are significant predictors, the effect is very small with the unsurprising exception of Year, suggesting that profiles recently acquired are *less* likely to be put on 'sale'. The coefficient for β_4 shows a positive, albeit small, effect of the number of browsers provided in a profile on the likelihood of profile sale. As indicated by the small R^2, we do not find a clear rationale explaining this effect.

Table A5.2 reports correlation coefficients between `Resources` types in our dataset. No high correlation is found, suggesting that no autocorrelation problem should affect the regression analysis provided in Section 5.5.2.

Table A5.3 reports all regression models on the expected (full) profile price. The main insight is that model coefficients are relatively stable as `Resources` are added in. When including bots on sale in the regression (Table A5.4), coefficients appear relatively stable and in line with those reported in Table A5.3, both in terms of trend and magnitude. An exception is $\log(GDP)$ in Model 5a, where the respective coefficient is not significant and drops in value when compared to Model 4a and Model 6a. This may suggest a correlation between $\log(GDP)$ and the presence of `Commerce` resources for profiles on sale, that is not present or weaker for profiles at full price.

A

The malware shift

In Nov 2019, Genesis Market reported an update in the malware, necessary to deal with changes introduced in the Chrome browser that appear to have affected the malware functionality. Confirmation of massive phishing campaigns in that period associated with the AZORult malware comes independently from Kaspersky and other researchers [105, 39, 146] and we found evidence of increased activity from the market in late Nov 2019 (Figure A5.3).

Table A5.3: Statistical models for profiles sold at full price.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
β_0	8.71***	−12.11***	−11.64***	−2.08*	2.24**	−2.15**	−5.54***	−5.57***	−3.70***
	(0.08)	(1.21)	(1.13)	(0.87)	(0.80)	(0.80)	(0.82)	(0.81)	(0.63)
Real Fngrpr	1.06***	0.69***	0.80***	1.15***	1.30***	1.30***	1.30***	1.31***	1.11***
	(0.14)	(0.16)	(0.15)	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)	(0.07)
log(GDP)		2.29***	2.19***	0.87***	0.24**	0.49***	0.44***	0.46***	0.42***
		(0.12)	(0.11)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.06)
Crypto			21.74***	15.19***	14.15***	13.87***	13.72***	13.62***	10.12***
			(0.65)	(0.51)	(0.46)	(0.45)	(0.44)	(0.44)	(0.34)
Money Transfer				12.30***	9.91***	9.21***	9.07***	8.86***	6.20***
				(0.17)	(0.17)	(0.17)	(0.16)	(0.16)	(0.13)
Commerce					5.94***	5.25***	5.27***	5.06***	3.22***
					(0.15)	(0.15)	(0.15)	(0.15)	(0.12)
Social						3.50***	3.52***	3.44***	1.68***
						(0.15)	(0.15)	(0.15)	(0.12)
Services							4.08***	3.95***	2.31***
							(0.29)	(0.29)	(0.22)
Other								4.22***	0.89***
								(0.31)	(0.24)
Resources									0.10***
									(0.00)
R^2	<0.01	0.04	0.18	0.52	0.60	0.63	0.64	0.65	0.79
Adj. R^2	<0.01	0.04	0.18	0.52	0.60	0.63	0.64	0.65	0.79
Num. obs.	7123	7123	7123	7123	7123	7123	7123	7123	7123

$^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05$

Table A5.4: Statistical models for all profiles (sold at full price and on sale).

	Model 1a	Model 2a	Model 3a	Model 4a	Model 5a	Model 6a	Model 7a	Model 8a	Model 9a	Model 10a
β_0	8.74***	−8.76***	−8.29***	−0.17	3.52***	−0.29	−3.49***	−3.46***	−2.26***	−2.10***
	(0.08)	(0.85)	(0.80)	(0.63)	(0.58)	(0.58)	(0.60)	(0.59)	(0.48)	(0.44)
Real Fngrpr	1.07***	1.22***	1.29***	1.52***	1.66***	1.69***	1.69***	1.70***	1.49***	1.03***
	(0.14)	(0.14)	(0.13)	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)	(0.07)	(0.07)
log(GDP)		1.77***	1.69***	0.58***	0.03	0.25***	0.21***	0.21***	0.24***	0.36***
		(0.09)	(0.08)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.05)	(0.04)
Crypto			19.23***	13.76***	12.72***	12.48***	12.34***	12.25***	9.08***	8.91***
			(0.49)	(0.39)	(0.36)	(0.35)	(0.34)	(0.34)	(0.28)	(0.25)
Money Transfer				10.95***	8.82***	8.25***	8.12***	7.94***	5.58***	5.37***
				(0.13)	(0.12)	(0.12)	(0.12)	(0.12)	(0.10)	(0.09)
Commerce					5.21***	4.60***	4.60***	4.42***	2.71***	2.74***
					(0.11)	(0.11)	(0.11)	(0.11)	(0.09)	(0.08)
Social						3.02***	3.08***	3.01***	1.42***	1.50***
						(0.11)	(0.11)	(0.11)	(0.09)	(0.08)
Services							3.83***	3.72***	2.13***	2.17***
							(0.21)	(0.21)	(0.17)	(0.16)
Other								3.54***	0.59**	0.69***
								(0.22)	(0.19)	(0.17)
Resources									0.10***	0.09***
									(0.00)	(0.00)
Sale										−3.40***
										(0.07)
R^2	<0.01	0.04	0.15	0.49	0.57	0.59	0.60	0.61	0.74	0.79
Adj. R^2	<0.01	0.04	0.15	0.49	0.57	0.59	0.60	0.61	0.74	0.79
Num. obs.	11683	11683	11683	11683	11683	11683	11683	11683	11683	11683

$^{***}p < 0.001, ^{**}p < 0.01, ^{*}p < 0.05$

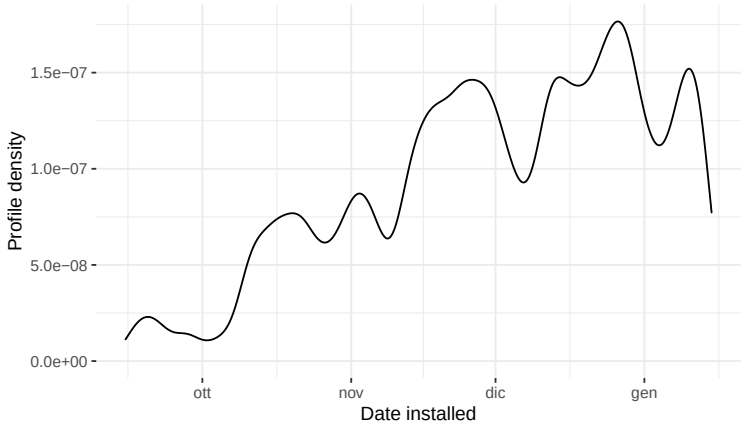


Figure A5.3: New profiles published on the market over time.

By analyzing the structure and the length of profile names sold on the market, it is possible to identify whether such profile derives from AZORult or other malware [254]; we further validated the claims of [254] by infecting our systems with AZORult and noting that the C2 server receives the stolen information labeled with a unique ID structured in a '8-8-8-8-8' pattern. From such analysis, we noted that between late Dec 2019 and Jan 2020, the market (temporarily) started dismissing AZORult in favor of another unknown malware (Figure A5.4) (underground platforms rumor about Racoon Stealer and Smokebot, but such claims are not confirmed from our side). In early Feb 2020, Google released Chrome 80, an update that provides AES encryption to the local db of saved credentials, *de facto* denying AZORult and similar malware to perform credential stealing; from our investigation, we found confirmation that the market operators definitely shifted from AZORult to another unknown infostealer malware in late Feb 2020 (Figure A5.4).

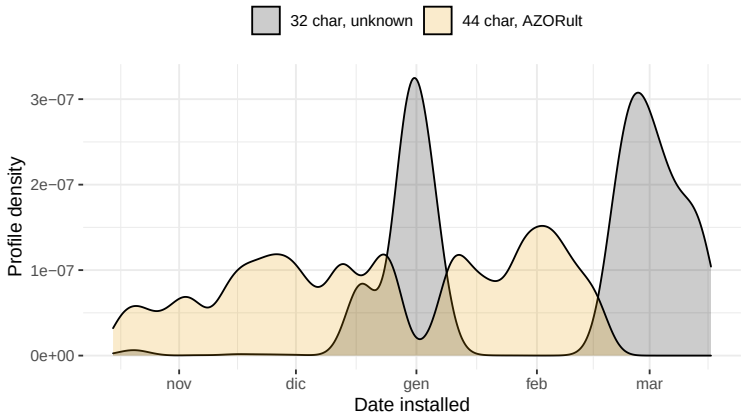


Figure A5.4: New profiles published on the market over time, divided by malware strain.

Chapter 6 appendix

Time window selection

Table A6.1 reports the relative fraction of overall observations that can be derived up to each monitoring day n. We note that $L_0^d \cup L_{1\cap\ldots\cap 6}^d$ (short for $L_0^d \cup (L_0^d \cap L_1^d) \cup \ldots \cup (L_0^d \cap L_1^d \cap \ldots \cap L_6^d)$) would allow us to maximize the number of appeared profiles as well as observations of sales while satisfying modelling constraints (Section 6.3.2). However, the sales model would have to account for all the variability in available alternatives at the purchase decision for any day d (Section 6.3.2), which proved to be computationally unfeasible for $n > 2$ (the model fails to converge). $L_0^d \cup L_{1\cap\ldots\cap 6}^d$ does, however, represent the overall empirical evidence we have of actual profile appearances and sales; hence, use it as a benchmark to evaluate the trade-off between the fraction of available profiles and the fraction of remaining sales up to observation day n.

Table A6.1: Available data points across monitoring periods.

	Available data points		
data $\forall d \in D$	obs days (%)	profiles (%)	sales (%)
$L_0^d \cup L_{1\cap\ldots\cap 6}^d$	107 (100.0%)	12′149 (100.0%)	2′051 (100.0%)
$L_0^d \cap L_1^d$	101 (94.4%)	11′357 (93.5%)	1′193 (58.2%)
$\ldots \cap L_2^d$	89 (83.2%)	9′778 (80.5%)	1′423 (69.4%)
$\ldots \cap L_3^d$	86 (80.4%)	9′445 (77.7%)	1′593 (77.7%)
$\ldots \cap L_4^d$	77 (72.0%)	8′071 (66.4%)	1′501 (73.2%)
$\ldots \cap L_5^d$	72 (67.3%)	7′560 (62.2%)	1′520 (74.1%)
$\ldots \cap L_6^d$	67 (62.6%)	6′860 (56.5%)	1′432 (69.8%)

Distinguishing sales from reservations

To verify the market's claims stating that the disappearance of a product exclusively depends on sales and that there are no other stochastic processes involved in the re-appearing of profiles besides the race condition between our crawling and a profile becoming reserved, we check for every listing day L_0^d if a disappeared product reappears in subsequent listing days $L_{1..6}^d$ and how often. Under the assumption that profiles only disappear if they're sold, L_{n+1}^d shall always contain a subset of L_n^d; the reservation mechanism introduces violations of this hypothesis, so we measure how often it occurs to evaluate how it compares to our expectations and to understand if it poses concerns on the validity of the sales detection technique. We check $L_{n+1}^d \cap L_n^d$, $\forall n \in [1..5]$. By analyzing the dataset containing information about $L_{0..6}^d$, out of the $6'860$ profiles available, only 74 reappeared over the next monitored days, representing the 1.08% of the total, against the expected 2.08%[1]. Further, we do not identify any profile that disappeared for more than 1 day; these results seem compatible with

[1]Given the maximum duration of a reservation being 30 minutes, the probability of missing a reserved profile is at most 2.08%, if it remains unsold.

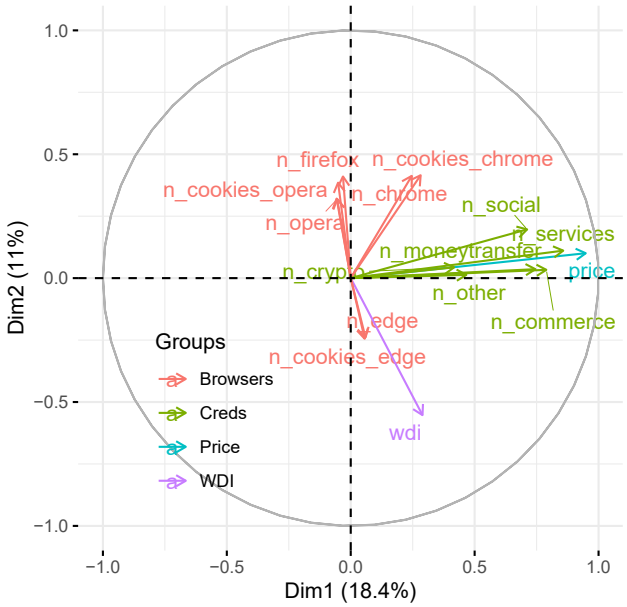


Figure A6.1: Two-dimensional representation of the variable vector space in output of the MFA.

the assumption that no other stochastic processes are involved in this phenomenon. As we consider $L^d_{0,1}$, we can identify false positives introduced when labelling a profile as 'sold' in the case of a profile reappearing on L^d_2; we identify 23 profiles of this type and correct their label accordingly.

Interpreting features against MFA dimensions

Table A6.2 reports the original variable variance captured by all MFA dimensions.

Table A6.2: Captured variance by MFA dimensions.

Dim.	1	2	3	4	5	6	7	8	9	10
var (%)	18.41	11.02	9.57	9.37	9.08	8.83	8.28	7.37	7.31	2.58
tot (%)	18.41	29.44	39.01	48.38	57.46	66.29	74.57	81.94	89.25	91.83
Dim.	11	12	13	14	15	16	17	18	19	20
var (%)	2.37	1.47	1.23	1.10	0.63	0.48	0.37	0.35	0.18	0.00
tot (%)	94.20	95.67	96.90	98.00	98.62	99.10	99.48	99.82	100.00	100.00

Figure A6.1 provides a representation over the two predominant dimensions (`Dim.1,2`, accounting for 29.44% of the overall data variance) of the (quantitative, as opposed to categorical) variable vector space; the projection of each vector onto each dimension represents how influential that variable is on the dimension, normalized at a group level; the closer variables are over a specific dimension, the more that dimension captures correlation among those variables. Colors represent variable groups; unsurprisingly, variables within

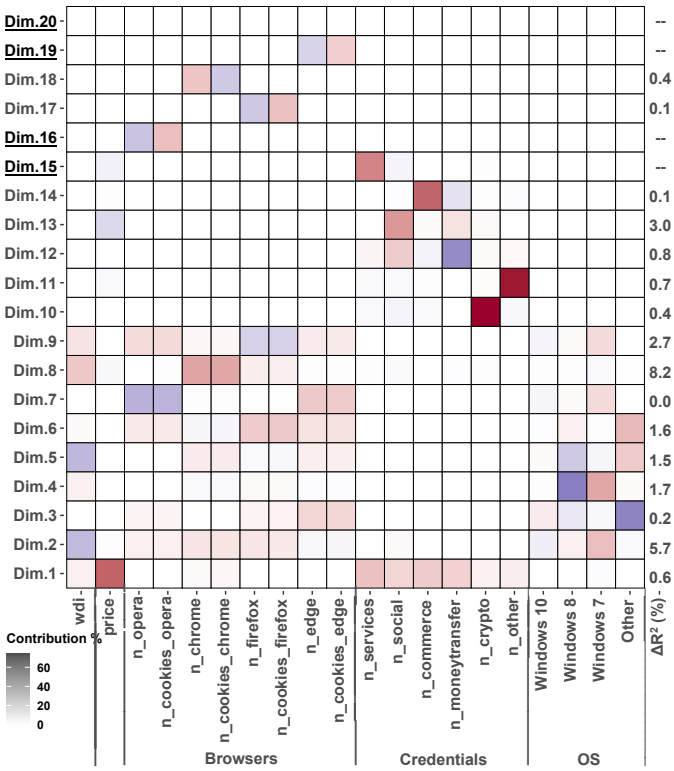


Figure A6.2: Variables' contribution to MFA dimensions. Underlined dimensions are not included in the final model. For brevity we include here the amount of variance explained by each dimension (ΔR^2) in the final model in the last column of the matrix.

the same group tend to be closely related to each other. `Browsers` and `WDI` appear to be of main relevance for `Dim.2`, whereas `Dim.1` represents mostly price and available credentials. Price and available resources seem to be highly and positively correlated (in agreement with [Campobasso7]); interestingly, `WDI` is highly but negatively correlated to the browser(s) characteristics of the affected user over these two dimensions; in particular, it emerges that profiles abundant in Edge profiles and cookies originate from more wealthy countries. From Figure A6.2, it is possible to observe that dimensions from 10 to 15 predominantly characterize profiles in terms of their available credentials; however, due to their low eigenvalues and variance captured, they fail to provide remarkable qualitative insights on the profile construction. Nonetheless, a few considerations can be made. `Dim.12` shows that generally profiles present an inverse correlation between the number of social and moneytransfer credentials. On the other hand, `Dim.13` indicates that whenever they are associated, and credentials from social media platforms are predominant, those tend to have a lower price than the average. For a more complete perspective on the relations between variables across different dimensions, we provide a repository containing all the possible combinations of two dimensions (as in Figure A6.1)[2].

[2]Link to the resources: `https://gitlab.tue.nl/impaas-mfa-plots/impaas-mfa-plots`

A

Model evaluation

We report the coefficients for the full model in the table below. In parenthesis, the standard error for each added dimension.

β_0	Dim.8	Dim.2	Dim.13	Dim.9	Dim.4	Dim.6	Dim.5	Dim.12
-2.51^{***}	0.62^{***}	-0.41^{***}	1.02^{***}	0.32^{***}	0.19^{***}	0.38^{***}	-0.17^{***}	-0.52^{***}
(0.05)	(0.04)	(0.04)	(0.09)	(0.04)	(0.03)	(0.04)	(0.04)	(0.08)
	Dim.11	Dim.1	Dim.10	Dim.18	Dim.3	Dim.14	Dim.17	Dim.7
	0.44^{***}	0.06^{*}	0.28^{***}	-0.76^{***}	-0.35^{***}	-0.30^{**}	0.38^{*}	0.09^{*}
	(0.06)	(0.02)	(0.05)	(0.18)	(0.03)	(0.10)	(0.17)	(0.04)

$std(c|day) = 0.25$, AIC=6564.8, BIC=6696.9, $R2m = 0.264$, $R2c = 0.278$, # Obs=$11'357$,
$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

We proceeded to build our final model adding one variable at a time and performing the analysis of the variance (ANOVA) for each new model. The final model explains the 27.8% of the data theoretical variance.

Model performance. In Section 6.4.2, we report model performance in terms of R^2; R^2_m represents the fraction of variance explained by the fixed effects; R^2_c includes variance explained by both fixed and random effects. $std(c|day)$ is the standard deviation of the random effect at the intercept.

We also compare the adopted model (accounting for purchase alternatives on a given day) to the same fixed effect model over the dataset with all the sales (regardless of how reliable the data collection was up to day n, first row of the same table). This results in a fitted model with the same coefficient directionality, but achieving only an $R^2 = 0.14$ against the obtained $R^2 = 0.278$ of the regression accounting for daily profile clusters for purchase alternatives. This further corroborates the importance of modelling the sales process with (tractable) factors accounting for the stochasticity introduced by alternative options on the purchase decision. To evaluate the discriminatory power of our sales prediction model, we run $1'000$ simulations to cross-validate our model with a randomly selected training set accounting for $2/3$ of the full dataset and validate it with the remaining records. For each simulation, we calculate the related area under curve (AUC). The median AUC value amounts to 0.757, and the related ROC curve is reported in Figure A6.3; the performance of the model is stable across simulations (68.2% $CI = [0.746, 0.768]$).

Data reconstruction and simulation

A

As mentioned in Section 6.4.1, data collected from January $21st$ to June $30th$ 2021 spans over a period of 161 days; the considered dataset $L^d_{0,1}$ offers 101 days with complete observations, leaving 60 days d to recreate or simulate as follows: (a) 6 days d have L^d_0 and no L^d_1; (b) 42 do not have L^d_0, but have L^d_1, (c) 6 have L^d_2, 1 has L^d_3, while (d) 4 have the last 24 hours market recap. For one day only (Feb 14th, 2021) we have no information at all; by looking at expected or collected profiles in the surrounding days (from 13 to 16 Feb), it appears that there was only 1 profile listed, thus leading us to conclude that for that day we would not

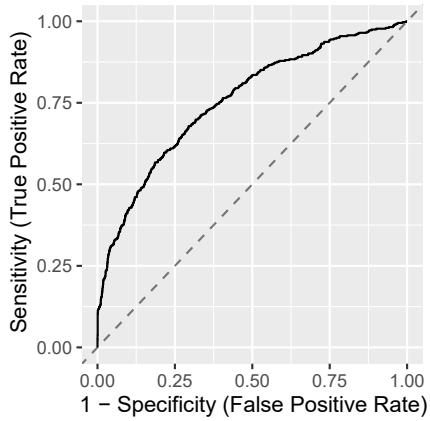


Figure A6.3: ROC curve with median AUC (0.757) over $1'000$ simulations.

have captured anything regardless. As mentioned in Section 6.3.2, we predict the sale outcome for the profiles in (a) using two different cutoff values for a more rigorous evaluation, one calculated to minimize false negatives ('stringent cutoff' - spec: 0.95, sens: 0.304), and another more 'generous' to improve the prediction's sensitivity (spec: 0.80, sens: 0.604); respectively, predictions report $53/792$ sold profiles in the former case and $176/792$ in the latter. To simulate products in (b), we compute the expected number of profiles by looking at the ratio of profiles counted during L_1^d and the offered L_0^d (17.6%, which is below the expected 25% due to **Ch4**); in case we miss L_1^d (c), we rely on the first L_i^d available by adjusting the previously calculated ratio with an additive factor approximating sales until that day, calculated from dataset $L_{0..6}^d$. Finally, to simulate products in (d), we calculate the average expected products as the ratio between the days for which we have available both market report and L_0^d profiles (16.20% of reported profiles are collected during L_0^d based on 95 observations; the low percentage is caused again from **Ch4**). With this information, we can now simulate the listings: we simulate the full market $10'000$ times by sampling with replacement an inversely proportional number of profiles from the 3 right-most and 3 left-most days for which L_0^d is available, including days from (a), and we do the whole process twice using the two different threshold values for (a).

A

Simulation validation. Our simulation strategy assumes that profiles appearing on the market on a certain day are similar to those appearing immediately before or after that day. To verify this assumption, we extract with replacement from the six closest L_0 listing days the total expected profiles; the extraction process is weighted based on how close a listing day is to the target listing day to refill. We simulate $1'000$ times the product listings of days for which we already have complete information, and we compare the MFA dimensions for the simulated profiles to the actual profiles. To visualize, we select four random L_0, two for which have all neighbor days with full information (e.g., Jan $24th$ is the L_0 to simulate and we have all L_0 from Jan $21st$ to Jan $27th$), and two for which we do not (e.g., the simulation extracts profiles from the six closest L_0 which may be 'further' than three days distance from the L_0 to

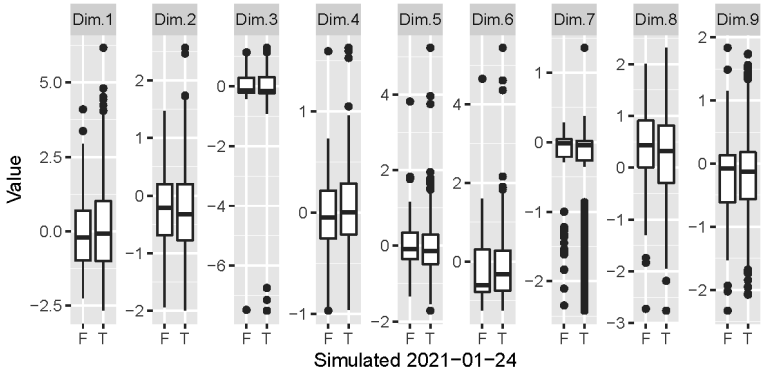


Figure A6.4: Simulation for January 24th, all neighbors.

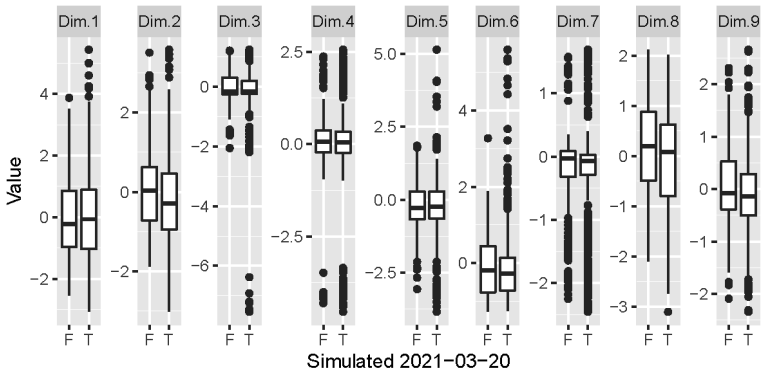


Figure A6.5: Simulation for March 20th, all neighbors.

simulate). The results are reported in Figures A6.4, A6.5, A6.6 and A6.7: for each dimension, the left boxplot shows the dimensions of the actual profiles, while the right one shows those from simulated profiles. The figures suggest that distributions across different dimensions show no significant differences between the expected and simulated values, suggesting that profiles appearing over contiguous days are similar to each other. A set of Wilcoxon Sign-ranked tests confirms this observation for the observed days across all dimensions used in the model in the 86% of cases, implying that our simulation strategy can reproduce a similar distribution of profiles for the days for which we have no observation.

A

Round up factor for market volumes

When estimating the market size and revenue, we have to account that (a) we consider only sales up to one day of market activity and (b) we sample only the 25% of the available products at Moscow's midnight. In Section 6.4.1, we discussed the dimensions' similarity between products sold within 24 hours and those sold later; to estimate sales happening after L_1^d, we

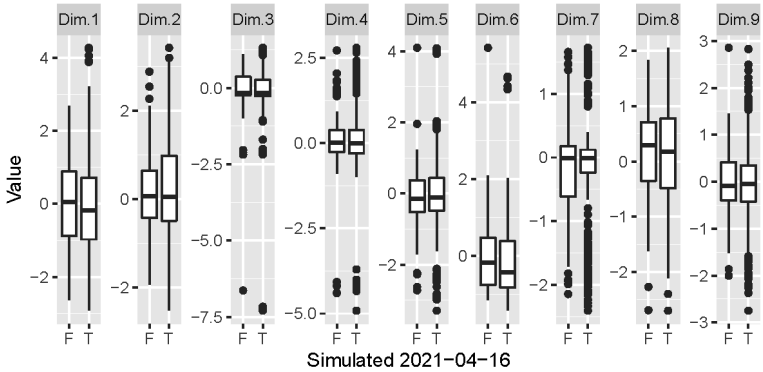


Figure A6.6: Simulation for April 16th, partial neighbors.

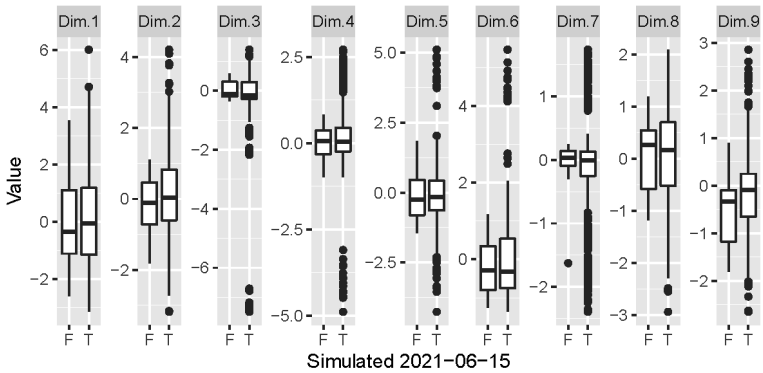


Figure A6.7: Simulation for June 15th, partial neighbors.

consider $L^d_{0..6}$ and compute the fraction of observed sold profiles during $L^d_{2..6}$ over the whole period, accounting for the 49% of all sales that would occur during the full six days period of monitoring. Therefore, we'll adjust our sales estimation to cover this fraction of unmeasured sales. With regards to (b), we empirically observed that the listing of a profile can be delayed (up to) some hours; comparing L^d_0 and L^d_1 cardinality, we note that we sample 17.58% of products on average, instead of the expected 25%. These two factors scale our estimations of 11.14 to represent the actual volumes of Genesis Market. To err on the conservative side and to aid comparisons, we round it down to 10 in the presentation.

A

Chapter 7 appendix

Market mechanisms

In this section, we enumerate and briefly describe the specific implementations we identified in the markets under analysis, for each of the mechanisms outlined earlier in this section. Unless otherwise specified, the features are considered binary.

Restricted access model. We find several levels of access restriction in the cybercrime forums under analysis:

F1. *Restricted sections - pull-in.* Access to one or more sections of the market can be granted by a pull-in mechanism (administrator authorization, application, community vote, invite from member).

F2. *Restricted sections - payment.* Access to one or more sections of the market can be granted upon the payment of a recurrent or one-off fee.

F3. *Restricted registration - pull-in.* Access to the entire market can be granted by a pull-in mechanism (administrator authorization, application, community vote, invite from member).

F4. *Restricted registration - payment.* Access to the entire market that could be granted upon the payment of a recurrent or one-off fee.

Dispute resolution system. Markets in our selection implement dispute resolution systems in the following ways:

F5. *Scammers banned.* We find evidence that scammers are be banned from the market when found guilty.

F6. *Working dispute resolution system.* We find evidence that the market offers an active dispute resolution system where users and administrators interact on the basis of provided evidence to evaluate each case.

F7. *Neutral mediator.* We do not find any evidence that the dispute mediator(s) have other functions in the market or act as seller themselves.

Reputation systems. Reputation systems in our market selection are characterized by:

F8. *Seller status - verification.* The market allows sellers to obtain a privileged 'seller status' via manual verification.

F9. *Seller status - payment.* The market requires sellers to pay a fee to obtain a privileged 'seller status'.

F10. *User status - verification.* The market allows any users to obtain a generic privileged status via manual verification.

F11. *User status - payment.* The market requires any users to pay a fee to obtain a generic privileged status.

F12. *Reputation change on trade.* A reputation score can exclusively be changed as a consequence of trading activities, as opposed to arbitrarily at any time and by any user.

F13. *Reputation change by VIP.* A reputation score can exclusively be changed by members with a high status on the market, as opposed to by any user in the forum.

Mitigation of perverse incentives. Forum markets in our selection implement several features that may mitigate perverse incentives leading to, for example, exit scams by the market administrators.

F14. *Seller status - recurrent fee.* By payment of a periodic fee to the market administrators, sellers can maintain a seller privileged status.

F15. *User status - recurrent fee.* By payment of a periodic fee to the market administrators, users can maintain a generic privileged status.

F16. *Escrow fee.* The market imposes an escrow fee imposed to all transactions using the escrow service. Fees can be a fixed amount in fiat currency, and a fixed or variable percentage of the transaction value. This feature is excluded from the analysis.

F17. *Sponsored ads.* The market offers the possibility to pay for sponsored ads.

Rule system. To provide a common baseline for trade, there should be rules and these should be enforced. Within our markets, we find the following:

F18. *Clear trade rules.* Trade activities are regulated by a dedicated and enforceable set of rules.

F19. *Moderator roles.* The role of moderators in the market is defined by specific rules in the market. We define a non-binary scale to classify the possible outcomes:

0. no moderation
1. moderators exist
2. moderators exist and can be looked up from a member list
3. moderators exist and operate in specific sections
4. moderators exist and rules explicitly mention their roles
5. combines the properties of 2 and 4
6. combines the properties of 3 and 4
7. combines the properties of 2, 3 and 4.

This feature is excluded from the analysis.

F20. *Active moderation.* Moderators appear to be active on the market (e.g.,conducting moderating activities during trials).

Strong authentication and anti-bot features To limit unwanted activity and verify the human nature of users, markets in our selection implement the following:

F21. *2FA available.* The market supports two-factor-authentication for its users.

F22. *CAPTCHA on authentication.* The authentication procedure to the market is protected by CAPTCHAs.

F23. *CAPTCHA on access.* CAPTCHAs are prompted at each market access.

Escrow system. Escrowing systems in our market selection can be:

F24. *Escrow available.* The market provides its members with an escrow system for transactions.

F25. *Escrow recommended.* The market explicitly recommends its members to use an escrow system for transactions.

Interaction model. Interaction among participants in our selection of forum markets can be:

F26. *Public interaction.* The market allows users to publicly interact with all members.

F27. *Private interaction.* The market allows users to privately interact with all members.

F28. *Pay to show content.* Authors of a post or comment can require other users to pay a fee to show posted content.

My Publications

[Campobasso1] **Michele Campobasso and Luca Allodi**. Know Your Cybercriminal: Evaluating Attacker Preferences by Measuring Profile Sales on an Active, Leading Criminal Market for User Impersonation at Scale. *32nd USENIX Security Symposium (USENIX Security 2023)*, August 2023.

[Campobasso2] **Michele Campobasso and Luca Allodi**. You Can Tell a Cybercriminal by the Company they Keep: A Framework to Infer the Relevance of Underground Communities to the Threat Landscape. *22nd Workshop on the Economics of Information Security (WEIS 2023)*, July 2023.

[Campobasso3] **Michele Campobasso and Luca Allodi**. `THREAT/crawl`: a Trainable, Highly-Reusable, and Extensible Automated Method and Tool to Crawl Criminal Underground Forums. *17th Symposium on Electronic Crime Research (APWG eCrime 2022)*, November 2022.

[Campobasso4] **Martin Rosso, Michele Campobasso, Ganduulga Gankhuyag, and Luca Allodi**. `SAIBERSOC`: A Methodology and Tool for Experimenting with Security Operation Centers. *Digital Threats: Research and Practice (DTRAP)*, 3(2):1–29, February 2022.

[Campobasso5] **Winnona DeSombre, Michele Campobasso, Luca Allodi, James Shires, JD Work, Robert Morgus, Patrick Howell O'Neill, and Trey Herr**. A Primer on the Proliferation of Offensive Cyber Capabilities. *In-Depth Research & Reports, 2021, Atlantic Council*, March 2021.

[Campobasso6] **Martin Rosso, Michele Campobasso, Ganduulga Gankhuyag, and Luca Allodi**. `SAIBERSOC`: Synthetic Attack Injection to Benchmark and Evaluate the peRformance of Security Operation Centers. *2020 Annual Computer Security Applications Conference (ACSAC)*, pages 141–153, December 2020.

[Campobasso7] **Michele Campobasso and Luca Allodi**. Impersonation-as-a-Service: Characterizing the Emerging Criminal Infrastructure for User Impersonation at Scale. *2020 ACM SIGSAC Conference on Computer and Communications Security (CCS)*, pages 1665–1680, November 2020.

[Campobasso8] **Michele Campobasso, Pavlo Burda, and Luca Allodi**. `CARONTE`: Crawling Adversarial Resources Over Non-Trusted, high-profile Environments. *1st Workshop on Attackers and Cyber-Crime Operations (WACCO) – 2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, pages 433–442, July 2019.

References

[1] **Abbasi, A., Li, W., Benjamin, V., Hu, S., and Chen, H.** Descriptive analytics: Examining expert hackers in web forums. In *Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint* (2014), IEEE, pp. 56–63.

[2] **Abdi, H., and Valentin, D.** Multiple Factor Analysis (MFA). *Encyclopedia of Measurement and Statistics* (Jan 2007).

[3] **Ablon, L., and Libicki, M.** Hacker's Bazaar: The Markets for Cybercrime Tools and Stolen Data. *Defense Counsel Journal 82* (2015), 143.

[4] **Abrams, L.** Dutch Police Post "say no to cybercrime" warnings on hacker forums, Online: `https://www.bleepingcomputer.com/news/security/dutch-police-post-say-no-to-cybercrime-warnings-on-hacker-forums/`.

[5] **Acar, G., Juarez, M., and individual contributors**. tor-browser-selenium - Tor Browser automation with Selenium, Online: `https://github.com/webfp/tor-browser-selenium`.

[6] **Adams, D.** *The Hitchhiker's Guide to the Galaxy*. Pan Books, 1979.

[7] **Afroz, S., Garg, V., McCoy, D., and Greenstadt, R.** Honor among thieves: A common's analysis of cybercrime economies. In *2013 APWG eCrime Researchers Summit* (2013), IEEE, pp. 1–11.

[8] **Agresti, A.** *Categorical data analysis*, vol. 482. John Wiley & Sons, 2003.

[9] **Aiken, M., Davidson, J., Amann, P., et al.** Youth pathways into cybercrime. *European Cybercrime Center (EC3)* (2016).

[10] **Akanle, O., and Shadare, B. R.** Why has it been so difficult to Counteract Cyber Crime in Nigeria? Evidence from an Ethnographic Study. *International Journal of Cyber Criminology 14*, 1 (2020), 29–43.

[11] **Akerlof, G. A.** The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics 84*, 3 (1970), 488–500.

[12] **Akyazi, U., van Eeten, M., and Gañán, C. H.** Measuring Cybercrime as a Service (CaaS) Offerings in a Cybercrime Forum. In *20th Workshop on the Economics of Information Security (WEIS 2021)*.

[13] **Alaca, F., and Van Oorschot, P. C.** Device fingerprinting for augmenting web authentication: classification and analysis of methods. In *32nd Annual Conference on Computer Security Applications* (2016), pp. 289–301.

[14] **Aliapoulios, M., Ballard, C., Bhalerao, R., Lauinger, T., and McCoy, D.** Swiped: Analyzing Ground-truth Data of a Marketplace for Stolen Debit and Credit Cards. In *30th USENIX Security Symposium (USENIX Security 2021)*.

[15] **Allodi, L.** Economic factors of vulnerability trade and exploitation. In *2017 ACM SIGSAC Conference on Computer and Communications Security (CCS)* (oct 2017), pp. 1483–1499.

[16] **Allodi, L., Corradin, M., and Massacci, F.** Then and now: On the maturity of the cybercrime markets the lesson that black-hat marketeers learned. *IEEE Transactions on Emerging Topics in Computing 4*, 1 (2015), 35–46.

[17] **Allodi, L., Massacci, F., and Williams, J.** The Work-Averse Cyberattacker Model: Theory and Evidence from Two Million Attack Signatures. *Risk Analysis 42*, 8 (2022), 1623–1642.

[18] **Alrwais, S., Liao, X., Mi, X., Wang, P., Wang, X., Qian, F., Beyah, R., and McCoy, D.** Under the shadow of sunshine: Understanding and detecting bulletproof hosting on legitimate service provider networks. In *2017 IEEE Symposium on Security and Privacy (S&P)* (2017), IEEE, pp. 805–823.

[19] **Anderson, L. K., Taylor, J. R., and Holloway, R. J.** The consumer and his alternatives: An experimental approach. *Journal of Marketing Research 3*, 1 (1966), 62–67.

[20] **Anderson, R., Barton, C., Böhme, R., Clayton, R., Van Eeten, M. J., Levi, M., Moore, T., and Savage, S.** Measuring the Cost of Cybercrime. *The Economics of Information Security and Privacy* (2013), 265–300.

[21] **Anderson, R., Clayton, R., Böhme, R., and Collier, B.** Silicon den: Cybercrime is entrepreneurship. In *19th Workshop on the Economics of Information Security (WEIS 2021)*.

[22] **Anderson, R., and Moore, T.** The economics of information security. *Science 314* (2006).

[23] **Andrioaie, A.** Redline malware is wreaking havoc with passwords stored in web browsers, Online: `https://heimdalsecurity.com/blog/redline-malware-steals-your-password-from-the-browser/`.

[24] **Arghire, B.** Iranian hackers exploit recent office 0-day in attacks: Report. Online: `https://www.securityweek.com/iranian-hackers-exploit-recent-office-0-day-attacks-report`. *SecurityWeek* (May 2017).

[25] **Armin, J., Thompson, B., Ariu, D., Giacinto, G., Roli, F., and Kijewski, P.** 2020 Cybercrime Economic Costs: No Measure no Solution. In *2015 10th International Conference on Availability, Reliability and Security* (2015), IEEE, pp. 701–710.

[26] **Baeza-Yates, R., Ribeiro, B. d. A. N., et al.** *Modern Information Retrieval*. New York: ACM Press; Harlow, England: Addison-Wesley, 2011.

[27] **Bai, Q., Xiong, G., Zhao, Y., and He, L.** Analysis and Detection of Bogus Behavior in Web Crawler Measurement. *Procedia Computer Science 31* (2014), 1084–1091.

[28] **Balla, A., Stassopoulou, A., and Dikaiakos, M. D.** Real-time web crawler detection. In *2011 18th International Conference on Telecommunications (ICT)* (2011), IEEE, pp. 428–432.

[29] **Bamford, J., and De Chant, T.** Exclusive: Edward Snowden on cyber warfare. Online: https://www.pbs.org/wgbh/nova/article/snowden-transcript/. *PBS - Exclusive: Edward Snowden on Cyber Warfare* (Jan 2015).

[30] **Bank, T. W.** World Development Indicators, Online: https://datacatalog.worldbank.org/dataset/world-development-indicators.

[31] **BBC News**. Cyber-attack: Europol says it was unprecedented in scale. Online: https://www.bbc.com/news/world-europe-39907965. *BBC News* (May 2017).

[32] **Benjamin, V., Li, W., Holt, T., and Chen, H.** Exploring threats and vulnerabilities in hacker web: Forums, IRC and carding shops. In *Intelligence and Security Informatics (ISI), 2015 IEEE International Conference* (2015), IEEE, pp. 85–90.

[33] **Benjamin, V., Samtani, S., and Chen, H.** Conducting large-scale analyses of underground hacker communities. In *Cybercrime Through an Interdisciplinary Lens*. Routledge, 2016, pp. 70–89.

[34] **Benjamin, V., Valacich, J. S., and Chen, H.** DICE-E: A Framework for Conducting Darknet Identification, Collection, Evaluation with Ethics. *MIS Quarterly 43*, 1 (2019).

[35] **Bermudez-Villalva, A., and Stringhini, G.** The shady economy: Understanding the difference in trading activity from underground forums in different layers of the web. In *16th APWG Symposium on Electronic Crime Research (APWG eCrime)* (2021), IEEE, pp. 1–10.

[36] **Bhalerao, R., Aliapoulios, M., Shumailov, I., Afroz, S., and McCoy, D.** Mapping the Underground: Supervised Discovery of Cybercrime Supply Chains. In *14th APWG Symposium on Electronic Crime Research (APWG eCrime)* (2019), IEEE, pp. 1–16.

[37] **Bing, C., and Schectman, J.** Exclusive: Ex-NSA cyberspies reveal how they helped hack foes of UAE. Online: https://www.reuters.com/investigates/special-report/usa-spying-raven/. *Reuters*.

[38] **Binsalleeh, H., Ormerod, T., Boukhtouta, A., Sinha, P., Youssef, A., Debbabi, M., and Wang, L.** On the analysis of the zeus botnet crimeware toolkit. In *Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on* (2010), IEEE, pp. 31–38.

[39] **Bisson, D.** AZORult Trojan Disguised Itself as Fake ProtonVPN Installer, Online: https://www.tripwire.com/state-of-security/featured/azorult-trojan-disguised-itself-as-fake-protonvpn-installer/.

[40] **BlackBerry**, Online: https://blogs.blackberry.com/en/2018/09/unpacking-a-packer-powershell-obfuscation-using-securestring.

[41] **Bonneau, J., Herley, C., Van Oorschot, P. C., and Stajano, F.** The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In *2012 IEEE Symposium on Security and Privacy (S&P)* (2012), IEEE, pp. 553–567.

[42] **Bonneau, J., Herley, C., Van Oorschot, P. C., and Stajano, F.** Passwords and the evolution of imperfect authentication. *Communications of the ACM 58*, 7 (2015), 78–87.

[43] **Bouwman, X., Griffioen, H., Egbers, J., Doerr, C., Klievink, B., and Van Eeten, M.** A different cup of {TI}? The added value of commercial threat intelligence. In *29th USENIX Security Symposium (USENIX Security 2020)*, pp. 433–450.

[44] **Box, G. E., and Hunter, J. S.** The 2 k-p fractional factorial designs. *Technometrics 3*, 3 (1961), 311–351.

[45] **Brunt, R., Pandey, P., and McCoy, D.** Booted: An Analysis of a Payment Intervention on a DDoS-for-Hire Service. In *16th Workshop on the Economics of Information Security (WEIS 2017)*.

[46] **Bursztein, E., Benko, B., Margolis, D., Pietraszek, T., Archer, A., Aquino, A., Pitsillidis, A., and Savage, S.** Handcrafted fraud and extortion: Manual account hijacking in the wild. In *2014 Conference on Internet Measurement Conference* (2014), pp. 347–358.

[47] **Caballero, J., Grier, C., Kreibich, C., and Paxson, V.** Measuring Pay-per-Install: The Commoditization of Malware Distribution. In *20th USENIX Security Symposium (USENIX Security 2011)* (2011).

[48] **Cai, R., Yang, J.-M., Lai, W., Wang, Y., and Zhang, L.** iRobot: An intelligent crawler for Web forums. In *17th International Conference on World Wide Web* (2008), ACM, pp. 447–456.

[49] **Chen, Y., Pavlov, D., and Canny, J. F.** Large-scale behavioral targeting. In *15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2009), pp. 209–218.

[50] **CHEQ**. What are the different types of anti-crawler protection?, Online: https://cheq.ai/blog/what-are-the-different-types-of-anti-crawler-protection/.

[51] **Chohan, R.** *Opportunistic Behavior in Industrial Marketing Relationships*. PhD thesis, Luleå University of Technology, 2020.

[52] **Christin, N.** Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. In *22nd International Conference on World Wide Web* (2013), pp. 213–224.

[53] **Cimpanu, C.** Two Trend Micro zero-days exploited in the wild by hackers. Online: https://www.zdnet.com/article/two-trend-micro-zero-days-exploited-in-the-wild-by-hackers/. *ZDNET* (mar 2020).

[54] **Citizen Lab**. BLASTPASS: NSO Group iphone zero-click, Zero-day exploit captured in the wild, Online: https://citizenlab.ca/2023/09/blastpass-nso-group-iphone-zero-click-zero-day-exploit-captured-in-the-wild/.

[55] **CleanTalk**. CleanTalk Anti-Flood and Anti-Crawler (Bot protection) Options, Online: https://cleantalk.org/help/anti-flood-and-anti-crawler.

[56] **Cloud, H.** Configuring Anti-Crawler Rules to Prevent Crawler Attacks, Online: https://support.huaweicloud.com/intl/en-us/bestpractice-waf/waf_06_0006.html.

[57] **Cloudflare**. Using privacy pass with cloudflare, Online: https://support.cloudflare.com/hc/en-us/articles/115001992652-Using-Privacy-Pass-with-Cloudflare. cloudflare help center.

[58] **Cobalt Strike**. Cobalt Strike training, Online: https://www.cobaltstrike.com/training.

[59] **Coleman, J. L.** Rational Choice and Rational Cognition. *Legal Theory 3*, 2 (1997), 183–203.

[60] **Collier, B., and Clayton, R.** A "sophisticated attack"? innovation technical sophistication and creativity in the cybercrime ecosystem. In *21st Workshop on the Economics of Information Security (WEIS 2022)*.

[61] **Collier, B., Clayton, R., Hutchings, A., and Thomas, D.** Cybercrime is (often) boring: maintaining the infrastructure of cybercrime economies. *19th Workshop on the Economics of Information Security (WEIS 2020)* (2020).

[62] **Collier, B., Thomas, D. R., Clayton, R., and Hutchings, A.** Booting the booters: Evaluating the effects of police interventions in the market for denial-of-service attacks. In *2019 ACM SIGCOMM Conference on Internet Measurements* (2019), pp. 50–64.

[63] **Collins, M. E., and Loughran, T. A.** Rational choice theory, heuristics, and biases. *The Oxford handbook of offender decision making 6* (2017), 10.

[64] **Corporation, M.** MITRE ATT&CK Framework Tactics, Online: https://attack.mitre.org/tactics/enterprise/.

[65] **CTRLBOX Consulting**. Cyber Crime Incident Tracker - Cached copy (website offline), Online: https://web.archive.org/web/20221205110731/https://www.arresttracker.com/pages.

[66] **Cylance**. Threat Spotlight: Analyzing AZORult Infostealer Malware, Online: https://threatvector.cylance.com/en_us/home/threat-spotlight-analyzing-azorult-infostealer-malware.html.

[67] **Das, A., Bonneau, J., Caesar, M., Borisov, N., and Wang, X.** The tangled web of password reuse. In *21st Network and Distributed System Security Symposium (NDSS)* (2014), vol. 14, pp. 23–26.

[68] **Davidson, A., Goldberg, I., Sullivan, N., Tankersley, G., and Valsorda, F.** Privacy Pass: Bypassing Internet Challenges Anonymously. *Privacy Enhancing Technologies 2018*, 3 (2018), 164–180.

[69] **Décary-Hétu, D., and Aldridge, J.** Sifting through the net: Monitoring of online offenders by researchers. *European Review of Organised Crime 2*, 2 (2015), 122–141.

[70] **Décary-Hétu, D., and Dupont, B.** The Social Network of Hackers. *Global Crime 13*, 3 (2012), 160–175.

[71] **Décary-Hétu, D., and Leppänen, A.** Criminals and signals: An assessment of criminal performance in the carding underworld. *Security Journal 29* (2016), 442–460.

[72] **Department of Justice**. Criminal Indictment, Online: `https://web.archive.org/web/20150211203259/https://cis.uab.edu/forensics/blog/Kostyukov.Indictment.pdf`.

[73] **Department of Justice**. First Superseeding Indictment, Online: `https://web.archive.org/web/20221029212509/https://dd80b675424c132b90b3-e48385e382d2e5d17821a5e1d8e4c86b.ssl.cf1.rackcdn.com/external/infraudsupersedingindictment.pdf`.

[74] **Department of Justice**. Text of the indictment charging 36 Defendants for Alleged Roles in Transnational Criminal Organization Responsible for Cybercrimes. Online: `http://www.derechos.org/nizkor/corru/doc/infraud2.html`. *Equipo Nizkor* (2018).

[75] **Department of Justice**. United States v. Andrey Turchin. Online: `https://www.justice.gov/usao-wdwa/united-states-v-andrey-turchin`. *Western District of Washington* (Jul 2020).

[76] **DeSombre, W., and Byrnes, D.** Thieves and Geeks: Russian and Chinese Hacking Communities. *Recorded Future* (2018).

[77] **Dingledine, R., Mathewson, N., and Syverson, P.** Tor: The Second-Generation Onion Router. In *13th USENIX Security Symposium (USENIX Security 2004)*, vol. 4, pp. 303–320.

[78] **Dmitrienko, A., Liebchen, C., Rossow, C., and Sadeghi, A.-R.** On the (In)Security of Mobile Two-Factor Authentication. In *Financial Cryptography and Data Security* (Berlin, Heidelberg, 2014), N. Christin and R. Safavi-Naini, Eds., Springer Berlin Heidelberg, pp. 365–383.

[79] **Doran, D., and Gokhale, S. S.** Web robot detection techniques: overview and limitations. *Data Mining and Knowledge Discovery 22*, 1-2 (2011), 183–210.

[80] **Doran, D., Morillo, K., and Gokhale, S. S.** A comparison of web robot and human requests. In *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (2013), ACM, pp. 1374–1380.

[81] **Dupont, B., Côté, A.-M., Boutin, J.-I., and Fernandez, J.** Darkode: Recruitment patterns and transactional features of "the most dangerous cybercrime forum in the world". *American Behavioral Scientist 61*, 11 (2017), 1219–1243.

[82] **Dupont, B., Côté, A.-M., Savine, C., and Décary-Hétu, D.** The ecology of trust among hackers. *Global Crime 17*, 2 (2016), 129–151.

[83] **Dupont, B., and Lusthaus, J.** Countering distrust in illicit online networks: The dispute resolution strategies of cybercriminals. *Social Science Computer Review 40*, 4 (2022), 892–913.

[84] **Dupret, G., and Liao, C.** A Model to Estimate Intrinsic Document Relevance from the Clickthrough Logs of a Web Search Engine. In *3rd ACM International Conference on Web Search and Data Mining* (2010), ACM, pp. 181–190.

[85] **Dupret, G. E., and Piwowarski, B.** A User Browsing Model to Predict Search Engine Click Data from Past Observations. In *31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (2008), ACM, pp. 331–338.

[86] **Eastern District of Wisconsin**. Seizure warrant for Genesis Market domains, Online: `https://s3.documentcloud.org/documents/23742615/genesis-market.pdf`.

[87] **Eisenhardt, K. M.** Agency theory: An assessment and review. *Academy of management review 14*, 1 (1989), 57–74.

[88] **EUROPOL**. Takedown of notorious hacker marketplace selling your identity to criminals, Online: `https://www.europol.europa.eu/media-press/newsroom/news/takedown-of-notorious-hacker-marketplace-selling-your-identity-to-criminals`.

[89] **EUROPOL**. Massive blow to criminal dark web activities after globally coordinated operation, Online: `https://www.europol.europa.eu/media-press/newsroom/news/massive-blow-to-criminal-dark-web-activities-after-globally-coordinated-operation`.

[90] **F-Secure**. Vulnerability, Online: `https://www.f-secure.com/v-descs/articles/vulnerability.shtml`.

[91] **Fallmann, H., Wondracek, G., and Platzer, C.** Covertly probing underground economy marketplaces. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment* (2010), Springer, pp. 101–110.

[92] **Fang, Y., Guo, Y., Huang, C., and Liu, L.** Analyzing and Identifying Data Breaches in Underground Forums. *IEEE Access 7* (2019), 48770–48777.

[93] **Felson, M.** The process of co-offending. *Crime Prevention Studies 16* (2003), 149–168.

[94] **Figueroa, M., Bing, N., and Silvestrini, B.** The conti leaks - insight into a ransomware unicorn, Online: `https://www.breachquest.com/blog/conti-leaks-insight-into-a-ransomware-unicorn/`.

[95] **Fischerkeller, M. P., and Harknett, R. J.** Cyber Persistence Intelligence Contests and Strategic Competition. *Part 5 of 'Policy Roundtable: Cyber Conflict as an Intelligence Contest in 'Cyber Competition, Special Issue* (2020).

[96] **Ford, R., and Ray, H.** Googling for gold: Web crawlers, hacking and defense explained. *Network Security 2004*, 1 (2004), 10–13.

[97] **Franceschi-Bicchierai, L.** Who are the NSA's elite hackers? Online: `https://www.vice.com/en_us/article/bmvyxw/nsa-hacking-unit-tao-cyberwar`. *VICE* (Aug 2016).

[98] **Franklin, J., Perrig, A., Paxson, V., and Savage, S.** An inquiry into the nature and causes of the wealth of internet miscreants. In *ACM Conference on Computer and communications security* (2007), pp. 375–388.

[99] **Freeman, D., Jain, S., Dürmuth, M., Biggio, B., and Giacinto, G.** Who Are You? A Statistical Approach to Measuring User Authenticity. In *23rd Network and Distributed System Security Symposium (NDSS)* (2016), pp. 1–15.

[100] **Gailmard, S.** Accountability and Principal–Agent Theory. In *The Oxford Handbook of Public Accountability*. Oxford University Press, 05 2014.

[101] **Gallagher, S.** Helpful(?) coding tips from the CIA's School of Hacks. Online: `https://arstechnica.com/information-technology/2017/03/malware-101-the-cias-dos-and-donts-for-tool-developers/`. *Ars Technica* (Mar 2017).

[102] **Gallagher, S.** German police seize 'Bulletproof' hosting data center in former NATO bunker. Online: `https://arstechnica.com/information-technology/2019/09/german-police-seize-bulletproof-hosting-data-center-in-former-nato-bunker/`. *Ars Technica* (2019).

[103] **Gambetta, D.** *Codes of the Underworld*. Princeton University Press, 2011.

[104] **Gao, H., Hu, J., Wilson, C., Li, Z., Chen, Y., and Zhao, B. Y.** Detecting and characterizing social spam campaigns. In *10th ACM SIGCOMM Conference on Internet Measurement* (2010), pp. 35–47.

[105] **Gatlan, S.** AZORult Malware Infects Victims via Fake ProtonVPN Installer, Online: `https://www.bleepingcomputer.com/news/security/azorult-malware-infects-victims-via-fake-protonvpn-installer/`.

[106] **Geiger, M., Bauer, J., Masuch, M., and Franke, J.** An analysis of black energy 3, Crashoverride, and Trisis, three malware approaches targeting operational technology systems. In *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)* (2020), vol. 1, IEEE, pp. 1537–1543.

[107] **Georgoulias, D., Pedersen, J. M., Falch, M., and Vasilomanolakis, E.** A qualitative mapping of Darkweb marketplaces. In *16th APWG Symposium on Electronic Crime Research (APWG eCrime)* (2021), IEEE, pp. 1–15.

[108] **Goldsmith, A., and Wall, D. S.** The seductions of cybercrime: Adolescence and the thrills of digital transgression. Online: https://doi.org/10.1177/1477370819887305. *European Journal of Criminology 19*, 1 (2022), 98–117.

[109] **Goncharov, M.** Russian underground 101. *Trend Micro incorporated research paper* (2012), 51.

[110] **Goncharov, M.** Criminal hideouts for lease: Bulletproof hosting services. *Forward-Looking Threat Research (FTR) Team, A TrendLabsSM Research Paper 28* (2015).

[111] **Gray, I. W., Cable, J., Brown, B., Cuiujuclu, V., and McCoy, D.** Money Over Morals: A Business Analysis of Conti Ransomware. In *17th APWG Symposium on Electronic Crime Research (APWG eCrime)* (2022), IEEE, pp. 1–12.

[112] **Greenwald, G.** NSA collecting phone records of millions of Verizon customers daily. Online: https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order. *The Guardian* (Jun 2013).

[113] **Grier, C., Ballard, L., Caballero, J., Chachra, N., Dietrich, C. J., Levchenko, K., Mavrommatis, P., McCoy, D., Nappa, A., Pitsillidis, A., Provos, N., Rafique, M. Z., Rajab, M. A., Rossow, C., Thomas, K., Paxson, V., Savage, S., and Voelker, G. M.** Manufacturing compromise: the emergence of exploit-as-a-service. In *Proc. of CCS'12* (2012), ACM, pp. 821–832.

[114] **Guo, F., Liu, C., and Wang, Y. M.** Efficient Multiple-Click Models in Web Search. In *2nd ACM International Conference on Web Search and Data Mining* (2009), ACM, pp. 124–131.

[115] **Guo, Q., and Agichtein, E.** Beyond Dwell Time: Estimating Document Relevance from Cursor Movements and other Post-click Searcher Behavior. In *21st International Conference on World Wide Web* (2012), ACM, pp. 569–578.

[116] **Hald, S. L., and Pedersen, J. M.** An updated taxonomy for characterizing hackers according to their threat properties. In *2012 14th International Conference on Advanced Communication Technology (ICACT)* (2012), IEEE, pp. 81–86.

[117] **Haslebacher, A., Onaolapo, J., and Stringhini, G.** All your cards are belong to us: Understanding online carding forums. In *12th Symposium on Electronic Crime Research (APWG eCrime)*, IEEE, pp. 41–51.

[118] **Havron, S., Freed, D., Chatterjee, R., McCoy, D., Dell, N., and Ristenpart, T.** Clinical Computer Security for Victims of Intimate Partner Violence. In *28th USENIX Security Symposium (USENIX Security 2019)* (Santa Clara, CA), USENIX Association, pp. 105–122.

[119] **Herley, C.** Why do nigerian scammers say they are from nigeria? In *11th Workshop on the Economics of Information Security (WEIS 2012)*, Berlin.

[120] **Herley, C.** So long, and no thanks for the externalities: the rational rejection of security advice by users. In *Proc. of NSPW'09* (2009), NSPW '09, ACM, pp. 133–144.

[121] **Herley, C.** When does targeting make sense for an attacker? *2013 IEEE Symposium on Security and Privacy (S&P) 11*, 2 (2013), 89–92.

[122] **Herley, C., and Florêncio, D.** Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. In *Economics of Information Security and Privacy* (2010), Springer, pp. 33–53.

[123] **Hine, G. E., Onaolapo, J., De Cristofaro, E., Kourtellis, N., Leontiadis, I., Samaras, R., Stringhini, G., and Blackburn, J.** Kek, Cucks, and God Emperor Trump: A Measurement Study of 4chan's Politically Incorrect Forum and Its Effects on the Web. In *11th International AAAI Conference on Web and Social Media* (2017), vol. 11, pp. 92–101.

[124] **Ho, G., Cidon, A., Gavish, L., Schweighauser, M., Paxson, V., Savage, S., Voelker, G. M., and Wagner, D.** Detecting and Characterizing Lateral Phishing at Scale. In *28th USENIX Security Symposium (USENIX Security 2019)* (Santa Clara, CA), USENIX Association, pp. 1273–1290.

[125] **Ho, G., Javed, A. S. M., Paxson, V., and Wagner, D.** Detecting Credential Spearphishing Attacks in Enterprise Settings. In *26th USENIX Security Symposium (USENIX Security 2017)* (2017), pp. 469–485.

[126] **Holt, T. J., Smirnova, O., and Hutchings, A.** Examining signals of trust in criminal markets online. *Journal of Cybersecurity* (2016), tyw007.

[127] **Holt, T. J., Strumsky, D., Smirnova, O., and Kilger, M.** Examining the Social Networks of Malware Writers and Hackers. *International Journal of Cyber Criminology 6*, 1 (2012).

[128] **Huang, C., Guo, Y., Guo, W., and Li, Y.** HackerRank: Identifying key hackers in underground forums. *International Journal of Distributed Sensor Networks 17*, 5 (2021).

[129] **Huang, K., Siegel, M., and Madnick, S.** Systematically understanding the cyber attack business: A survey. *ACM Computing Surveys (CSUR) 51*, 4 (2018), 1–36.

[130] **Huggins, J., Gross, P., Wang, J. T., and individual contributors**. Selenium, A suite of tools for browser automation., Online: `https://www.selenium.dev/`.

[131] **Hutchings, A., and Clayton, R.** Exploring the provision of online booter services. *Deviant Behavior 37*, 10 (2016), 1163–1178.

[132] **Hutchings, A., and Holt, T. J.** A crime script analysis of the online stolen data market. *British Journal of Criminology 55*, 3 (2015), 596–614.

[133] **Hutchings, A., and Holt, T. J.** Interviewing cybercrime offenders. *Journal of Qualitative Criminal Justice and Criminology* (2018).

[134] **Hutchings, A., and Pastrana, S.** Understanding ewhoring. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)* (2019), IEEE, pp. 201–214.

[135] **Insikt Group**, Online: `https://go.recordedfuture.com/hubfs/reports/cta-2019-0724.pdf`.

[136] **Intel471**. Here's who is powering the bulletproof hosting market. Online: https://intel471.com/blog/top-bulletproof-hosting-providers-yalishanda-ccweb-brazzzers-2021.

[137] **IOActive**. Technical White Paper: Reversal and Analysis of Zeus and SpyEye Banking Trojans, Online: https://ioactive.com/pdfs/ZeusSpyEyeBankingTrojanAnalysis.pdf.

[138] **Ion, I., Reeder, R., and Consolvo, S.** "... no one can hack my mind": Comparing Expert and Non-Expert Security Practices. In *11th Symposium On Usable Privacy and Security (SOUPS 2015)* (2015), pp. 327–346.

[139] **Jacob, G., Kirda, E., Kruegel, C., and Vigna, G.** PUBCRAWL: Protecting Users and Businesses from CRAWLers. In *21st USENIX Security Symposium (USENIX Security 2012)* (2012), pp. 507–522.

[140] **Jacobs, J., Romanosky, S., Adjerid, I., and Baker, W.** Improving vulnerability remediation through better exploit prediction. *Journal of Cybersecurity 6*, 1 (2020), tyaa015.

[141] **Jiang, J., Song, X., Yu, N., and Lin, C.-Y.** Focus: learning to crawl web forums. *IEEE Transactions on knowledge and Data Engineering 25*, 6 (2012), 1293–1306.

[142] **Kadlecová, L.** Russian-speaking cybercrime: reasons behind its success. *Eur Rev Organised Crime 2*, 2 (2015), 104–121.

[143] **Kahneman, D., and Tversky, A.** Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*, 2 (1979), 263–292.

[144] **Karami, M., and McCoy, D.** Understanding the Emerging Threat of DDoS-as-a-Service. In *6th USENIX Workshop on Large-Scale Exploits and Emergent Threats* (2013).

[145] **Kaspersky**. Are your passwords stored securely? Kaspersky finds 60% rise in users hit by password stealers in 2019, Online: https://me-en.kaspersky.com/about/press-releases/2019_kaspersky-finds-rise-in-users-hit-by-password-stealers.

[146] **Kaspersky Labs**. New AZORult campaign abuses popular VPN service to steal cryptocurrency, Online: https://www.kaspersky.com/about/press-releases/2020_new-azorult-campaign-abuses-popular-vpn-service-to-steal-cryptocurrency.

[147] **Kawaguchi, Y., Yamada, A., and Ozawa, S.** AI Web-Contents Analyzer for Monitoring Underground Marketplace. In *International Conference on Neural Information Processing* (2017), Springer, pp. 888–896.

[148] **Korakakis, M., Magkos, E., and Mylonas, P.** Automated CAPTCHA Solving: An Empirical Comparison of Selected Techniques. In *SMAP* (2014), pp. 44–47.

[149] **Kovacs, E.** South Korea-linked hackers targeted Chinese government via VPN Zero-Day. Online: https://www.securityweek.com/south-korea-linked-hackers-targeted-chinese-government-vpn-zero-day. *SecurityWeek* (Jan 2023).

[150] **Krebs, B.** German Cops Raid 'Cyberbunker 2.0', Arrest 7 in Child Porn, Dark Web Market Sting. Online: https://krebsonsecurity.com/2019/09/german-cops-raid-cyberbunker-2-0-arrest-7-in-child-porn-dark-web-market-sting/. *Krebs on Security* (Sep 2019).

[151] **Krebs, B.** Why were the Russians so set against this hacker being extradited? Online: https://krebsonsecurity.com/2019/11/why-were-the-russians-so-set-against-this-hacker-being-extradited/. *Krebs on Security* (Nov 2019).

[152] **Krebs, B.** Krebs on Security, Online: https://krebsonsecurity.com/2020/03/live-coronavirus-map-used-to-spread-malware/.

[153] **Krebs, B.** Three top Russian cybercrime forums hacked. Online: https://krebsonsecurity.com/2021/03/three-top-russian-cybercrime-forums-hacked/. *Krebs on Security* (Mar 2021).

[154] **Krebs, B.** Breach exposes users of Microleaves Proxy Service, Online: https://krebsonsecurity.com/2022/07/breach-exposes-users-of-microleaves-proxy-service/.

[155] **Kwon, S., Kim, Y.-G., and Cha, S.** Web robot detection based on pattern-matching technique. *Journal of Information Science 38*, 2 (2012), 118–126.

[156] **Kwon, S., Oh, M., Kim, D., Lee, J., Kim, Y.-G., and Cha, S.** Web robot detection based on monotonous behavior. *Information Science and Industrial Applications 4* (2012), 43–48.

[157] **Laferrière, D., and Décary-Hétu, D.** Examining the uncharted dark web: Trust signalling on single vendor shops. *Deviant Behavior 44*, 1 (2023), 37–56.

[158] **Lai, Y.-M., Zheng, X., Chow, K., Hui, L. C., and Yiu, S.-M.** Automatic online monitoring and data-mining internet forums. In *2011 7th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)* (2011), IEEE, pp. 384–387.

[159] **Laliwala, Z., and Shaikh, A.** *Web Crawling and Data Mining with Apache Nutch*. Packt Publishing, 2013.

[160] **Larin, B.** Magnitude exploit kit – evolution. Online: https://securelist.com/magnitude-exploit-kit-evolution/97436/. *Securelist English Global* (May 2021).

[161] **Lê, S., Josse, J., and Husson, F.** FactoMineR: A Package for Multivariate Analysis. *Journal of Statistical Software 25*, 1 (2008), 1–18.

[162] **Leotta, M., Stocco, A., Ricca, F., and Tonella, P.** ROBULA+: An algorithm for generating robust XPath locators for web testing. *Journal of Software: Evolution and Process 28*, 3 (2016), 177–204.

[163] **Leukfeldt, E. R.** Cybercrime and social ties: Phishing in Amsterdam. *Trends in Organized Crime 17* (2014), 231–249.

[164] **Leukfeldt, E. R.** Organised cybercrime and social opportunity structures: A proposal for future research directions. *The European Review of Organised Crime 2*, 2 (2015), 91–103.

[165] **Leukfeldt, E. R., Kleemans, E. R., and Stol, W. P.** A typology of cybercriminal networks: from low-tech all-rounders to high-tech specialists. *Crime, Law and Social Change 67* (2017), 21–37.

[166] **Leukfeldt, R., Kleemans, E., and Stol, W.** The Use of Online Crime Markets by Cybercriminal Networks: A View From Within. Online: `https://doi.org/10.1177/0002764217734267`. *American Behavioral Scientist 61*, 11 (2017), 1387–1402.

[167] **Li, V. G., Dunn, M., Pearce, P., McCoy, D., Voelker, G. M., Savage, S., and Levchenko, K.** Reading the tea leaves: A comparative analysis of threat intelligence. In *28th USENIX Security Symposium (USENIX Security 2019)*.

[168] **Lim, W.-Y., Raja, V., and Thing, V. L.** Generalized and lightweight algorithms for automated web forum content extraction. In *2013 IEEE International Conference on Computational Intelligence and Computing Research* (2013), IEEE, pp. 1–8.

[169] **Lin, X., Ilia, P., Solanki, S., and Polakis, J.** Phish in Sheep's Clothing: Exploring the Authentication Pitfalls of Browser Fingerprinting. In *31st USENIX Security Symposium (USENIX Security 2022)*, pp. 1651–1668.

[170] **Lusthaus, J.** Trust in the world of cybercrime. *Global Crime 13*, 2 (2012), 71–94.

[171] **Lusthaus, J.** Beneath the dark web: Excavating the layers of cybercrime's underground economy. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (2019), IEEE, pp. 474–480.

[172] **Lusthaus, J., and Varese, F.** Offline and local: The hidden face of cybercrime. *Policing: A Journal of Policy and Practice 15*, 1 (2021), 4–14.

[173] **MalwareBytes Threat Intelligence Team**. Multi-stage APT attack drops Cobalt Strike using Malleable C2 feature. Online: `https://blog.malwarebytes.com/threat-analysis/2020/06/multi-stage-apt-attack-drops-cobalt-strike-using-malleable-c2-feature/`. *Malwarebytes* (jun 2020).

[174] **Mandiant**, Online: `https://www.mandiant.com/resources/reports/supply-chain-analysis-quartermaster-sunshop`.

[175] **Marczak, B., Scott-Railton, J., McKune, S., Razzak, B. A., and Deibert, R.** Hide and seek: Tracking NSO group's pegasus spyware to operations in 45 countries. Online: `https://citizenlab.ca/2018/09/hide-and-seek-tracking-nso-groups-pegasus-spyware-to-operations-in-45-countries/`. *The Citizen Lab* (May 2020).

[176] **Martin, J., and Christin, N.** Ethics in cryptomarket research. *International Journal of Drug Policy 35* (2016), 84–91.

[177] **Matt, W.** The scammers who scam scammers on Cybercrime Forums: Part 1. Online: https://news.sophos.com/en-us/2022/12/07/the-scammers-who-scam-scammers-on-cybercrime-forums-part-1/. *Sophos News* (Dec 2022).

[178] **McAfee, Centre for Strategic & International Studies**. Net losses: Estimating the global cost of cybercrime. *McAfee, Centre for Strategic & International Studies* (2014).

[179] **McNair, J.** What is the Average Reading Speed and the Best Rate of Reading? *http://ezinearticles.com/?What-is-the-Average-Reading-Speed-and-the-Best-Rate-of-Reading?&id=2298503* (2009).

[180] **Metrick, K., Najafi, P., and Semrau, J.** Zero-Day Exploitation Increasingly Demonstrates Access to Money, Rather than Skill—Intelligence for Vulnerability Management. Tech. rep., Technical Report, FireEye Technical Report. Available online: https://www.mandiant.com/resources/blog/zero-day-exploitation-demonstrates-access-to-money-not-skill, 2020.

[181] **Milka, G.** Anatomy of Account Takeover. In *Enigma 2018 (Enigma 2018)* (Santa Clara, CA, Jan. 2018), USENIX Association.

[182] **Miller, C.** The Legitimate Vulnerability Market: The Secretive World of 0-Day Exploit Sales. In *6th Workshop on the Economics of Information Security (WEIS 2007)*.

[183] **Minozzi, B.** Block bad bots and stop bad bots crawlers and spiders and anti spam protection, Online: https://wordpress.org/plugins/stopbadbots/.

[184] **Mohr, G., Stack, M., Rnitovic, I., Avery, D., and Kimpton, M.** Introduction to heritrix. In *4th International Web Archiving Workshop* (2004), Citeseer, pp. 109–115.

[185] **Moneva, A., Leukfeldt, E. R., and Klijnsoon, W.** Alerting consciences to reduce cybercrime: a quasi-experimental design using warning banners. *Journal of Experimental Criminology* (2022), 1–28.

[186] **Montalbano, E.** Angry affiliate leaks Conti Ransomware Gang Playbook, Online: https://threatpost.com/affiliate-leaks-conti-ransomware-playbook/168442/.

[187] **Morris, R., and Thompson, K.** Password security: A case history. *Communications of the ACM 22*, 11 (1979), 594–597.

[188] **Motoyama, M., McCoy, D., Levchenko, K., Savage, S., and Voelker, G. M.** An analysis of underground forums. In *2011 ACM SIGCOMM Conference on Internet Measurement Conference* (2011), ACM, pp. 71–80.

[189] **Mulliner, C., Borgaonkar, R., Stewin, P., and Seifert, J.-P.** SMS-based one-time passwords: attacks and defense. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment* (2013), Springer, pp. 150–159.

[190] **National Security Agency**. Defense Language Institute and National Cryptologic School Agreement Helps U.S. Service Personnel Earn Associate Degree. Online: https://www.nsa.gov/Press-Room/News-Highlights/Article/

Article/1874058/defense-language-institute-and-national-cryptologic-school-agreement-helps-us-s/. *National Security Agency/Central Security Service* (Jun 2019).

[191] **Nayak, K., Marino, D., Efstathopoulos, P., and Dumitraş, T.** Some Vulnerabilities Are Different Than Others. In *Proc. of RAID'14*. Springer, 2014, pp. 426–446.

[192] **Noroozian, A., Koenders, J., Van Veldhuizen, E., Ganan, C. H., Alrwais, S., McCoy, D., and Van Eeten, M.** Platforms in everything: analyzing ground-truth data on the anatomy and economics of bullet-proof hosting. In *28th USENIX Security Symposium (USENIX Security 2019)*, pp. 1341–1356.

[193] **Norton**. What is bulletproof hosting? Online: https://us.norton.com/internetsecurity-emerging-threats-what-is-bulletproof-hosting.html. *Emerging Threats*.

[194] **Norton**. Norton Cybercrime Report 2013, Online: https://www.hearst.com/documents/33329/653025/2013+Norton+Report.pdf.

[195] **NSO Group**. Home page of the NSO Group, Online: http://www.nsogroup.com/.

[196] **Nunes, E., Diab, A., Gunn, A., Marin, E., Mishra, V., Paliath, V., Robertson, J., Shakarian, J., Thart, A., and Shakarian, P.** Darknet and deepnet mining for proactive cybersecurity threat intelligence. *arXiv preprint arXiv:1607.08583* (2016).

[197] **Nurmi, J., Niemelä, M., and Brumley, B. B.** Malware Finances and Operations: a Data-Driven Study of the Value Chain for Infections and Compromised Access. In *18th International Conference on Availability, Reliability and Security* (2023), pp. 1–12.

[198] **Oerlemans, J.-J.** *Investigating cybercrime*. PhD thesis, Leiden University, 2017.

[199] **Oest, A., Safei, Y., Doupé, A., Ahn, G.-J., Wardman, B., and Warner, G.** Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis. In *13th APWG Symposium on Electronic Crime Research (APWG eCrime)* (2018), IEEE, pp. 1–12.

[200] **Oguzhantopgul**. Oguzhantopgul/mod_anticrawl: Mod_anticrawl: An anti-crawling module for Apache Web Servers, Online: https://github.com/oguzhantopgul/mod_antiCrawl.

[201] **Onaolapo, J., Mariconti, E., and Stringhini, G.** What happens after you are pwnd: Understanding the use of leaked webmail credentials in the wild. In *2016 Internet Measurement Conference* (2016), pp. 65–79.

[202] **O'Neill, P. H.** Inside NSO, Israel's billion-dollar spyware giant. Online: https://www.technologyreview.com/2020/08/19/1006458/nso-spyware-controversy-pegasus-human-rights/. *MIT Technology Review* (Aug 2020).

[203] **Osborne, C.** FXMSP hacker indicted by Feds for selling backdoor access to hundreds of companies. Online: https://www.zdnet.com/article/

fxmsp-hacker-indicted-by-feds-for-selling-network-access-impacting-hundreds-of-companies/. *ZDNET*.

[204] **Osterwalder, A., and Pigneur, Y.** *Business model generation: a handbook for visionaries, game changers, and challengers*, vol. 1. John Wiley & Sons, 2010.

[205] **Overdorf, R., Troncoso, C., Greenstadt, R., and McCoy, D.** Under the underground: Predicting private interactions in underground forums. *arXiv preprint arXiv:1805.04494* (2018).

[206] **Ozment, A.** The Likelihood of Vulnerability Rediscovery and the Social Utility of Vulnerability Hunting. In *4th Workshop on the Economics of Information Security (WEIS 2005)*.

[207] **Pastrana, S., Hutchings, A., Caines, A., and Buttery, P.** Characterizing eve: Analysing cybercrime actors in a large underground forum. In *International Symposium on Research in Attacks, Intrusions, and Defenses* (2018), Springer, pp. 207–227.

[208] **Pastrana, S., Hutchings, A., Thomas, D., and Tapiador, J.** Measuring ewhoring. In *2019 ACM SIGCOMM Conference on Internet Measurements* (2019), pp. 463–477.

[209] **Pastrana, S., Thomas, D. R., Hutchings, A., and Clayton, R.** Crimebb: Enabling cybercrime research on underground forums at scale. In *27th International Conference on World Wide Web* (2018), ACM, pp. 1845–1854.

[210] **Pearson, G., and Hobbs, D.** King pin? A case study of a middle market drug broker. *The Howard Journal of Criminal Justice 42*, 4 (2003), 335–347.

[211] **Peretti, K. K.** Data Breaches: What the Underground World of Carding Reveals. *Santa Clara Computer High Technology Law Journal 25* (2008), 375.

[212] **Perlroth, N., and Shane, S.** In Baltimore and beyond, a stolen N.S.A. tool wreaks havoc. Online: https://www.nytimes.com/2019/05/25/us/nsa-hacking-tool-baltimore.html. *The New York Times* (May 2019).

[213] **Pete, I., Hughes, J., Caines, A., Vu, A. V., Gupta, H., Hutchings, A., Anderson, R., and Buttery, P.** PostCog: A Tool for Interdisciplinary Research into Underground Forums at Scale. In *2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (2022), IEEE, pp. 93–104.

[214] **Pimenta, B., and Blancas, J.** The Rising Threat of Traffers. Online: https://outpost24.com/resources/the-rising-threat-of-traffers/. *Outpost24* (Aug 2023).

[215] **Piquero, A. R., Paternoster, R., Pogarsky, G., and Loughran, T.** Elaborating the individual difference component in deterrence theory. *Annual Review of Law and Social Science 7* (2011), 335–360.

[216] **Ponemon Institute LLC**. The Value of Threat Intelligence: Annual Study of North American & United Kingdom Companies, Online: https://stratejm.com/wp-content/uploads/2019/08/2019_Ponemon_Institute-Value_of_Threat_Intelligence_Research_Report_from_Anomali.pdf.

[217] **Portnoff, R. S., Afroz, S., Durrett, G., Kummerfeld, J. K., Berg-Kirkpatrick, T., McCoy, D., Levchenko, K., and Paxson, V.** Tools for Automated Analysis of Cybercriminal Markets. In *26th International Conference on World Wide Web* (2017), International World Wide Web Conferences Steering Committee, pp. 657–666.

[218] **ProWebScraper**. Top 10 Captcha Solving Services Compared, Online: `http://www.prowebscraper.com/blog/top-10-captcha-solving-services-compared/`.

[219] **Qassrawi, M. T., and Zhang, H.** Client honeypots: Approaches and challenges. In *New Trends in Information Science and Service Science (NISS), 2010 4th International Conference* (2010), IEEE, pp. 19–25.

[220] **Research, C. P.** Leaks of Conti Ransomware Group Paint Picture of a surprisingly normal tech start-up... sort of, Online: `https://research.checkpoint.com/2022/leaks-of-conti-ransomware-group-paint-picture-of-a-surprisingly-normal-tech-start-up-sort-of/`.

[221] **Robot, S.** Web crawler detected: How to scrape under the Radar, Online: `https://scrapingrobot.com/blog/crawler-detected/`.

[222] **Sabillon, R., Cavaller, V., Cano, J., and Serra-Ruiz, J.** Cybercriminals, cyberattacks and cybercrime. In *2016 IEEE International Conference on Cybercrime and Computer Forensic (ICCCF)* (2016), IEEE, pp. 1–9.

[223] **Santanna, J. J., van Rijswijk-Deij, R., Hofstede, R., Sperotto, A., Wierbosch, M., Granville, L. Z., and Pras, A.** Booters—An analysis of DDoS-as-a-service attacks. In *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)* (2015), IEEE, pp. 243–251.

[224] **Sardar, T. H., and Ansari, Z.** Detection and confirmation of web robot requests for cleaning the voluminous web log data. In *IMpact of E-Technology on US (IMPETUS), 2014 International Conference on the* (2014), IEEE, pp. 13–19.

[225] **Schäfer, M., Fuchs, M., Strohmeier, M., Engel, M., Liechti, M., and Lenders, V.** BlackWidow: Monitoring the dark web for cyber security information. In *2019 11th International Conference on Cyber Conflict (CyCon)* (2019), vol. 900, IEEE, pp. 1–21.

[226] **Schectman, J., and Bing, C.** Special report: White House Veterans helped Gulf Monarchy Build Secret Surveillance Unit. Online: `https://www.reuters.com/article/us-usa-raven-whitehouse-specialreport/special-report-white-house-veterans-helped-gulf-monarchy-build-secret-surveillance-unit-idUSKBN1YE1OB`. *Reuters* (Dec 2019).

[227] **Schelling, T. C.** *The Strategy of Conflict: with a new Preface by the Author*. Harvard university press, 1980.

[228] **Scrapy**. A fast and powerful scraping and web crawling framework, Online: `https://scrapy.org/`.

[229] **Scrapy**. Avoiding getting banned, Online: https://docs.scrapy.org/en/latest/topics/practices.html#bans.

[230] **Security Tools**. Open-source Red Team Tools. Online: https://github.com/topics/security-tools?q=red%2Bteam&unscoped_q=red%2Bteam. *GitHub*.

[231] **SecurityWeek**. Zerodium Discloses Flaw That Allows Code Execution in Tor Browser, Online: https://www.securityweek.com/zerodium-discloses-flaw-allows-code-execution-tor-browser.

[232] **Segal, A.** The Code not Taken: China, the United States, and the Future of Cyber Espionage. *Bulletin of the Atomic Scientists 69*, 5 (2013), 38–45.

[233] **Shah, A., Ganesan, R., Jajodia, S., and Cam, H.** Understanding tradeoffs between throughput, quality, and cost of alert analysis in a CSOC. *IEEE Transactions on Information Forensics and Security 14*, 5 (2018), 1155–1170.

[234] **Shirey, R. W.** RFC 4949: Internet security glossary, version 2. Online: https://tools.ietf.org/html/rfc4949. *IETF Datatracker* (Aug 2007).

[235] **Sinigaglia, F., Carbone, R., Costa, G., and Zannone, N.** A survey on multi-factor authentication for online banking in the wild. *Computers & Security 95* (2020), 101745.

[236] **SOCRadar**. Discord: The new playground for Cybercriminals, Online: https://socradar.io/discord-the-new-playground-for-cybercriminals/.

[237] **Sood, A. K., and Enbody, R. J.** Crimeware-as-a-service—A survey of commoditized crimeware in the underground market. Online: https://www.sciencedirect.com/science/article/pii/S1874548213000036. *International Journal of Critical Infrastructure Protection 6*, 1 (2013), 28–38.

[238] **Soska, K., and Christin, N.** Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem. In *24th USENIX Security Symposium (USENIX Security 2015)* (2015), pp. 33–48.

[239] **Soudijn, M. R., and Zegers, B. C. T.** Cybercrime and Virtual Offender Convergence Settings. *Trends in Organized Crime 15*, 2-3 (2012), 111–129.

[240] **Stevanovic, D., An, A., and Vlajic, N.** Feature evaluation for web crawler detection with data mining techniques. *Expert Systems with Applications 39*, 10 (2012), 8707–8717.

[241] **Stobert, E.** The agony of passwords: Can we learn from user coping strategies? In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM New York, NY, USA, 2014, pp. 975–980.

[242] **Stringhini, G., and Thonnard, O.** That ain't you: Blocking spearphishing through behavioral modelling. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment* (2015), Springer, pp. 78–97.

[243] **Tabarrok, A., and Cowen, T.** The end of asymmetric information. *Cato Unbound 6* (2015).

[244] **The Hacker News**. Firefox Zero-Day Exploit used by FBI to shutdown Child porn on Tor Network hosting; Tor Mail Compromised, Online: https://thehackernews.com/2013/08/Firefox-Exploit-Tor-Network-child-pornography-Freedom-Hosting.html.

[245] **Thomas, K., Amira, R., Ben-Yoash, A., Folger, O., Hardon, A., Berger, A., Bursztein, E., and Bailey, M.** The abuse sharing economy: Understanding the limits of threat exchanges. In *International Symposium on Research in Attacks, Intrusions, and Defenses (RAID)* (2016), Springer, pp. 143–164.

[246] **Thomas, K., Huang, D., Wang, D., Bursztein, E., Grier, C., Holt, T. J., Kruegel, C., McCoy, D., Savage, S., and Vigna, G.** Framing Dependencies Introduced by Underground Commoditization. In *14th Workshop on the Economics of Information Security (WEIS 2015)*.

[247] **Thomas, K., Li, F., Grier, C., and Paxson, V.** Consequences of connectivity: Characterizing account hijacking on twitter. In *2014 ACM SIGSAC Conference on Computer and Communications Security (CCS)* (2014), pp. 489–500.

[248] **Thomas, K., Li, F., Zand, A., Barrett, J., Ranieri, J., Invernizzi, L., Markov, Y., Comanescu, O., Eranti, V., Moscicki, A., et al.** Data breaches, phishing, or malware? Understanding the risks of stolen credentials. In *2017 ACM SIGSAC Conference on Computer and Communications Security (CCS)* (2017), pp. 1421–1434.

[249] **Thomas, R., and Martin, J.** The Underground Economy: Priceless. *; login:: the magazine of USENIX & SAGE 31*, 6 (2006), 7–16.

[250] **Thomas, W.** Mercenary APTs – An Exploration. Online: https://www.cyjax.com/mercenary-apts-an-exploration/. *CYJAX* (Oct 2021).

[251] **Tor Project**. The Tor Project: Privacy and Freedom Online, Online: https://www.torproject.org/.

[252] **Trend Micro**. Russian underground 101. *Cupertino, CA: Trend Micro* (2012).

[253] **Trend Micro**. North Korean hackers allegedly exploit Adobe Flash Player vulnerability (CVE-2018-4878) against South Korean targets. Online: https://www.trendmicro.com/vinfo/us/security/news/vulnerabilities-and-exploits/north-korean-hackers-allegedly-exploit-adobe-flash-player-vulnerability-cve-2018-4878-against-south-korean-targets. *Trend Micro - Security News* (feb 2020).

[254] **TrustWave**. Messing with AZORult Part 1: Malware Breakdown, Online: https://www.trustwave.com/en-us/resources/blogs/spiderlabs-blog/messing-with-azorult-part-1-malware-breakdown/.

[255] **Tunnell, K. D.** Choosing crime: Close your eyes and take your chances. *Justice Quarterly 7*, 4 (1990), 673–690.

[256] **Tunnell, K. D.** Choosing crime: The criminal calculus of property offenders. *The American Journal of Sociology 98* (1993), 1497–1499.

[257] **Turk, K., Pastrana, S., and Collier, B.** A tight scrape: Methodological approaches to cybercrime research data collection in adversarial environments. In *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)* (2020), IEEE, pp. 428–437.

[258] **Van de Bunt, H. G., Kleemans, E. R., De Poot, C., Bokhorst, R., Huikeshoven, M., Kouwenberg, R., Nassou, M. v., and Staring, R.** Georganiseerde criminaliteit in Nederland.

[259] **van de Laarschot, J., and van Wegberg, R.** Risky Business? Investigating the Security Practices of Vendors on an Online Anonymous Market using Ground-Truth Data. In *30th USENIX Security Symposium (USENIX Security 2021)*, pp. 4079–4095.

[260] **Van Hardeveld, G. J., Webber, C., and O'Hara, K.** Deviating from the cybercriminal script: Exploring tools of anonymity (mis) used by carders on cryptomarkets. *American Behavioral Scientist 61*, 11 (2017), 1244–1266.

[261] **van Wegberg, R., Miedema, F., Akyazi, U., Noroozian, A., Klievink, B., and van Eeten, M.** Go see a specialist? Predicting cybercrime sales on online anonymous markets from vendor and product characteristics. In *29th International Conference on World Wide Web* (2020), pp. 816–826.

[262] **Van Wegberg, R., Tajalizadehkhoob, S., Soska, K., Akyazi, U., Ganan, C. H., Klievink, B., Christin, N., and Van Eeten, M.** Plug and Prey? Measuring the Commoditization of Cybercrime via Online Anonymous Markets. In *27th USENIX Security Symposium (USENIX Security 2018)* (2018), pp. 1009–1026.

[263] **Von Ahn, L., Blum, M., Hopper, N. J., and Langford, J.** CAPTCHA: Using hard AI problems for security. In *International Conference on the Theory and Applications of Cryptographic Techniques* (2003), Springer, pp. 294–311.

[264] **Vx-Underground**. The staff of XSS appear to be mildly frustrated with threat intelligence companies scraping their forum.they are now allowing companies the ability to scrape the forum for an annual fee of $2,000. pic.twitter.com/pffkhxlnfr, Online: `https://twitter.com/vxunderground/status/1585368524901748736?s=20`.

[265] **wallas.anthony**. Scam report – 0day.Today (INJ3CT0R), Online: `https://bitcointalk.org/index.php?topic=1611835.0`.

[266] **Weulen Kranenbarg, M.** When do they offend together? Comparing co-offending between different types of cyber-offenses and traditional offenses. *Computers in Human Behavior 130* (2022), 107186.

[267] **Weulen Kranenbarg, M., Van Gelder, J.-L., Barends, A. J., and de Vries, R. E.** Is there a cybercriminal personality? Comparing cyber offenders and offline offenders on

HEXACO personality domains and their underlying facets. *Computers in Human Behavior 140* (2023), 107576.

[268] **Wiefling, S., Iacono, L. L., and Dürmuth, M.** Is this really you? An empirical study on risk-based authentication applied in the wild. In *IFIP International Conference on ICT Systems Security and Privacy Protection* (2019), Springer, pp. 134–148.

[269] **Wikipedia contributors**. Democratization of technology — Wikipedia, The Free Encyclopedia, Online: `https://w.wiki/7k3k`. Accessed 5 October 2023.

[270] **Wikipedia contributors**. Phrases from The Hitchhiker's Guide to the Galaxy: The Answer to the Ultimate Question of Life, the Universe, and Everything is 42 — Wikipedia, The Free Encyclopedia, Online: `https://w.wiki/7k2k`. [accessed 11 October 2023].

[271] **Wittes, B., Poplin, C., Jurecic, Q., and Spera, C.** Sextortion: Cybersecurity, teenagers, and remote sexual assault. *Center for Technology at Brookings* (2016).

[272] **Woods, D. W., and Böhme, R.** SoK: Quantifying cyber risk. In *2021 IEEE Symposium on Security and Privacy (S&P)* (2021), IEEE, pp. 211–228.

[273] **Yan, Q., Han, J., Li, Y., DENG, H., et al.** On limitations of designing usable leakage-resilient password systems: Attacks, principles and usability. In *19th Network and Distributed System Security Symposium (NDSS)* (2012).

[274] **Yip, M., Shadbolt, N., and Webber, C.** Why forums? An empirical analysis into the facilitating factors of carding forums. In *5th Annual ACM Web Science Conference* (2013), pp. 453–462.

[275] **Yip, M., Webber, C., and Shadbolt, N.** Trust among cybercriminals? Carding forums, uncertainty and implications for policing. *Policing and Society 23*, 4 (2013), 516–539.

[276] **Zhang, D., Zhang, D., and Liu, X.** A novel malicious web crawler detector: Performance and evaluation. *International Journal of Computer Science Issues (IJCSI) 10*, 1 (2013), 121.

[277] **Zhang, Y., Fan, Y., Hou, S., Liu, J., Ye, Y., and Bourlai, T.** iDetector: Automate Underground Forum Analysis Based on Heterogeneous Information Network. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (2018), IEEE, pp. 1071–1078.

[278] **Zhang, Y., Fan, Y., Ye, Y., Zhao, L., and Shi, C.** Key player identification in underground forums over attributed heterogeneous information network embedding framework. In *28th ACM International Conference on Information and Knowledge Management* (2019), pp. 549–558.

[279] **Zhao, Q., Willemsen, M. C., Adomavicius, G., Harper, F. M., and Konstan, J. A.** Interpreting User Inaction in Recommender Systems. In *12th ACM Conference on Recommender Systems* (New York, NY, USA, 2018), RecSys '18, ACM, pp. 40–48.

[280] **Zhuge, J., Holz, T., Song, C., Guo, J., Han, X., and Zou, W.** Studying malicious websites and the underground economy on the Chinese web. In *Managing Information Risk and the Economics of Security*. Springer, 2009, pp. 225–244.

Curriculum Vitæ

Michele, originally from Castellaneta, Italy, earned both his BSc (2016) and MSc cum laude (2019) in Computer Engineering from the University of Bologna. His master's project at Eindhoven University of Technology resulted in a scientific publication presented at the WACCO workshop, held at the IEEE European Symposium on Security and Privacy 2019 (Euro S&P). In 2019, he began his PhD at Eindhoven under the supervision of dr. Luca Allodi, focusing on the cybercriminal underground ecosystem's impact on the threat landscape, with an emphasis on emerging threats and threat modeling.

His research gained industry and media outlets attention; furthermore, he contributed to a report on the NSO Group for the U.S. Department of Commerce, and his research on emerging threats has been instrumental to the 'Cookie Monster' investigation by the Dutch National Police and EUROPOL, leading to global arrests and the shutdown of Genesis Market. Michele co-organized Ei/PSi seminars, supervised numerous students and research projects, and led lab activities for the Offensive Computer Security course during his PhD.