Browser Anti-Fingerprinting as a pathway to Impersonation and Fraud.

Presented by

O.N.R. Haalstra (Olaf)

Supervisor	:	Dr. A. Sperotto (Anna)
Co-supervisor	:	Dr. Ir. A. Continella (Andrea)
Company-supervisor	:	Ir. J. van Lenthe (Jarmo)

UNIVERSITY OF TWENTE.

POLITIE

O.N.R. Haalstra (Olaf): *Browser Anti-Fingerprinting as a pathway to Impersonation and Fraud.*, © 25. August 2020

ABSTRACT

Browser fingerprinting can be used to uniquely identify users with very high accuracy. This is why companies, like banks, online merchandise and payment providers are using this additionally to passwords to protect users against fraud by cybercriminals. However, cybercriminals have found ways to circumvent these checks by utilising so called anti-fingerprinting browsers. With anti-fingerprinting browsers cybercriminals can circumvent browser fingerprinting and impersonate their victim effectively. In this study the market that revolves around these anti-fingerprinting browsers is studied. Stolen digital identities that can be used to configure anti-fingerprinting browsers are bought and subjected to a testing system that analyses the fingerprint. This leads to a proposed defence mechanism which increases the fingerprinted surface and shows that none of the anti-fingerprinting browsers were able to 100% accurately spoof another browsers and therefore fail to correctly impersonate one's digital identity. This study shows that anti-fingerprinting browsers pose a serious threat within the cybercriminal landscape, but that this is not insurmountable as results have shown that a defence mechanism against browser anti-fingerprinting can be constructed.

It does not do harm to the mystery to know a little about it.

— Richard Feynman [14]
Theoretical physicist and
Nobel Prize winner

ACKNOWLEDGMENTS

Ever since March 2020 extraordinary times have begun during which I finished most, if not all, of the writing for my master thesis. This period is of course known for the lack of social interactions and therefore also less distractions than ever before. I am grateful for having finished it, especially given that writing a thesis can be horrible at times. Nevertheless, I enjoyed performing the research at all times and learned a lot about the every day threats we face within the cyber security landscape, a challenge which I am happy to dedicate my career to.

First of all, I would like to thank Jarmo van Lenthe for always being in a good mood, providing me this research opportunity and helping me with all the red tape I was presented with at the start of my internship. Even though your colleagues seem to think they all have right to some of your valuable time, you were always up for a chat, a brainstorm session or a technical question.

Secondly, I would like to thank Anna Sperotto who really got me to the level of quality the thesis is as presented, you hopefully agree with me: is great. With your feedback I was always able to move forward, and your questions always pushed me to explain myself better: also in text.

Next, I would like to thank Jeroen van der Ham for critically assessing my ethical analysis and unintentionally giving me some last minute, post green light, stress. Which was of course very helpful in improving the piece. Additionally, I would like to thank Andrea Continella to relieve some of this stress by becoming part of the graduation committee upon request.

Finally, I would like to thank my parents who never once doubted the career choices I made. They gave me all the opportunities that led up to this day and have a very relaxed attitude towards the student life I fully enjoyed the past 7 years.

Since the current crisis forced me to work solely from home, I would like to express a sincere gratitude towards my girlfriend and the ladies of *Huize Pluim* which tolerated my stay for a prolonged period of time and lent me a desk in exchange for wine and Tony's Chocolonely.

CONTENTS

I THESIS

1 INTRODUCTION 2
- 1.1 Motivation . . . 2
- 1.2 Objectives . . . 2
- 1.3 Contribution . . . 3
- 1.4 Structure . . . 3

2 RELATED WORK 4
- 2.1 Terminology . . . 4
  - 2.1.1 Fingerprinting . . . 4
  - 2.1.2 Anti-Fingerprinting . . . 4
- 2.2 Fingerprinting . . . 4
  - 2.2.1 Application . . . 5
  - 2.2.2 Tracking techniques . . . 5
  - 2.2.3 Countermeasures . . . 6
  - 2.2.4 Limitations . . . 6
  - 2.2.5 Overview . . . 8
- 2.3 Malicious intent . . . 8
- 2.4 Summary . . . 11

3 METHOD 12
- 3.1 Approach . . . 12
- 3.2 Objectives . . . 12
  - 3.2.1 State of the Market . . . 12
  - 3.2.2 Threat Analysis . . . 12
  - 3.2.3 Anti Fingerprinting . . . 13
  - 3.2.4 Protection . . . 13

4 STATE OF THE MARKET 14
- 4.1 Overview . . . 14
- 4.2 Products . . . 14
  - 4.2.1 Categories . . . 14
- 4.3 Landscape . . . 16
- 4.4 Discussion . . . 17
  - 4.4.1 Limitations . . . 17
- 4.5 Conclusion . . . 17

5 THREAT ANALYSIS 19
- 5.1 Scope . . . 19
- 5.2 Genesis Market . . . 19
  - 5.2.1 Analysis . . . 19
  - 5.2.2 Caveat . . . 20
  - 5.2.3 Digital Identity . . . 21
  - 5.2.4 Offerings . . . 21
  - 5.2.5 Collection of digital identities . . . 21

5.2.6 Maturity of service . . . 23
5.2.7 Software . . . 24
5.3 Threat for the Netherlands . . . 24
5.4 Impact . . . 28
5.5 Discussion . . . 29
5.5.1 Limitations . . . 29
5.5.2 Suggestions . . . 30
5.6 Conclusion . . . 30
6 ANTI FINGERPRINTING 31
6.1 Data acquisition . . . 31
6.1.1 Ethical implications . . . 31
6.1.2 Legal implications . . . 34
6.1.3 Ethical & Legal Conclusion . . . 35
6.2 Analysis . . . 35
6.2.1 Fingerprinting . . . 36
6.3 Results . . . 36
6.3.1 Fingerprint attributes . . . 36
6.3.2 Groups . . . 38
6.3.3 Subgroups . . . 38
6.4 Discussion . . . 40
6.4.1 Limitations . . . 40
6.4.2 Suggestions . . . 40
6.5 Conclusion . . . 41
7 PROTECTION 42
7.1 Motivation . . . 42
7.2 Approach . . . 42
7.2.1 Feature support . . . 42
7.2.2 Sequences . . . 43
7.2.3 Similarity scoring . . . 43
7.3 Results . . . 44
7.3.1 Features . . . 44
7.3.2 Verify Reported feature . . . 45
7.3.3 Reported features pairwise scored on similarity . . . 45
7.3.4 Browser comparison . . . 48
7.3.5 Tests pairwise comparison . . . 48
7.4 Discussion . . . 52
7.4.1 Limitations . . . 52
7.4.2 Suggestions . . . 54
7.5 Conclusion . . . 54
8 CONCLUSION 56
8.1 Summary . . . 56
8.1.1 State of the Market . . . 56
8.1.2 Threat Analysis . . . 56
8.1.3 Anti Fingerprinting . . . 57
8.1.4 Protection . . . 57
8.2 Limitations . . . 57

8.3 Future Suggestions 59

II APPENDIX

A EXAMPLES OF PRODUCTS AVAILABLE ON THE MARKET 61

B GENESIS MARKET ADVERTISEMENT 63

B.1 Translation 63

C FULL LIST OF FINGERPRINTED ATTRIBUTES 65

D FULL LIST OF FEATURE SUPPORT 66

E ADDITIONAL CORRELATION FIGURES 68

BIBLIOGRAPHY 71

LIST OF FIGURES

Figure 4.1 Fingerprinting landscape 15
Figure 5.1 Overview of number of digital identities on the Genesis Market as reported by various sources over time. . 20
Figure 5.2 Overview of the *Bots* (profiles) page where all digital identities can be filtered. 22
Figure 5.3 Detailed information about a profile before purchase. Displays whether there are fingerprints and cookies available, time of malware infection, time of latest information update, operating system, country and a IP 16 prefix. 22
Figure 5.4 Detailed information about which sites are available for a profile. Displays the full URLs including the availability of a username and password if available, which will become visible after purchase. 23
Figure 5.5 Screenshot of a news article on the Genesis Market included with some of the user feedback on this article, image was taken in August 2020. 24
Figure 5.6 YouTube instruction video how the Genesis software can be used in order to load profile information into the browser in order to impersonate the victim. 25
Figure 5.7 Number of profiles originating from various countries [43], state of February 2020. 28
Figure 5.8 User feedback on the Genesis Market showing that not all users succeed in flawless impersonation, while other users offer their help to use the service better. . . 29
Figure 6.1 List of attributes that are fingerprinted by FingerprintJS2, each number indicates which *subgroup* it belongs to read from left to right. If the number is the same for multiple browsers it means the reported value is the same. 39
Figure 7.1 Sequences of supported features per browser 46
Figure 7.2 List of correctness of reported feature support per browser 47

Figure 7.3 Correlation matrix with 30 browser test scored on similarity with 103 most recent browsers from the "Can I Use" database. Each square displays the similarity between those, the brighter the color the higher the similarity score is. To calculate the similarity score the *Weighted Hamming Distance* is used. Particular interest is taken in the fact that Group 5 and 6 (ID 0 - 11) show that all versions (regardless of their reported version), resemble Chrome 77 closer than their *identified browser*. 49

Figure 7.4 Correlation matrix which displays how similar specific browsers and versions are based on their feature support. The 103 browser from the "Can I Use" database have been pairwise scored on similarity. This provides a baseline on which browser versions are the most similar. Identical to Figure 7.3 the similarity score is calculated using the *Hamming Distance*. It can be observed that especially the latest versions of Edge, Chrome and Opera show a high resemblance. . . 50

Figure 7.5 Correlation matrix which displays the similarities scores from the 30 test results compared pairwise. The *Weighted Hamming Distance* similarity scoring is used. Special attention can be focused on the column with ID 16 (Chrome 77) and rows with ID 0-11 (Group 5 and 6), this seems to indicate a high similarity with Chrome 77 (Group 1). 51

Figure A.1 Linken Sphere promotes ability to investigate antifraud systems. 61

Figure A.2 Marketed at $100 per month, with no up front payment. 62

Figure A.3 Promotes the ability to 'Print your own money', marketed at $100 per month with an $2999 up front payment. 62

Figure E.1 Correlation matrix which displays the similarities scores from the 30 test results compared pairwise. The *Weighted Hamming Distance* similarity scoring is used. Special attention can be focused on the column with ID 16 (Chrome 77) and rows with ID 0-11 (Group 5 and 6), this seems to indicate a high similarity with Chrome 77 (Group 1). However, the results from Group 4 'pollute' the results. The similarity described above is still visible however it is less clear than in Figure 7.5. 68

Figure E.2 Correlation matrix with 30 browser test scored on similarity with 103 most recent browsers from the "Can I Use" database. Each square displays the similarity between those, the brighter the color the higher the similarity score is. To calculate the similarity score the *Hamming Distance* is used. Particular interest is taken in the fact that Group 5 and 6 (ID 0 - 11) show that all versions (regardless of their reported version), resemble Chrome 77 closer than their *identified browser*. However, due to the fact that the unweighted distance method is used here this resemblance is visually less strong than it is in Figure 7.3. 69

Figure E.3 Correlation matrix which displays the similarities scores from the 30 test results compared pairwise. The *Unweighted Hamming Distance* similarity scoring is used. Special attention can be focused on the column with ID 16 (Chrome 77) and rows with ID 0-11 (Group 5 and 6), this seems to indicate a higher similarity with Chrome 77 (Group 1). However, due to the fact that the unweighted distance method is used here this resemblance is visually less strong than it is in Figure 7.5. 70

LIST OF TABLES

Table 2.1 Examples of legitimate and illegitimate use of (anti) fingerprinting. 5
Table 2.2 Comparison of attributes used by various fingerprinting libraries and the capability of anti-fingerprinting products to block or spoof them. 9
Table 2.3 Overview of fingerprinters and countermeasures with their respective sources for Table 2.2 10
Table 4.1 Overview of anti-fingerprinting products 18
Table 5.1 Overview of available login portals on the Genesis Market from vital companies of the Netherlands 26
Table 5.2 Overview of available subdomain names on the Genesis Market 27
Table 6.1 Overview of browsers subjected to the testing system . 37
Table 6.2 Subgrouping process 39
Table 7.1 Browsers categorised into groups and subgroups. Belonging to the same subgroups means that the compatibility sequence from Figure 7.1, is exactly the same for those browsers. 53

ACRONYMS

PII Personal Identifiable Information

NHTCU National High Tech Crime Unit

VM Virtual Machine

VPN Virtual Private Network

GUID Global Unique IDentifier

2FA Two Factor Authentication

Part I

THESIS

1

INTRODUCTION

It is possible to uniquely identify a web visitor with 99.1% accuracy utilising browser fingerprinting [11]. Browser fingerprinting is used by companies such as: banks, online merchandise and payment providers. It is used to enhance the security and verify whether the user that tries to authenticate is really the correct user, and not a cybercriminal with stolen credentials. In addition, browser fingerprinting is also used to track users across the web [49]. This technique became increasingly more accurate, since the fingerprintable surface has expanded over time because of increasing browser functionality [37]. This lead to privacy-aware solutions that try to combat browser fingerprinting by either spoofing or blocking attributes. However, not only privacy-minded started to implement anti-fingerprinting measures: also cybercriminals started to develop this technology, not to prevent them from being uniquely identified, but rather to identify as someone else.

Because browser fingerprinting can be utilised to uniquely identify individuals, companies started using this as a protection measure against fraud [55]. Cybercriminals invented ways to spoof the browser fingerprint in order to impersonate a victim and being able to commit fraud on their behalf, this phenomenon which is the illegitimate application of anti-fingerprinting is studied and described in this thesis.

1.1 MOTIVATION

This study was performed because of recent developments, where payed anti-fingerprinting products have surfaced. The first articles about anti-fingerprinting products date back to 2015 [23] [44] , and were reported again in 2017 [31]. As of mid 2019 this phenomenon was more and more picked up by others [3] [9] [22] [36] [38] [42] [50]. While writing this thesis more and more articles surfaced on the same topic [8] [20] [21] [27] [32] [43] [57]. These products at first, were allegedly used by cybercriminals for protecting their identity. However, later it became more and more apparent that these products were used for digital impersonation to enable fraud. The most discussed product is the *Genesis Market*, which is first described in Section 5.2.

1.2 OBJECTIVES

The goal of this study is to establish whether anti-fingerprinting products pose a serious threat within the cyber criminal landscape.

To achieve this goal, four research questions have been identified and are formulated as follows:

- What is the current state of the anti-fingerprinting market, which products are offered and how do they compare?
- How large is the upcoming threat of anti-fingerprinting? Does it enable mass cybercrime or targeted attacks?
- How do anti-fingerprinting solutions successfully circumvent browser fingerprinting?
- How can we achieve protection from malicious use of anti-fingerprinting?

Ultimately the main research question: 'To what extent do anti-fingerprinting products pose a threat within the cyber criminal landscape?', is answered.

The method and approach for each objective is outlined in Chapter 3.

1.3 CONTRIBUTION

The contribution of this research is twofold. Firstly, the current threat landscape of anti-fingerprinting is extensively researched. This lead to insights that are utilised by the Dutch police. Secondly, a new way to detect browser spoofing is proposed which could be used to defend against browser which utilise anti-fingerprint technology.

1.4 STRUCTURE

After this chapter, Chapter 2 explores the current landscape of browser fingerprinting with the associated countermeasures based upon scholarly articles. Chapter 3 introduces the method divided into four objectives which correspond to the research questions above. These four objectives are subsequently introduced, discussed and concluded in the following four chapters. Chapter 4 describes the current state of the anti-fingerprinting market with its available products. Chapter 5 takes a deep dive into one of the available products and quantifies the available goods. Chapter 6 describes the analysis of the browser fingerprint that are produced by anti-fingerprinting browsers and how these relate to normal browsers. In Chapter 7 an analysis of the feature support of browsers is analysed and tested with normal and anti-fingerprinting browsers similar to the previous chapter. Although each of these chapters have a conclusion, in Chapter 8 the conclusions and limitations are summarised, the main research question is answered and there is briefly touched upon future suggestions.

2

RELATED WORK

In this chapter the preliminary literature review that was used to prepare this research will be discussed. In order to gain a better understanding of browser (anti-)fingerprinting current landscape and application has been researched [18].

2.1 TERMINOLOGY

In this chapter legitimate use and illegitimate use of both fingerprinting and anti-fingerprinting are considered, for examples of each category see Table 2.1.

2.1.1 *Fingerprinting*

First there is the legitimate application of fingerprinting, this aims to prevent digital impersonation of users in order to prevent fraud.

Secondly there is the illegitimate use of fingerprinting, which aims to track users across the web and hurting their privacy. The first two techniques are described in Section 2.2.1.

2.1.2 *Anti-Fingerprinting*

Then there is the legitimate application of anti-fingerprinting, aimed towards prevention of the illegitimate application of fingerprinting. This is for privacy preservation of the user and to prevent cross domain tracking, this is described in Section 2.2.3.

Finally there is the illegitimate application of anti-fingerprinting, this aims to defeat the legitimate application of fingerprinting. Namely, to bypass anti-fraud systems which utilise browser fingerprinting. This allows cybercriminals to digitally impersonate a user, in order to commit fraud on their behalf. This is seen as malicious intent, and is described in Section 2.3.

2.2 FINGERPRINTING

Browser fingerprinting is a widely used technique to uniquely identify web user and to track their online behaviour [5]. Since browsers are critical applications for most end-users because this provides access to online web applications, this encourages advertisers to continuously track user sessions for profiling purposes. Cookies remain the most widespread technique for tracking, however the use of cookies is limited by regulations in Europe and

	Fingerprinting	Anti Fingerprinting
Legitimate use	Anti-fraud systems, performance monitoring, targeted advertisement, related content suggestion	Privacy preservation, multi accounts management
Illegitimate use	Cross domain tracking	Bypassing fingerprinting systems by impersonation to commit fraud

Table 2.1: Examples of legitimate and illegitimate use of (anti) fingerprinting.

the United States [2]. To sidestep these limitations researchers came up with the complementary technique browser fingerprinting [55].

2.2.1 *Application*

Companies are applying tracking for various reasons, such as fraud prevention by identifying illegitimate usage attempts, suggesting related content and better targeting advertisements. There are two types of tracking: *regular tracking* and *third-party tracking*. Regular tracking is confined to the tracker's own website, third party tracking is tracking the user across the web [49]. Torres [49] argues that enabling cross-domain tracking is of little benefit to the user, while negatively impacting the user's privacy. While tracking based on cookies is easily detectable: they can be inspected and deleted. In contrast browser fingerprinting is harder to detect and even hard to opt-out. It works just as well in the "private-mode" of modern browsers, and being able to uniquely identified by browser fingerprinting implies that also users without an account can be tracked. It has been shown that third-party tracking based on fingerprinting is widespread, 50% of the top-100.000 websites contain resilient third-party tracking [19]. Not only is it possible to uniquely identify many users within a large dataset [11], it is also possible to track users for a prolonged period of time [54].

2.2.2 *Tracking techniques*

For many business it is important to limit the impact on the performance, while creating a browser fingerprint. In order to construct a valuable fingerprint, it is important to collect as many independent properties as possible [25]. Popular browser properties to collect are, including, but not limited to, cookies, canvas and plugins. An overview of used techniques can be found in Table 2.2, this list of attributes that can be used in fingerprints is rapidly growing [10]. Most attributes can be extracted via JavaScript but in some cases it is also possible to collect attributes for fingerprinting via, for example, TCP/IP stack, CSS, HTML5 [5]. In many cases the fingerprint is created

by concatenating a string of multiple properties [60] [25] and in some cases hashed [6] [24]. For more advanced cases such as audio and canvas fingerprinting the output is always hashed [12]. Fingerprinting services are offered by companies to others, either embedded by the visited site, or by a third party via advertisements.

2.2.3 *Countermeasures*

Several existing products directly aim to stop trackers, including commercially developed plugins (e.g. Ghostery, uBlock, etc), and academically developed plugins (e.g. DCB [5], Privaricator [34], etc.), browser (e.g. Firefox, Tor). While these plugins are mainly focused towards the preservation of privacy, there are also commercially available products (e.g. Anti Detect, see Figure A.3 and Linken Sphere, see Figure A.1) which are focused towards privacy and multi account management. While the privacy preservation and multi account management are both legitimate uses of anti-fingerprinting. The latter products also enable the evasion of anti fraud systems and impersonation, which allows for illegitimate use of anti-fingerprinting product which enables cyber crime.

All of the solutions named above work by either blocking or spoofing certain attributes to hinder the tracking thereof. For a full overview see Table 2.2. Both techniques have their limitations, since neither are sufficient to prevent tracking and both impact user experience [49]. In addition, it can be expected that parties that provide fingerprinting services will try to use new forms of tracking that evade the blocking or spoofing. This means that both techniques need to be continuously updated in order to achieve a complete fingerprinting countermeasure. Paradoxically, anti-fingerprinting privacy technologies can be self-defeating if they are not used by a sufficient number of people [11].

2.2.4 *Limitations*

Over time several countermeasures have emerged, finding an absolute approach that can prevent fingerprinting while maintaining the richness of the modern browser is a challenging task. The following recurring problems can be found [5]:

1. Protection is only implemented for a subset of attributes.
2. Randomisation causes unrealistic parameters increasing detection.
3. Functionality is altered severely, limiting usability.
4. Non deterministic behaviour is introduced which can be detected by fingerprinting twice.

The following defence strategies can be classified [17], each with their limitations:

BLOCKING FINGERPRINTING *Limitations: 1,3* A straightforward solution is to block access to certain attributes or block certain techniques in order to reduce the possibilities to create a browser fingerprint. This can be achieved via blocking the scripts that perform the fingerprinting. However, many web pages rely on these scripts in order to function properly for legitimate purposes: to enable fraud prevention, suggest related content, collect performance statistics which are arguably of benefit to the user. Another possibility is to block access to specific attributes and therefore reducing the number of fingerprintable attributes which reduces the fingerprintability of the browser. However it is only possible to block access to some specific attributes without severely limiting usability.

ATTRIBUTE SWITCHING *Limitations: 1,4* Another defensive strategy is to alter the value of certain attributes from a predefined set of configurations. These profiles can be extracted from real browser configurations, this is important in order to preserve fingerprint consistency to avoid the second limitation. However, due to the richness of modern browsers it is very hard to include every attribute and the switching of attributes could be detected if the fingerprint would be taken at different instances.

SPOOFING ATTRIBUTE *Limitations: 1,2,4* Attributes that are used in order to construct a fingerprint can be spoofed. This is mostly achieved as an extension of attribute switching to also adding random noise to those attributes that are the result of some rendering process, such as audio and canvas fingerprinting. By doing this the fingerprint changes, which in turn changes the perceived identity of the user by the tracker. Due to the randomisation unrealistic parameters might occur together with non deterministic behaviour. Both lead to the detection of spoofing and can make the user's fingerprint stand out more due to these inconsistencies. As for all techniques holds that it is very hard to include every attribute in the process.

RECONFIGURATION *Limitations: 1* Instead of trying to block or spoof certain attributes, this strategy is more focused on hiding amongst the crowd. By reconfiguring certain attributes, such as the browser language, the time and date and the list of plugins in combination with spoofing attributes values, such as audio and canvas fingerprint to resemble other users, fingerprinting is less valuable since the fingerprint is no longer unique. In order to overcome the limitations two and four it is important to base profiles on real browser fingerprints and adopt the same fingerprint for the entirety of each session. Once again holds that it is very hard to include every attribute in the process.

The hardest limitation to overcome is to determine the entire fingerprinting surface. In order to block fingerprinting via the spoofing of attributes and reconfiguration of the system, a list of fingerprintable attributes must continuously be kept up to date in order to effectively circumvent fingerprinting [49]. Consequently, due to the limitations described for each defensive strat-

egy, it becomes clear that taking countermeasures against browser fingerprinting is incredibly difficult: they either limit functionality or, by making small mistake, they still allow for fingerprinting the user.

2.2.5 *Overview*

Based on the combined results of previously conducted studies on browser fingerprinting it becomes apparent, see Table 2.2, that the most popular defensive technique is spoofing, this makes sense with regards to the limitations of blocking fingerprinting as described above. However, in Table 2.2 can also be seen that many attributes do not have a defensive strategy implemented at all. It can be noticed that the fingerprinting surface is large, and it is no surprise that as described in Section 2.2.4, *limitation 1* applies for every defensive strategy.

Table 2.2 has been constructed based on results from surveys [29] and [49], supplemented with [4], [5], [7], [11], see Table 2.3 for a full overview of sources. In comparison with Section 2.2.4 spoofing might be either *Attribute switching*, *Spoofing attribute* or *Reconfiguration*. Due to the innovative character of computer science, changes might have occurred so a blank space does not conclusively mean that the particular product possibly does not block or spoof that particular attribute. The list is also by no means inclusive, but rather an survey of attributes that have previously been reported in literature.

2.3 MALICIOUS INTENT

While these anti-fingerprinting products can protect the user to prevent them from being tracked amongst the web, most of these allegedly criminal products provide a natural copy of a real digital identity. It allows a malicious user to become a digital twin of another user, effectively bypassing anti-fraud fingerprinting detection mechanisms. This goes much beyond protecting your privacy and is focused towards committing crime. More specifically databases online are filled with stolen digital identities [32]. These digital identities allow cybercriminals to gain access to various online services normally accessed by the victim. In addition gaining access to digital identities on scale also opens up pathways to engagement and click fraud [20] In general cyber crime can be classified in two classes: mass cybercrime and targeted attacks [3] [21].

	Fingerprinters						Countermeasures										
Attribute	Pan	BC	IO	TM	Add	FPfs	BF	FA	DCB	FPB	FG	FPB	PV	Tor	BB	RAS	RG
Plugin Enumeration	⊗	⊗	⊗	⊗	⊗	⊗	○	×	○	○	×	×	?	×	×		×
Font Detection	⊗	⊗		⊗					○	×	×	?	?	○		?	
User-Agent	⊗	⊗	⊗	⊗	⊗	⊗	○	×	○		○	○		○	○	?	×
HTTP Header Accept	⊗						○										
HTTP Header Accept-Charset	⊗							×						?			
HTTP Header Accept-Encoding	⊗						○					?				?	
HTTP Header Accept-Language	⊗						○				○	○		○	○	?	
Screen Resolution	⊗	⊗	⊗	⊗	⊗	⊗	○	×	○	×	?	○		○		?	
Timezone	⊗	⊗	⊗	⊗	⊗	⊗	○		○		?	○		○		?	
Browser Language		⊗	⊗		⊗	⊗		×	○		○	○		○	○		×
OS & Kernel Version		⊗	⊗	⊗	⊗	⊗	○	×	○	×	?	○		○		?	×
DOM Storage	⊗	⊗	⊗	⊗	⊗	⊗						?				?	
IE userData	⊗	⊗															
Java Enabled	⊗				⊗							○					×
DNT User Choice					⊗	⊗						○				?	×
Cookies Enabled	⊗											○					×
JS detect: Flash Enabled	⊗	⊗	⊗	⊗	⊗	⊗					×	○		×			×
ActiveX + CLSIDs	⊗	⊗		⊗	⊗	⊗											
Date & Time		⊗	⊗	⊗	⊗				○								
CPU		⊗	⊗		⊗	⊗						○		○		?	×
System/User Language		⊗	⊗		⊗							?					
OpenDatabase			⊗		⊗	⊗						?					
Canvas Fingerprinting					⊗	⊗	○		○	○		?		?	○	?	
Mime-type Enumeration	⊗			⊗			○		○	○	×	×	?	×	×		×
HTTP Proxy Detection			⊗	⊗													
IndexedDB					⊗	⊗						?					
Math Constants		⊗			⊗												
Windows Registry		⊗	⊗														
TCP/IP Parameters		⊗	⊗											?			
Google Gears Detection		⊗															
Flash Manufacturer				⊗													
MSIE Security Policy		⊗															
AJAX Implementation		⊗															
MSIE Product key			⊗														
Device Enumeration		⊗															
Device Identifiers			⊗														
IP address		⊗						×						?			
HTML Body Behaviour						⊗											
Battery					⊗		×					?		?	×		
WebGLRenderingContext					⊗		×	×				?		?	×	?	
(WebKit-)AudioContext							×								×		
appCodeName												○		○			×
product												○		○			×
productSub														○			×
vender												○		○			×
venderSub														○			×
onLine												○					×
appVersion												○		○			×
Screen color and pixel depth									○	×		○		○			
Screen avail(Left/Top/Height/Width)														○			
Screen horizontal/vertical DPI														○			

⊗: Fingerprinted attribute
×: Attribute is blocked by countermeasure
○: Attribute is spoofed by countermeasure
? : Unknown whether attribute is blocked or spoofed by countermeasure

Table 2.2: Comparison of attributes used by various fingerprinting libraries and the capability of anti-fingerprinting products to block or spoof them.

Fingerprinters

Scientific		Adapted from
Pan	Panopticlick [11]	[49]

Commercial

BC	BlueCava	[49]
IO	Iovation	[49]
TM	ThreatMatrix	[49]
Add	AddThis	[49]
FPjs	FingerPrintJS	[49]

Countermeasures

Scientific		Adapted from
BF	BFingerprinting [7]	[7]
FA	FingerprintAlert [4]	[4]
DCB	Disguised Chromium Browser [5]	[5]
FPG	FPGuard [13]	[5]
FG	FireGloves [6]	[49][29]
FPB	FingerPrint-Block [49]	[49][29]
PV	PriVaricator [34]	[49]

Browsers

Tor	Tor Browser Bundle	[49][29]
BB	Brave Browser	[49]

Commercial

RAS	Random Agent Spoofer	[49]
RG	RubberGlove	[29]

Table 2.3: Overview of fingerprinters and countermeasures with their respective sources for Table 2.2

2.4 SUMMARY

Countermeasures against fingerprint tracking have been identified together with their associated limitations. Countermeasures were evaluated based on their effectiveness to prevent the fingerprinting of attributes. It became apparent that many existing countermeasures do not fully cover all attributes, leaving a large surface on which (mostly) unique fingerprints can be created. Finally anti-fingerprinting can also be used as a pathway to enable fraud which is either targeted cyber crime or mass cyber crime.

3

METHOD

In this chapter the methods from each research question are introduced here to facilitate the reader. Based on the method you can decide which part is of most interest to you. However, it is advisable to read the research in order as presented.

3.1 APPROACH

This research consists of four parts based upon the four research questions from Section 1.2. Each part will have it's own introduction, approach, result and conclusion. Finally, the conclusion will be drawn and opportunities for future work will be given.

3.2 OBJECTIVES

For each of the four objectives the method and scope will be shortly outlined below.

3.2.1 *State of the Market*

This part of the research resembles a 'literature study'. However, due to the lack of reporting of this phenomenon in scholarly articles, clear-web articles and darknet forums have been analysed in order to gain a good understanding of what types of products are out there and what they exactly do. Based on witness reports, instruction videos and threat intelligence reports, the landscape of anti-fingerprinting products was described and a comparison between the most popular products was made. Because in this research there is mainly an interest in illegitimate use of anti-fingerprinting. More sources have been considered to built an understanding of the anti-fingerprinting market than conventional in a literature study. In this part the study is limited to anti-fingerprinting markets with the intend to impersonate others.

3.2.2 *Threat Analysis*

Based on the popularity of one particular product, a deep dive into one of the anti-fingerprinting products was taken. This product is turned inside out in order to discover its full functionality and on top of that a quantitative and qualitative study was done into the information that is offered by this product. All information was gathered by manual information gathering either directly from the platform or combined from online articles. The approach

to gather all information manually was made in order to not stand out in terms of traffic due to the criminal nature of the platform. In this part the study is limited to results from only this product.

3.2.3 *Anti Fingerprinting*

Before performing the experiment the ethical and legal implications of the data collection were studied. After that the experiment is described. Based on the knowledge that was acquired by performing the study for the related work, various browsers, including those who utilise anti-fingerprinting technologies, were subjected to a fingerprinting system in order to establish an understanding of the spoofed attributes. The data was collected by developing a web-server which collects a browser fingerprint when visited. This server was hosted within the environment of National High Tech Crime Unit (NHTCU) and subsequently analysed in order to investigate any similarities between all tests. In this part the study is limited to only consider the most popular fingerprintable attributes.

3.2.4 *Protection*

While investigating the phenomenon of browser fingerprinting ideas arose to better defend against anti-fingerprinting. Famously put by [49]: "it is impossible in practice to determine and spoof the full set of fingerprintable characteristics", and as reported by Section 2.2.4 it is very hard to determine the entire fingerprinting surface, let alone spoofing all attributes correctly. The main idea for the additional protection was to increase this surface even more by creating an DNA-like sequence for each browser. The data was collected with the same web-server as described in Section 3.2.3. The analysis that was performed on this data compared the sequences in terms of similarity in order to attempt to separate browser who utilise anti-fingerprinting technology from those who do not. In this part the study is limited to only consider browser feature support to increase the fingerprinting surface.

4

STATE OF THE MARKET

In this chapter the state of the anti-fingerprinting market is reviewed, in order to provide an overview of the different variations of products that are purchasable on the internet. This study only regards anti-fingerprinting markets with the intend to impersonate others, in other words; those who abuse the functionality of altering one's fingerprint as described in Chapter 2.

4.1 OVERVIEW

In this study two essential parts to enable anti-fingerprinting have been identified:

- *Browser (plugin)* A customised browser or a plugin is provided by various parties that aid in configuring profiles from others.
- *Fingerprints database* A database filled with fingerprints from others serves as a shop where digital identities can be purchased.

This is regarded the 'new' way of using anti-fingerprinting [38], previously cybercriminals were mainly using virtual machines with all sorts of different browser configurations to try to closely match the configuration of their victim. This process has been improved and simplified by being able to select a browser configuration from a list and has expanded as far as to be able to select a stolen fingerprint from the victim's machine [8].

4.2 PRODUCTS

With consultation of experts, research into open sources and research on (darknet) forums, this list of popular anti-fingerprinting products was created with each their own unique features and properties, for a full overview see Table 4.1.

4.2.1 *Categories*

Roughly three categories have been identified:

- Browsers that spoofs attributes in order to hide the identity or match the configuration from a victim as close as possible
- Virtual Machine (VM)s configured with different browser configurations in order to match those as closely as possible to a victims configuration

- Browsers which can utilise databases containing stolen fingerprints where the victims' configuration can be copied for a near perfect copy

4.2.1.1 *Spoofing browsers*

In this category belong browser such as Antidetect and Linken Sphere. These browsers provide functionality to change all sorts of browser attributes in order to hide the identity of the user or impersonate a victim. These type of anti-fingerprinting products belong both in the *legitimate* and *illegitimate use* category. It can be used in order to protect your privacy, block cross domain tracking, but also for impersonation and fraud.

4.2.1.2 *Virtual Machines*

In this category belongs a product such as Fraudfox or a custom setup. By utilising the functionality of a virtual machine that can reset the configuration every time on startup a fresh identity can be generated every time in order to protect the identity of the user. However, with a product such as Fraudfox the functionality is built in to have as much as possible flexibility with creating a certain digital identity, which is aimed towards impersonating the digital identity of someone else. In the case of Fraudfox this anti-fingerprinting technology belongs in the *illegitimate use* category due to its clearly advertised functionality in order to enable fraud.

4.2.1.3 *Fingerprint databases*

In this category belong products such as the Genesis Market and Richlogs. These products enable the purchase of stolen real digital identities with one goal only: impersonation of the victim. This category also belongs to the *illegitimate use* category. Selling stolen profiles from victims is clearly meant for impersonation and not for preserving privacy of the user.

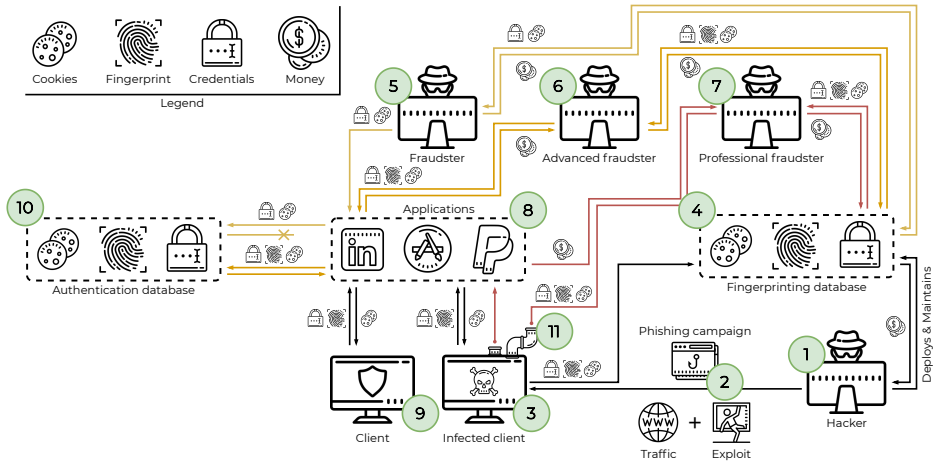

Figure 4.1: Fingerprinting landscape

4.3 LANDSCAPE

The anti-fingerprinting landscape roughly consists of *hackers* (1), *fraudsters* (5, 6, 7), *infected clients* (3), *fingerprinting databases* (4, 10) and *(targeted) applications* (8).

For a full overview see Figure 4.1.

1. The general idea is that the **hacker** (1) infects as many clients (3, 9) as possible with
2. *stealer malware* (2) [27] by either **traffic + exploit kits** or **mail spreading campaigns** (2) [8] [40]. Stealer Malware is malware that is delivered to someone either via an exploit or an malicious program, subsequently the malware extracts all information of interest and sends it to the server of the hacker.
3. Upon infecting the **client** (3)
4. the malware sends all available information contained by the local browsers on the client to a central **fingerprinting database** (4) which the hacker deploys and maintains. Once there are numerous profiles with stolen information available, fraudsters (5, 6, 7) can purchase these profiles and use them fro any kind of fraud, see Section 2.3. Before advanced fraud detection was available, using a correct combination of a password and username or stolen session cookies was enough to gain access to another account. However, this is no longer a viable modus operandi due to countermeasures which involve browser fingerprinting. Roughly three levels of sophistication are distinguished:
5. **Fraudster** (5) Uses a combination of correct password and username or stolen session cookies.
6. **Advanced Fraudster** (6) Uses the same as above in addition with a stolen fingerprint to circumvent fraud detection.
7. **Professional Fraudster** (7) Uses the same as above in addition with a proxy (11) to the infected client to bypass advanced countermeasures such as IP whitelisting.
8. All of these fraudster will attempt to access **applications** (8) on behalf of others. Databases containing stolen fingerprints, see Table 4.1, uniquely provide services for the advanced (6) and professional (7) fraudster in order to gain this access on behalf of infected clients (3).
9. Normally a **client** (9) would connect to the application, with a password and its fingerprint would be taken.
10. This fingerprint is subsequently compared with the **authentication database** (10) in which the password and fingerprint are authenticated, in return the client (9) receives a cookie. However in the case of the infected client (3) the password and fingerprint is provided by either the

advanced (6) or professional (7) fraudster. The normal fraudster (5) will no longer be able to execute the attack due to the countermeasures as described in point 4.

11. In very advanced cases where also IP filtering would be utilised the professional fraudster (7) can also utilise a **proxy** (11) which can be used to directly connect to the application from the infected client (3).

It has been shown that offering of digital identities can lead to serious threats. Identities from for example the Serbian Traffic Police, the New Zealand Internal Revenue Service, and the Qatar Government National Authentication System have been found [3]. In general, the victims are from all around the globe [50].

4.4 DISCUSSION

Over time various services that offer anti-fingerprinting products have been developed, each with their own working, pricing model and services. All of those try to accomplish the same goal to impersonate a victim in order to access services on behalf of someone else. The cat and mouse game between companies who utilise browser fingerprinting for anti-fraud purposes and the cybercriminals who purchase products for anti-fingerprinting solutions is clearly visible with the development of the services described above.

4.4.1 *Limitations*

For a more elaborate comparison between all services they should all be inspected in more detail, due to the pricing and limited possibilities to access the services this was not possible.

4.5 CONCLUSION

Referring back the research question 'What is the current state of the anti-fingerprinting market, which products are offered and how do they compare?' the anti-fingerprinting market has been analysed. As a result of this analysis, first the fingerprinting landscape has been identified where all of the identified products fit in. Subsequently the most popular services at the moment of writing have been identified and compared in terms of functionality, provided services and pricing. Although the functionality and pricing might be different in each case, the goal of all products is exactly the same: provide the functionality to be able to clone a victims' digital identity with minimal effort in order to enable fraud and impersonation.

DESCRIPTION	NAME	EXTRA SERVICES	PRICING	
Browser that spoofs attributes to hide identity	Antidetect[1]	Fingerprint database	Price	$2999
			Monthly	$100
			Fingerprint	free
	Linken Sphere [2]	Fingerprint database	Price	—
			Monthly	$100
			Fingerprint	$3-5
VM configured with different browser configurations	Fraudfox	Customizable fingerprints	Price	—
			Monthly	$100
			Fingerprint	— [3]
Databases containing stolen fingerprints	Genesis Market	Chromium plugin to configure fingerprints	Price	$25 [4]
			Monthly	— [5]
			Fingerprint	$1-95
	Richlogs	Chromium plugin and remote infected client access	Price	$50
			Monthly	—
			Fingerprint	$10 [6]

[1] Antidetect is (one of) the first commercial services to offer anti-fingerprinting. Many rip-offs of antidetect exist which are not considered in this research.
[2] The same makers also published the *Sphere* browser, which is free.
[3] Fingerprints must be configured yourself.
[4] Invite only, invites can be bought from other users for on average $25.
[5] At least one purchase every three months is required.
[6] No access to the richlogs database, price was taken from a screenshot [50].

Table 4.1: Overview of anti-fingerprinting products.

5

THREAT ANALYSIS

In this chapter an analysis is performed to estimate the threat that anti-fingerprinting markets pose. By assessing the number of available stolen identities in combination with the associated values of these identities the severeness of the threat of anti fingerprinting markets is evaluated.

5.1 SCOPE

In Table 4.1 can be seen that Antidetect, Linken Sphere and Fraudfox have a high associated cost, furthermore there was virtually no information available about these anti-fingerprinting products. Therefore these products were not subjected to the analysis. Based on various web articles the popularity of both the Genesis Market en Richlogs was analysed. The results can be found in Figure 5.1, here the total number of digital identities available for purchase can be found. Based on the large amount of identities available on the Genesis Market, the rest of the research is focused mainly towards the Genesis Market.

5.2 GENESIS MARKET

The Genesis Market is a large provider of stolen digital identities[8]. These digital identities are advertised in order to evade anti-fingerprinting and anti-fraud systems. Their purpose is to defeat the "We do not recognise your device" security check [38], by providing more information than only credentials but a full digital identity also containing a browser fingerprint. This enabled users of the platform to either monetise stolen user accounts or use the information for impersonation and targeted attacks. It is the Genesis Market which is the largest provider of this type of cyber crime [42]. The users of the Genesis Market fit the profile of advanced or professional fraudster, see Section 4.3. The Genesis Market is a shop with one seller as opposed to a market with supply and demand. Presumably the market is operated by a team of administrators, who are responsible to keep the shop running and stocked with digital identities [43].

5.2.1 *Analysis*

Based on the huge number of available number of stolen digital identities on the Genesis Market, this market was subjected to a deeper analysis into these available identities. The Genesis Market is reachable on the clearnet on `genesis.market` and is invite only. An invite is only obtainable through other members which spend at least $20 on their platform. The person who

Note: due to the criminal intent of the Genesis Market be advised to never visit this url without protecting your identity. Using the Tor browser or an anonymous VPN + VM is highly recommended.

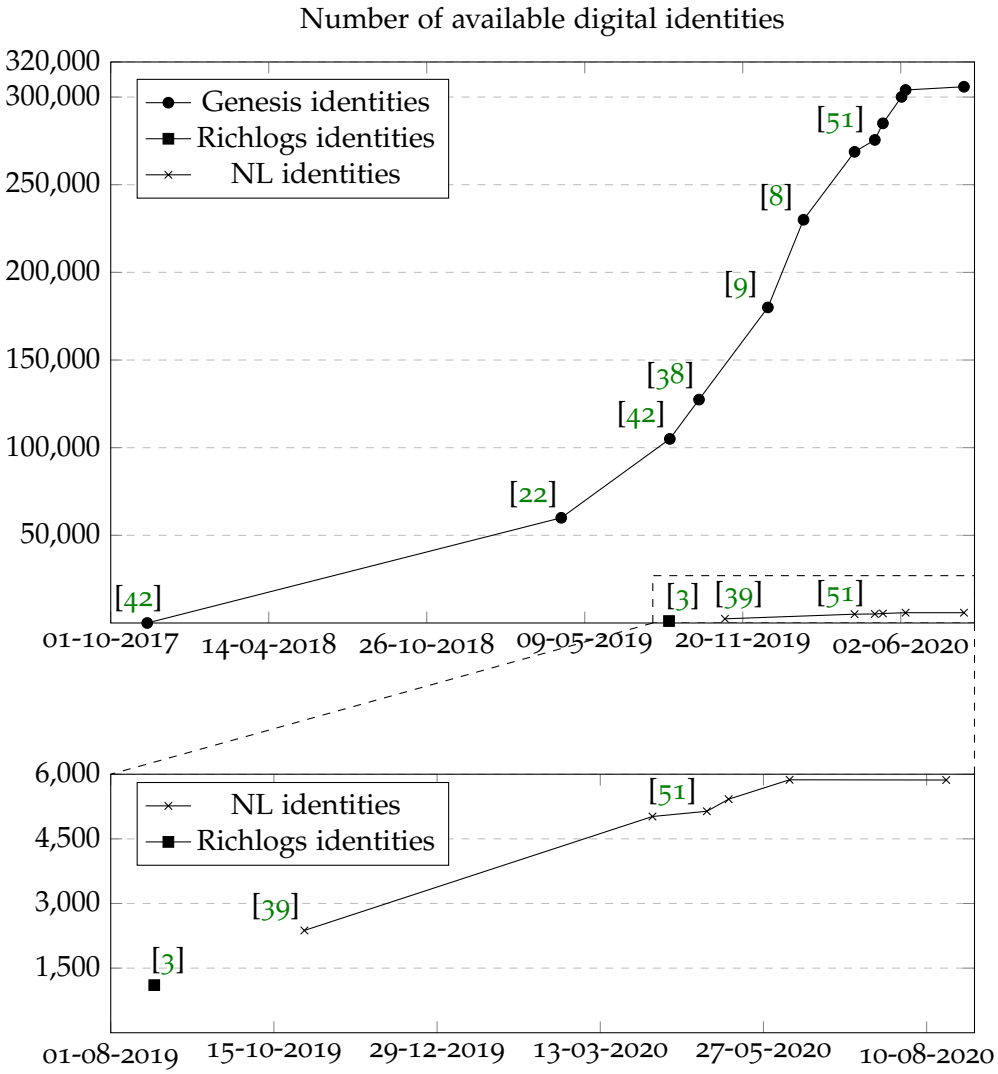

Figure 5.1: Overview of number of digital identities on the Genesis Market as reported by various sources over time.

referred you is subsequently responsible for your behaviour on the platform. The access to the Genesis Market can be revoked at any time. To prevent losing access to the Genesis Market all analyses that are performed for this study are performed manually and not programatically since this might stand out to the administrators of the market.

5.2.2 *Caveat*

Profiles that have been sold disappear from search on the Genesis Market. The numerical analysis that was performed only shows information about the profiles **that have not been purchased (yet)**, profiles disappear after 2 to 5 days which indicates millions of accounts have been sold on the Genesis Market [38]. Given that the Genesis Market is a criminal market operated by cybercriminals, no automatic analysis was performed in order to minimise

the risk of losing access. Due to this it is difficult to state exactly how many profiles are going in and out, what kind of profiles are more popular than others and if profiles from certain categories or countries are sold or offered more often than others. The analysis that was performed is therefore on what is still available on the market and online articles, which give an indication how many profiles are sold on the Genesis Market.

5.2.3 *Digital Identity*

Each digital identity consists of several parts, it was already common to sell stole credentials of individuals [38]. However, the Genesis Market takes it a step further, not only credentials are offered but also information about someones system including cookies and browser fingerprints.

5.2.4 *Offerings*

On the market all of the digital identities that are offered can be filtered on domain and country, see Figure 5.2. Once a profile has been selected there is more information available as visible in Figure 5.3 and Figure 5.4. The price of each profile is calculated with an unknown algorithm. The price seems to be based on the available information such as which services are available, 'Higher profile' services result in a higher price [22]. Service such as: Coinbase, Paypal, AliExpress and Binance are the most expensive [43]. These are all payment services and are apparently of greater value. Roughly the prices seem to built up in the following manner [43] or see Appendix B:

- Fingerprints+cookies+credentials for payment systems/email accounts: 50$ and up
- Fingerprints only: 50$
- Only credentials: 5$

Based on analysis of the Genesis Market most of the profiles that are available fall in to the lower two categories.

5.2.5 *Collection of digital identities*

The monthly influx of new digital identities is reported to be up to 50.000 profiles per month [27]. The question is: where do these profiles come from?

In contrast to other places were credentials can be bought, Genesis Market does not indicate which *Stealer Malware* is used in order to extract the private information from someone's computer.

Each profile on the Genesis Market has a Global Unique IDentifier (GUID) which can be used to identify which type of Stealer Malware was used in order to extract this information. The most recent addition of profiles are identified with a GUID that is formatted as a 32 character string, before that

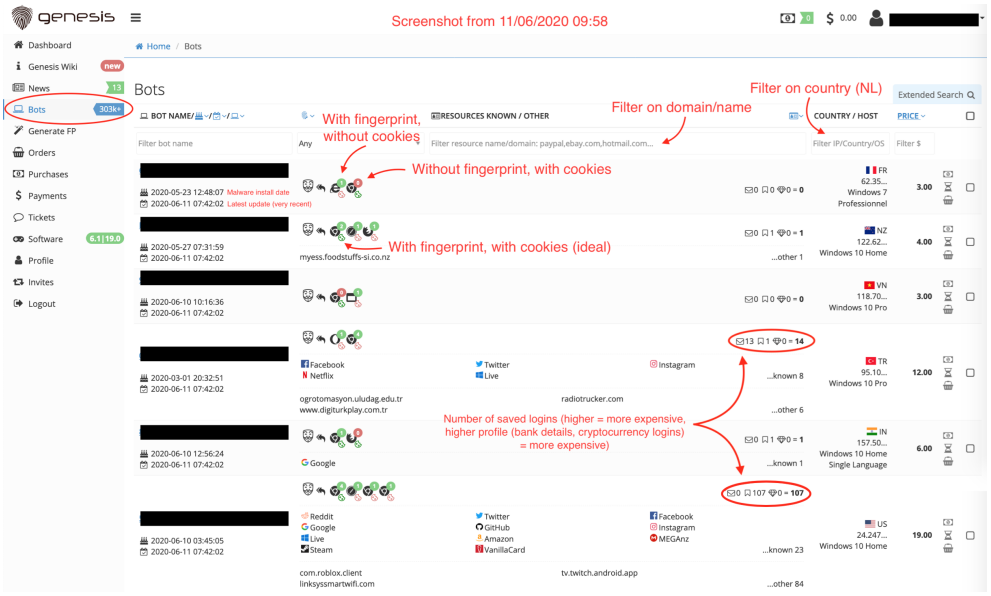

Figure 5.2: Overview of the *Bots* (profiles) page where all digital identities can be filtered.

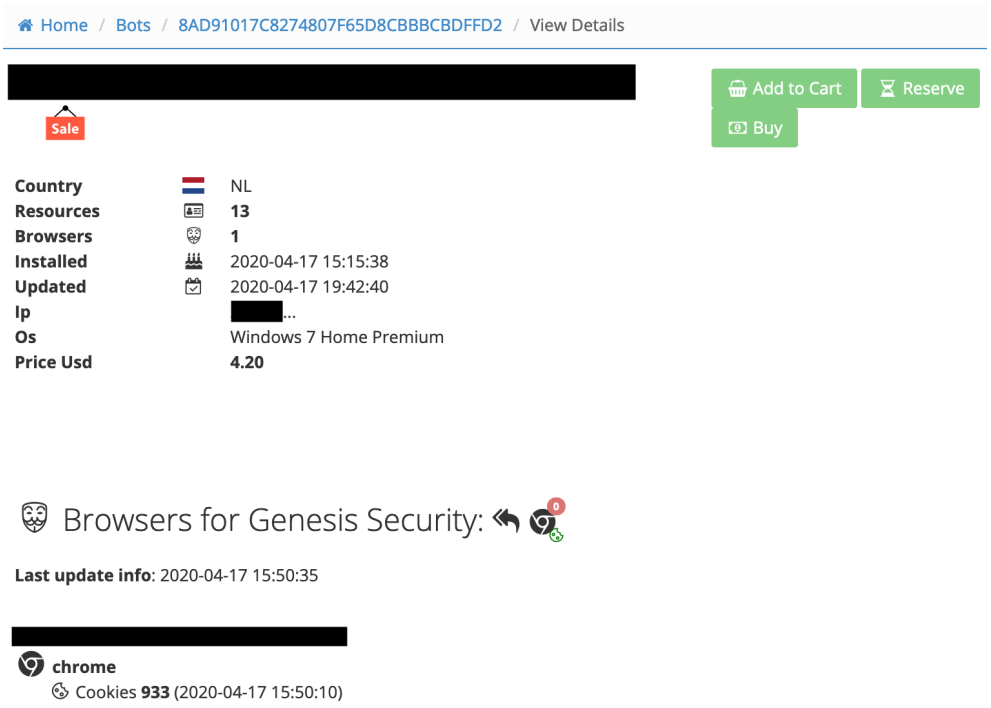

Figure 5.3: Detailed information about a profile before purchase. Displays whether there are fingerprints and cookies available, time of malware infection, time of latest information update, operating system, country and a IP 16 prefix.

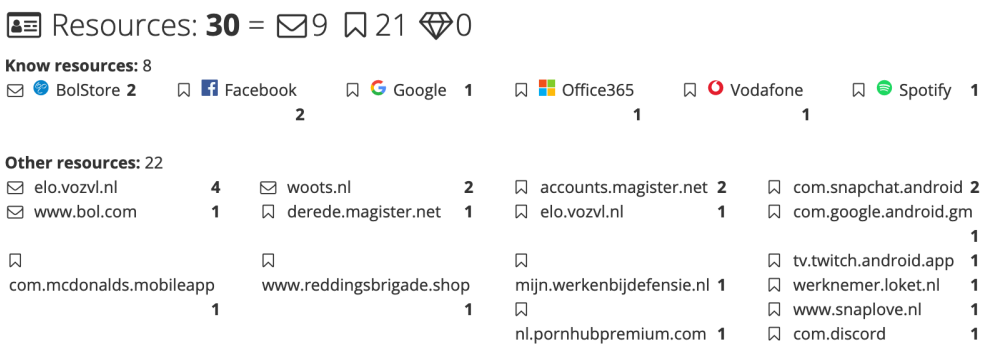

Figure 5.4: Detailed information about which sites are available for a profile. Displays the full URLs including the availability of a username and password if available, which will become visible after purchase.

the most popular format was 8-8-8-8-8, meaning 8 characters divided by a hyphen repeated 5 times.

Based on analysis [27] the 8-8-8-8-8 format points to AZORult stealer malware. However in February 2020 the influx of these type of profiles almost completely stopped due to the discontinued support by the developers and the broken compatibility with the newest version of Chrome 83 [26]. The Genesis Market administrators quickly found an alternative source of profiles by utilising stolen digital identities from the Raccoon stealer. The Raccoon stealer was confirmed to be used by cross referencing a profile on the Genesis Market with a database from Raccoon [39]. Raccoon extracts: Credit Card Data, Cryptocurrency Wallets, Passwords, Emails, Data from All Popular Browsers Including Credit Card Info, URLs, Usernames, Passwords, Cookies and System Information. Whether browser fingerprints are also extracted by Raccoon is unclear, however Genesis Market still offers fingerprints associated with accounts that originate from the Raccoon stealer. No fingerprints were found in the database from Raccoon.

5.2.6 *Maturity of service*

As opposed to dark web markets where account details are sold in a supply and demand manner, the Genesis Market is a full blown 'web shop' that can be filtered and searched for specific account details. The layout and appearance looks very professional, there is a wiki with explanations, tickets for customer service can be created and there is a news page with updates to their platform and their software. Compared to other criminal services everything is very well organised and appealing, the Genesis Market is named as on of the first *Account Takeover as a Service (ATaaS)* providers [28] [57]. This suggest that with increasing ease it is possible to utilise stolen credentials for fraud.

5.2.6.1 *News, Updates & User feedback*

Genesis Market incorporates news, (software) updates and user feedback within their platform, see Figure 5.2 on the left side of the screenshot. Updates with new functionality are presented in news articles where users can respond to. Users who respond on the platform mostly complement the development team for their work or ask for more functionality. Another small group of people offer others help how to effectively use the Genesis Market for fraud, see Figure 5.5.

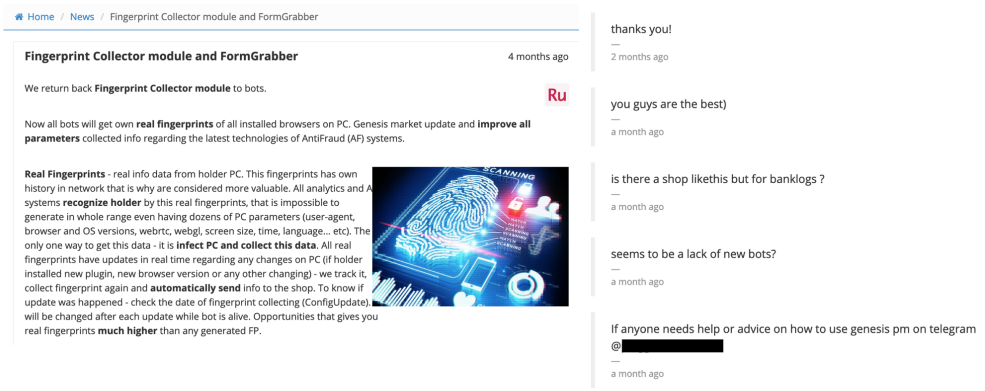

Figure 5.5: Screenshot of a news article on the Genesis Market included with some of the user feedback on this article, image was taken in August 2020.

5.2.7 *Software*

The most valuable additional service that is provided in addition to providing digital identities is the software that is provided by the Genesis Market, called the *Genesium browser*. The Genesium browser is a Chromium 77 based browser with a built in plugin, this plugin is their core software. Their software allows users to configure a browser plugin to directly load the bought information in order to use the *cloned browser identity* to immediately impersonate the victim. Upon brief inspection of the plugin, the code is heavily obfuscated and filled with anti-debugging measures, therefore cannot be easily reversed. Instructions how to use their browser plugin are available on their wiki but also on videos posted on YouTube, see Figure 5.6. The ease of use is clearly something the developers put in a lot of attention. Some profiles do not have an associated fingerprint available, if this is the case then it is possible for the user to create a fingerprint based on the information there is, this is not an organic fingerprint but does possibly provide a method to bypass anti fraud systems [43].

5.3 THREAT FOR THE NETHERLANDS

While analysing the profiles on the Genesis Market in particular attention was focused on the threat for Dutch accounts. Percentage wise not many

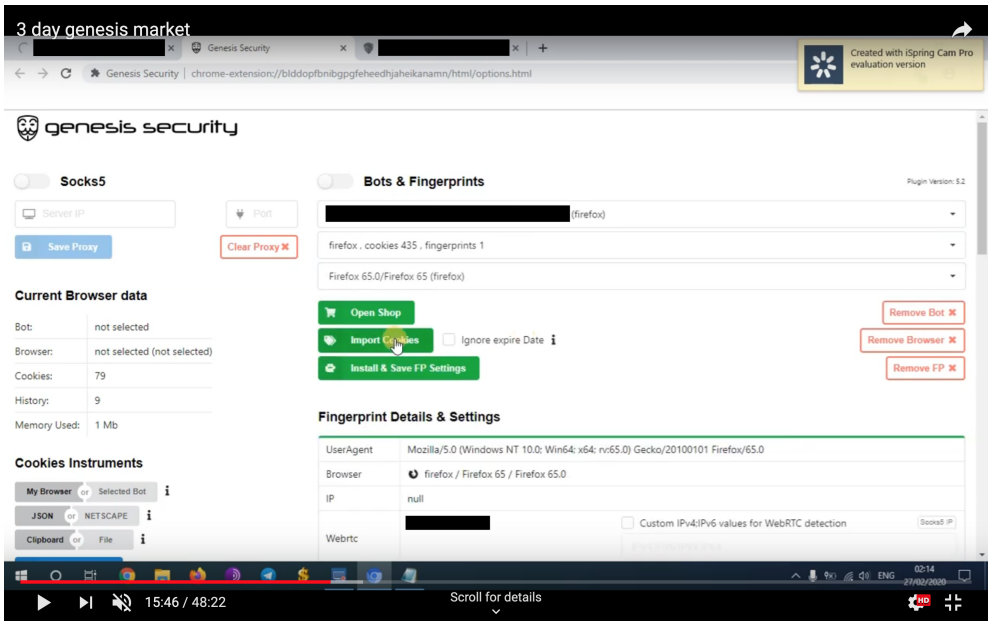

Figure 5.6: YouTube instruction video how the Genesis software can be used in order to load profile information into the browser in order to impersonate the victim.

profiles are from the Netherlands but it is still in the top 10 of countries that are currently offered on the Genesis Market, for the full overview of number of profiles available per country see Figure 5.7. Note that the caveat from Section 5.2.2 still holds and that these numbers only indicate the current number of available profiles. There seems to be a very high turnover rate [27] which puts the total amount of stolen identities worldwide somewhere between 1 and 3 million victims [38] and for the Dutch market between 12.000 to 50.000 victims.

The Genesis Market does not provide any profiles from the CIS (Commonwealth of Independent States) countries, the Raccoon stealer also does not collect information about systems from these states [40]. This has mostly a geopolitical reason, Russian cybercriminals operate with relative impunity inside Russia as long as they do not breach targets in their own country [45].

Based on manual analysis of profiles that were available some employee logins from companies that are in the sector of 'vital companies of the Netherlands' [56] have been identified [33] [35] [46] [61], for the results see Table 5.1 which contains an overview of how many profiles were available for purchase from companies in the vital sector. The customer and employee portals have been found through certificate transparency, which allows anyone to see registered SSL certificates for subdomains which lead to internal portals of companies [41]. Most interesting is to check for employee logins, since those could give cybercriminals easy access to internal systems [3]. In a broader search, a list of interesting subdomains was constructed based on their likeliness to contain employee logins, the number of results that were found for each subdomain are reported in Table 5.2.

	COMPANY	CONSUMER PORTAL	HITS	EMPLOYEE PORTAL	HITS
Drinkingwater companies	PWN Waterleidingbedrijf Noord-Holland	mijn.pwn.nl	21	pwnportal.nl	0
	Vitens	vitens.nl	52	mijnwerkplek.vitens.nl	0
	Waterbedrijf Groningen	mijn.waterbedrijfgroningen.nl	2	autodiscover.waterbedrijfgroningen.nl	0
	Waterleidingmaatschappij Drenthe	mijn.wmd.nl	4	autodiscover.wmd.nl	0
	Waternet	mijn.waternet.nl	11	werkt.waternet.nl	0
	Oasen	oasen.nl	0	-[1]	
	Dunea	mijndunea.onmicrosoft.com	5	-[1]	
	WML	wml.nl	11	wmlportal.sitel.com	1
				wml.maps.arcgis.com	1
	Evides	Evides.nl	17	-[1]	
	Brabant Water	mijn.brabantwater.nl	43	-[1]	
National transportation & distribution of electricity and gas	Rendo	-[1]		-[1]	
	Coteq	-[1]		-[1]	
	Liander	mijn-aansluiting.web.liander.nl	0	-[1]	
	Enexis	enexis.nl	5	adfs.enexis.nl	2
	Stedin	stedin.nl	0	-[1]	
	Westlandinfra	westlandinfra.nl	1	-[1]	
	Enduris	-[1]		-[1]	
	Eneco[2]	mijn.eneco.nl	20	-[1]	
		inloggen.eneco.nl[3]	54		
	Vattenfall[2]	accounts.vattenfall.nl	48	-[1]	
	Essent[2]	essent.nl	133	-[1]	
Storage, production and processing of nuclear material	COVRA	covra.nl	0	citrix.covra.nl	0
	EPZ	epz.nl	0	-[1]	
	NRG	nrg.eu	0	remote.nrg.eu	0
	PALLAS	pallasreactor.com	0	-[1]	
	Reactor Instituut Delft	tudelft.nl[4]	33	-[1]	
	URENCO	urenco.com	0	fs.urenco.com	0
Oil storage	COVA[5]	cova.nl	0	-[1]	
	Esso Nederland BV	exxonmobil.com	0	-[1]	
	Gunvor SA	gunvorgroup.com	0	-[1]	
	Hartree Partners UK	Hartreepartners	0	-[1]	
	Litasco SA	litasco	0	-[1]	
	Shell Nederland	Shell.com	2	-[1]	
	Varo Energy BV	varoenergy.com	0	-[1]	
	Vitol Netherlands BV	vitol.com	0	-[1]	
	BP Oil International	id.bp.com	11	pensionline.bp.com[6]	1
		bpes.bp.com[7]	2		

[1] No login found
[2] Electricity supplier
[3] Contains both consumer and employee logins
[4] These profiles originate from the University of Delft and are therefore not necessarily related to the reactor institute
[5] Independent organisation that oversights oil supplies
[6] Retirement funds
[7] Education platform

Table 5.1: Overview of available login portals on the Genesis Market from vital companies of the Netherlands. Collected on 11-05-2020.

SUBDOMAIN	HITS	RECORDED
Webmail		
adfs	844	11-05-2020
login	708	11-05-2020
mail	671	18-05-2020
webmail	476	18-05-2020
autodiscover	3	11-05-2020
(File) Sharing		
portal	1073	11-05-2020
intranet	104	11-05-2020
internal	1	11-05-2020
Remote Access		
remote	132	18-05-2020
citrix	25	18-05-2020
telewerken	6	03-06-2020
Online meeting & Collaboration		
werknemer	62	08-06-2020
jira	14	08-06-2020
meet	10	03-06-2020
Miscellaneous		
magister.net	596	11-05-2020
belastingdienst.nl	71	11-05-2020
utwente.nl	12	11-05-2020

Table 5.2: Overview of available subdomain names on the Genesis Market.

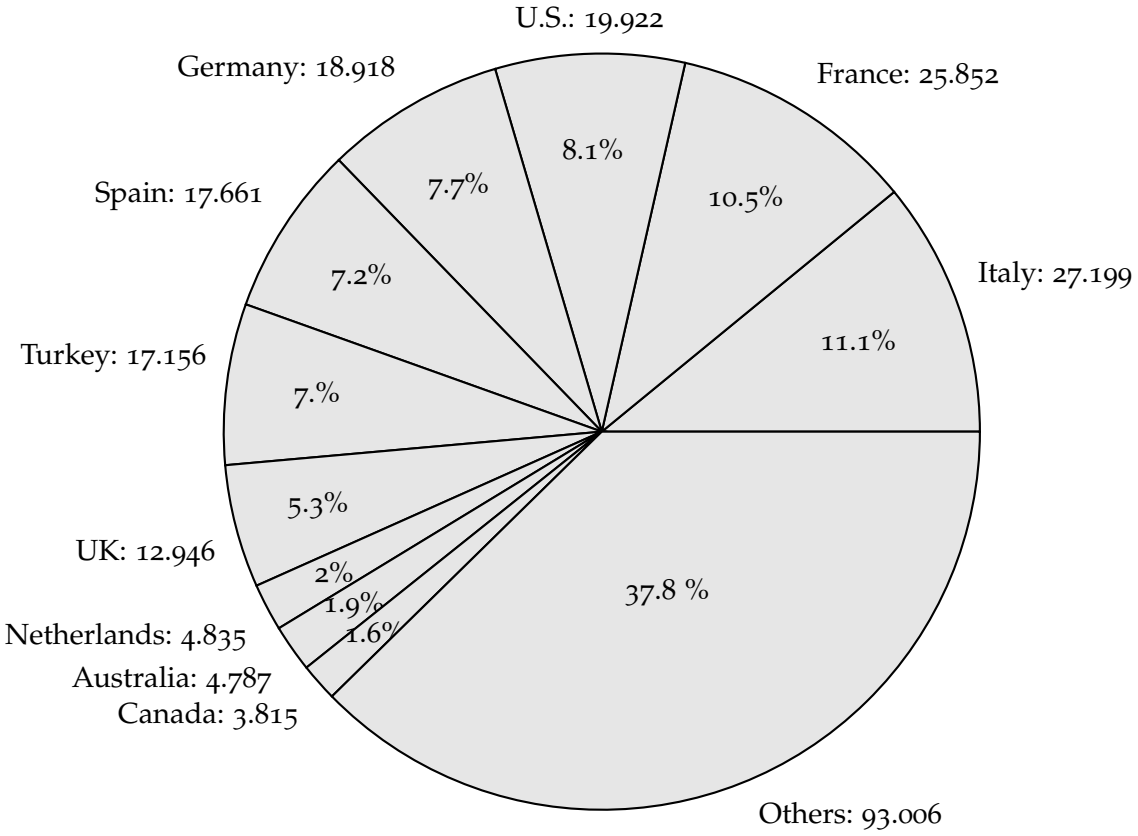

Figure 5.7: Number of profiles originating from various countries [43], state of February 2020.

5.4 IMPACT

Although the Genesis Market looks very professional and is easy to use, this does not automatically imply that it works in all cases. Users of the service have reported being unable to perform their desired fraud, in Figure 5.8 can be seen that users ask for instructions how to bypass Two Factor Authentication (2FA). They are being blocked by 2FA or security checks, while these security checks are exactly the checks that should be circumvented. For the Netherlands holds that all banking services always require 2FA, so even with credentials users of the Genesis Market will never be able to directly cash out on a Dutch bank account.

On the other side, the Genesis Market seems to be very popular with a high turnover rate of stolen information. For other countries and services it does not always hold that services which provide some sort of payment options are always protected with 2FA, so these might be easier to cash out for criminals and are therefore highly sought after profiles.

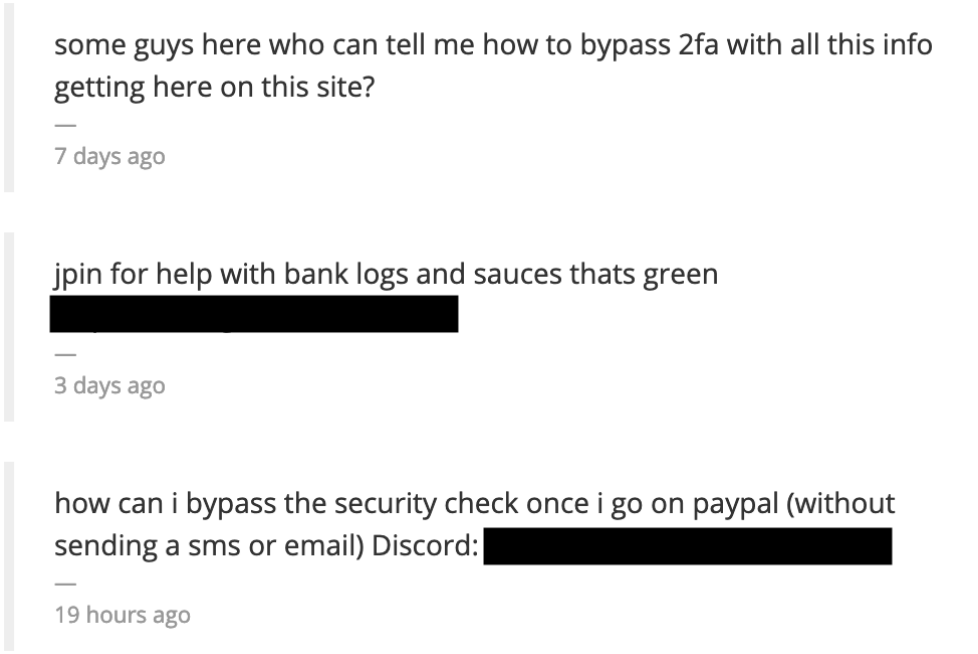

Figure 5.8: User feedback on the Genesis Market showing that not all users succeed in flawless impersonation, while other users offer their help to use the service better.

5.5 DISCUSSION

The Genesis Market is a clear example of the emerging threat of fraud with stolen credentials that has risen strongly past year. The way in which cybercriminals innovate, combining stolen information for more accurate impersonation, packaging everything nicely together poses a threat which has many victims.

5.5.1 *Limitations*

It is still difficult to state how many victims there are exactly and how severely they are impacted. Users of this service indicate that anti-fraud systems are not always fooled by the approach of the Genesis Market. Due to the phenomenon being of criminal origin, therefore being obscure, it is hard to tell in which occasions the information on the Genesis Market is most valuable and effective. However there seems to be solid evidence that the information is highly sought after and certainly offered in large volumes. Based on the pricing model and the low availability of highly priced profiles, see Section 5.2.4, this indicates that profiles that have higher chance of monetisation are mostly sought after.

This same problem relates to the results of Table 5.1 and Table 5.2, these numbers are a 'snapshot' on a certain date and only represent the profiles that have not been bought at that moment. It could very well be the case that many actors are looking for high value company employee profiles and

therefore are not there, or these companies are less likely to be in there due to stricter device policies.

5.5.2 *Suggestions*

In order to better understand what is the most valuable and effective information, victims should be investigated in order to establish what kind of fraud cybercriminals commit with the stolen information. In combination with an analysis on the differences and similarities between offered profiles from the various countries, ideas may arise for prevention and effectively raising awareness amongst victims.

5.6 CONCLUSION

Referring back to the research question 'How large is the upcoming threat of anti-fingerprinting? Does it enable mass cybercrime or targeted attacks?' it can be stated that the threat of anti-fingerprinting is significantly large. The number of victims are estimated to be millions, the market was growing very rapidly until recently, and there does not seem to be a way to stop cybercriminals from continuing this kind of crime. Whether it enables mass cybercrime or targeted attacks is a tougher question to answer. Much is unclear about the exact impact of offering so many credentials for relatively low prices with high ease of (mis)use. The information that is being offered is certainly very suitable for targeted attacks where the victim can be digitally impersonated, whether this happens at scale remains unanswered.

6

ANTI FINGERPRINTING

In this chapter the analysis of fingerprints from the market in the previous chapter will be described. First, a look at the ethical and legal implications of the data acquisition and the actual acquisition will be covered. Then, the analysis of the fingerprints will be described. After that, the results will be described and finally a conclusion will be drawn.

6.1 DATA ACQUISITION

In order to perform the research, samples are required from the Genesis Market, however these samples are stolen digital identities. To study the fingerprints, the stolen identities must be obtained. Due to the criminal nature of the Genesis Market, acquiring this data carries both ethical and legal implications. In some cases the ethical issues might also be illegal and have legal implications as well, some overlap can be expected.

6.1.1 *Ethical implications*

Due to the data acquisition of stolen digital identities two ethical issues arise: 1) By buying these identities money is transferred to a criminal organisation, 2) By obtaining these identities, Personal Identifiable Information (PII) including passwords is obtained. Based on the work of Thomas et al. [48], where he authors extracted ethical principles from existing advice and guidance and analyse 20 peer reviewed papers which deal with datasets from illicit origin. The authors list the set of ethical issues that require consideration when conducting research with data of illicit origin, which will be discussed in the following subsections. First the mitigations that were taken prior to the study will be briefly introduced, secondly the stakeholders will be identified, subsequently it will be discussed why informed consent is not possible, finally the justifications will be discussed with their belonging safeguards, and identified harms and benefits.

6.1.1.1 *Mitigations*

Based on the ethical analysis that is subsequently described, the following mitigations have been taken prior to performing this study. The mitigation are mainly focussed towards minimising the amount of sensitive data that is received by the researcher. The full list of harms with their respective mitigations can be found in Section 6.1.1.6. The most important points are:

- The risk of de-anonymisation is completely mitigated by never giving the researcher access to the sensitive PII of the victims. The profiles

were bought by the Dutch police, only the browser fingerprint was passed on to the researcher.

- No interaction with real systems was performed with the bought digital identities, only the controlled testing environment was subjected to the anti-fingerprinting browser.
- Only 11 identities in total were bought and tested in order to establish the functionality of the Genesis Market.
- All information used in this research is securely stored for the whole duration of the study and will be deleted afterwards.

6.1.1.2 *Stakeholders*

The primary, secondary and key stakeholders are identified in order to support the analysis of the potential harms and benefits of the research.

- **Primary** Primary stakeholders are those directly connected with data, such as those identified in it. In this case the primary stakeholders are the victims of the Genesis Market, their PII has been stolen from their infected machine and is subsequently sold on the market.
- **Secondary** The secondary stakeholders are intermediaries in the delivery of benefits or harms, such as service providers. Therefore, the administrators of the Genesis Market are identified as the secondary stakeholders. They are responsible for the delivery of benefits, for the users, and harms, for the victims.
- **Key** The key stakeholders are those such as the leaker or the researcher who are critical to the conduct of the research. Therefore the key stakeholders are the Dutch police and the researcher of this study. The Dutch police facilitated the data collection, which was critical in order to conduct this study.

6.1.1.3 *Informed Consent*

Difficulties arise when considering informed consent for the primary and secondary stakeholders. Firstly, the identity of the victims is unknown before the purchase has been made and the purchase immediately reveals their PII. Secondly, the administrators of the Genesis Market are involved with criminal activities and asking for their consent presumably leads to revoked access to the market. Therefore this study can only be performed without obtaining informed consent.

6.1.1.4 *Justifications*

The authors advice to implement safeguards, identify positive benefits and potential harms. And they indicate when ethical justifications are legitimate, the following justifications apply to this work:

- **No additional harm** The research is not identifying any natural persons and the data is stored and managed securely.
- **Fight malicious use** The research is using the data in order to describe the phenomenon and to study possible defensive mechanisms, on top of this the data that was acquired can no longer be used by someone else to afflict harm.
- **Necessary data** The research into defensive mechanisms cannot be conducted without using this data. Based on Thomas et al. [48] this is justified because there is public interest and there is no additional harm done. The public interest originates from the fact that the collected data aids this research and the Dutch police in investigating the phenomenon.

6.1.1.5 *Safeguards*

Also all safeguards have been implemented.

- **Secure Storage** The data is stored on secure storage, this means on an encrypted hard disk and on a NHTCU managed device in order to protect the integrity and confidentiality and to avoid accidental leakage. After the research all data on the device will be deleted.
- **Privacy** Respects privacy, this means no de-anonymisation is attempted and the study does not reveal any identities. In this study there is no access to the passwords or other sensitive PII of individuals, only the fingerprint is available.
- **Controlled Sharing** The data is shared in a controlled manner, this means that the data does not leave the environment of the NHTCU.

6.1.1.6 *Harms & Mitigations*

In order to perform a full review of the ethical implications, the harms and benefits are assessed according to [48].

- **Illicit measurement** The data for this research is obtained through paying the offenders, this is mitigated through legal means as described in Section 6.1.2.
- **Potential Abuse** Findings in this study, mainly with respect to defensive mechanisms, could be used by malicious actors to improve their anti-fingerprinting products as well. To mitigate this harm this research will be placed under a two year non disclosure agreement.
- **De-Anonymisation** This harm does not apply since no information from this study can be used to identify either individuals or groups, the data will not be published and PII will never be available to the researcher.

- **Sensitive Information** The data for this study only contains the browser fingerprint, which will not be disclosed to others, is securely stored and will be deleted after this study. The dataset that is held by the Dutch police containing the full digital identity, i.e.: passwords, cookies, configuration information and PII, will never be available to the researcher. Misuse is illegal by Dutch law, this will be described in Section 6.1.2.
- **Research Harm** This harm is twofold. Firstly, the data contains illegal material (stolen identities) which may lead to prosecution, this is mitigated as described in Section 6.1.2. Secondly, this research could lead to threats from criminals. Delayed publication might partially prevent this. Furthermore, the research is fully based on open sources and this is not the first study into these illegal services. However, research harm still poses a potential risk which cannot be completely mitigated.
- **Behavioural Change** In order to prevent behavioural changes amongst the stakeholder the data was never collected in an automised manner in order to prevent raising suspicion.

6.1.1.7 *Benefits*

Finally, the benefits are listed.

- **Reproducibility** In order to reproduce the results of this study, the next researcher must obtain the data through controlled sharing as explained above. This means this research can only be reproduced when the researcher will be placed under a non disclosure agreement and is provided access by the NHTCU.
- **Uniqueness** The data is not obtainable in any other way, it is useful to obtain the data because otherwise the defensive mechanisms cannot be designed.
- **Defence Mechanisms** This study studies stolen fingerprint in order to investigate the phenomenon and to design better anti-fingerprint defensive mechanisms.
- **Anthropology and Transparent** Does not apply.

6.1.2 *Legal implications*

The data obtained for this study contains stolen identities. By Dutch law Artikel 234 Wetboek van Strafrecht [58], obtaining, selling or possessing data which is intended to commit crimes as described by Dutch law Artikel 231b Wetboek van Strafrecht [58]: the deliberate action of using stolen PII with the purpose to steal or misuse this information in such a way that yields any negative consequence.

Note: This is loosely translated, for the exact definition the Wetboek van Strafrecht should be consulted.

For this research this means that no special permissions would be required, since by acquiring the information for this research there is no intent to misuse this information in any way. Otherwise, this would imply that performing this research would lead to a criminal offence which is for many reasons undesirable. However, the owner(s) of Genesis Market are certainly punishable by this law since they possess and sell this PII which is clearly purposed to steal and misuse someones (digital) identity. Therefor, the Dutch police can lawfully gain access to the information on the Genesis Market based on Artikel 126i Wetboek van Strafvordering [59], by means of a *pseudo-purchase* [1]. By Artikel 213b Wetboek van Strafrecht [58] it is only a criminal offence when the action is extrajudicial, in this case an injunction is provided to perform a *pseudo-purchase*, therefore it is not extrajudicial and subsequently it is not a criminal offence. On top of that, the *pseudo-purchase* protects the identity of the person that purchased the stolen identities and therefore partially mitigates researcher harm as described in Section 6.1.1.

The information that has been acquired by the NHTCU can now lawfully be used in their investigation as it can be used for this research. This effort is required in order to protect anyone from haphazard invasion of privacy by the police, although the data that is required only contains information about the victims and they are arguably better off when the information is bought by none criminal actors, the purchase still requires interaction with the owner of Genesis Market and therefore requires the interaction as described above. All problems with regards to ethics still hold, see Section 6.1.1.

6.1.3 *Ethical & Legal Conclusion*

In order to collect fingerprints from individuals these must be purchased directly from the Genesis Market. To obtain this data a *pseudo-purchase* was performed, see Section 6.1.2. Combined with the ethical justification, implemented safeguards and the identification of harms and benefits, see Section 6.1.1. The collection of data for this research is considered ethical and safe.

6.2 ANALYSIS

The Genesis Market can be filtered on profiles from the Netherlands, see Section 5.2, and now the profiles can be bought in the store. After purchase the profile becomes available to configure in your account via the plugin, see fig:genesis:youtube-instruction.

The setup is as follows:

- 11 Unique profiles were bought from the Genesis Market, these profiles must be able to digitally imitate the browser from another individual based on cookies, login information and in particular the browser fingerprint. Each of the profiles were configured in the Genesium browser and subjected to the testing system.

- 1 Test was performed with the Genesium browser without a configured profile, which serves as a control instance.
- 5 Spoofed instances were subjected to the testing system. With each time an unique fingerprint configuration, configured in the Sphere browser (free software from Linken Sphere, see Table 4.1: note 2).
- 1 Test was performed with the Sphere browser without a configured fingerprint, which serves as a control instance.
- 6 Versions of a Chromium based browser were subjected to the testing system.
- 5 versions of a Firefox based browser were subjected to the testing system.
- 1 Version of the Safari browser was subjected to the testing system.

In total 30 tests were executed, the overview which contains information about the experiment identifier, the tested version, the used operating system, whether the browser utilises spoofing and as which browser the identifies itself is summarised in Table 6.1.

6.2.1 *Fingerprinting*

In order to test the effectiveness of the anti-fingerprinting capabilities, the browser fingerprint is extracted. For this FingerprintJS2[1] is utilised, which collects 32 attributes: user agent, screen resolution, timezone, canvas fingerprint, etcetera. For all attributes see Appendix C. The fingerprinting script has some spoofing detection built in as well, lastly the user agent can be used for identification in order to determine spoofing in combination with the feature support as described above.

6.3 RESULTS

In total the test was executed 30 times using different combinations of browsers, versions and operating systems, for a full overview see Table 6.1. For each test a full fingerprint was extracted as provided by the FingerprintJS2 library.

6.3.1 *Fingerprint attributes*

Of the 32 attributes that were captured, 31 attributes were compared in terms of similarity. The 'user agent' was left out and only utilised for identification purposes. The full list of attributes can be found in Appendix C and the results are displayed in Figure 6.1.

1 15, https://github.com/fingerprintjs/fingerprintjs2.

Group 1: Non spoofed browsers on a Windows operating system

Identifier	13	14	15	16	17	25
Installed Browser	Iron 72.0	Iridium 2019.04	Iridium 2020.04	UnGoogled 77.0	Chrome 83.0	Firefox 78.0.2
Operating System	Windows 10	Windows 10	Windows 10	Windows 10	Windows 10	Windows 10
Spoofing	No	No	No	No	No	No
Identified Browser	Chrome 72	Chrome 73	Chrome 80	Chrome 77	Chrome 83	Firefox 78

Group 2: Non spoofed browsers on a Mac operating system

Identifier	12	26	27	29
Installed Browser	Firefox 78.0	Firefox 78.0.1	Chrome 83.0	Safari 13.1.1
Operating System	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6
Spoofing	No	No	No	No
Identified Browser	Firefox 78	Firefox 78	Chrome 83	Safari 13

Group 3: Tor browser with spoofed attributes on both Windows and Mac

Identifier	18	28
Installed Browser	Tor 9.0.5[1]	Tor 9.5.1[1]
Operating System	Windows 10	MacOS 10.15.6
Spoofing	Yes	Yes
Identified Browser	Firefox 68	Firefox 68

Group 4: Sphere browser with configured spoofed fingerprint profiles

Identifier	19	20	21	22	23	24
Installed Browser	Sphere 1.3[2]	Sphere 1.3	Sphere 1.3	Sphere 1.3	Sphere 1.3	Sphere 1.3
Operating System	Windows 10	Windows 10	Windows 10	Windows 10	Windows 10	Windows 10
Spoofing	No[4]	Yes	Yes	Yes	Yes	Yes
Identified Browser	Chrome 65	Firefox 54	Firefox 54	Firefox 54	Firefox 54	Firefox 54

Group 5: Genesium browser with configured spoofed fingerprint profiles

Identifier	1	2	3	4	5
Installed Browser	Genesium 19	Genesium 19	Genesium 19	Genesium 19	Genesium 19
Operating System	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6
Spoofing	Yes	Yes	Yes	Yes	Yes
Identified Browser	Chrome 83	Chrome 83	Chrome 83	Chrome 79	Chrome 83

Identifier	6	7	8	9	10
Installed Browser	Genesium 19	Genesium 19	Genesium 19	Genesium 19	Genesium 19
Operating System	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6	MacOS 10.15.6
Spoofing	Yes	Yes	Yes	Yes	Yes
Identified Browser	Chrome 80	Opera 66	Chrome 80	Chrome 79	IE 11

Group 6: Genesium browser without configured spoofed fingerprint profiles

Identifier	0	11
Installed Browser	Genesium 19[3]	Genesium 19
Operating System	MacOS 10.15.6	MacOS 10.15.6
Spoofing	No[4]	Yes[5]
Identified Browser	Chrome 77	Chrome 80

[1] The Tor 9 browser is based on Firefox 68.
[2] The Sphere 1.3 browser is based on Chrome 65.
[3] The Genesium 19 browser is based on Chrome 77.
[4] Browser which is suitable for spoofing but no profile has been selected.
[5] While no fingerprint profile was associated with this profile an auto generated spoofing profile was provided and therefore a different identified browser was reported as opposed to the real browser.

Table 6.1: Overview of browsers subjected to the testing system.

6.3.2 *Groups*

In total for 30 browsers fingerprint samples were extracted. The results were grouped into 6 groups that each have their own characteristics.

- Group 1 and 2 both contain samples from browsers who do not utilise spoofing techniques which are either installed on (1) Windows or (2) MacOS.
- Group 3 contains samples from the Tor browser both installed on Windows and Mac. The Tor browser does utilise spoofing techniques [62].
- Group 4 contains samples from the free Sphere browser which utilises browser spoofing and customisable fingerprint profiles as explained in Section 4.2.1.1. For samples with identifier 20 through 24 holds that the browser was configured with an artificial fingerprint, however for sample 19 no fingerprint was configured, therefore the reported browser value is the Chrome 65 upon which the Sphere browser is based.
- Group 5 contains samples from the Genesium browser which utilise a configured spoofed fingerprint. Opposed to the artificially crafted fingerprints of group 4, the fingerprints from group 5 result from real users.
- Group 6 contains samples from the Genesium browser were in the case of sample with identifier 0 no fingerprint was configured. Therefore the reported browser value is the Chrome 77 upon which the Genesium browser is based. For sample with identifier 11 no 'actual user' fingerprint was available and received an artificial fingerprint similar to Group 4.

These groups will be referred to in figures and explanations throughout Chapter 6 and Chapter 7.

6.3.3 *Subgroups*

In order to compare the fingerprints in terms of similarity, each attribute was grouped. In order to prevent confusion with the groups from Table 6.1 these are called subgroups. The subgroups were determined as follows: each time a value corresponds exactly to another value they are grouped together. As can be seen in Table 6.2, "Browser 4" belongs to the same subgroup as "Browser 1" because the attribute value is the same. All values that have not been 'seen' before are allocated a new `n+1` subgroup.

Browser	**Browser 1**	**Browser 2**	**Browser 3**	**Browser 4**
Attribute value	"A"	"B"	"C"	"A"
Assigned subgroup	0	1	2	**0**

Table 6.2: Subgrouping process.

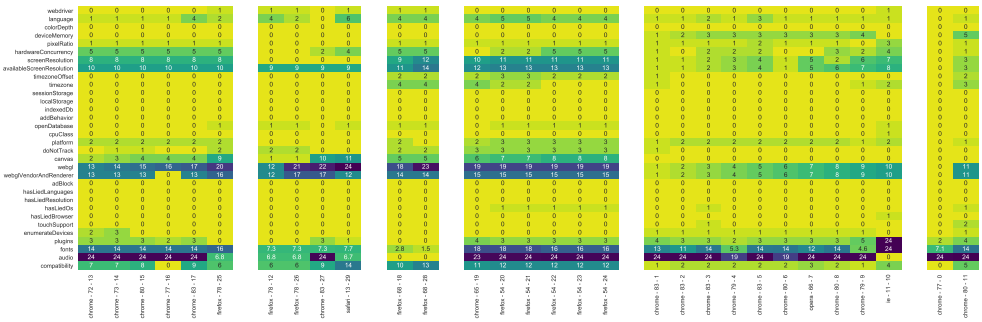

Figure 6.1: List of attributes that are fingerprinted by FingerprintJS2, each number indicates which *subgroup* it belongs to read from left to right. If the number is the same for multiple browsers it means the reported value is the same.

Note that the last three categories: plugins, fonts and audio are calculated differently and normalised between 0 and 24 to better be represented in the colour scale.

- Plugins displays the number of plugins installed in the browser
- Fonts displays the number of fonts normalised between 0 and 24
- Audio displays the audio fingerprint float, normalised between 0 and 24
- Compatibility has been included to subgroup the results from Chapter 7.

6.4 DISCUSSION

By observing the results from Figure 6.1, the following observations can be made:

1. In group 5, Genesium browser spoofed results, show a high variety amongst fingerprints: 'deviceMemory', 'screenResolution', 'availableScreenResolution', 'webgl', 'webglVendorAndRenderer' and 'fonts'.
2. Compared to other groups only 'webgl' shows this much variety amongst all results, for the rest can be observed that groups are formed.
3. In group 5 all 'canvas' fingerprints are exactly the same, even though this is one of the attributes that the Genesis Market claims to spoof, see Appendix B.
4. Some attributes do not add to the uniqueness of the fingerprints, because they are always the same: 'colorDepth', 'sessionStorage', 'localStorage', 'indexedDb', 'addBehavior', 'adBlock', 'hasLiedLanguages', 'hasLiedResolution'

It becomes visible that the test results from the Genesis Market software are showing a lot of variety in the fingerprints. However, it failed to spoof the canvas fingerprint for all 12 tests in group 5 and 6. This shows that not all attributes are correctly spoofed which could lead to fail to bypass anti-fraud systems, this could explain why some users are unable to user the Genesis Market in order to get access to certain accounts as described in Section 5.4.

6.4.1 *Limitations*

Although it can be shown that the fingerprinted attributes change with each test while using the Genesium browser, there has not been verified whether this is sufficient to bypass anti-fraud systems. It would be unethical to try this on real systems, since this could harm the victim of which the profile was stolen.

6.4.2 *Suggestions*

In order to test whether the Genesium browser loaded with a stolen digital identity would be able to successfully bypass anti-fraud systems a comprehensive experiment could be conducted where a fake system with fake accounts would be deliberately infected with the stealer malware. If subsequently this information would be passed to the Genesis Market, the profile could be bought and tested on real systems. It is not very trivial to execute this plan, since the steps to get the profile into the Genesis Market is out of our control and installing malware on a system must of course be done with great caution.

6.5 CONCLUSION

Referring back to the research question 'How do anti-fingerprinting solutions successfully circumvent browser fingerprinting?' it can be seen that fingerprint attributes are changed successfully while using the same browser loaded with different configurations from the Genesis Market. Not all attributes are altered such as the canvas fingerprint, which is considered as an advanced fingerprinting technique. It is not possible to conclude whether the altered fingerprint is enough in order to bypass anti-fraud systems, but it can be seen that the browser fingerprint is successfully altered each time.

7

PROTECTION

This chapter describes the approach of using browser features in order to distinguish spoofed browsers from real ones. This can aid in the fight against anti-fingerprinting fraud.

7.1 MOTIVATION

The administrators of Genesis Market claim that they developed the Genesium software by analysing the top 47 browser fingerprinting and tracking systems as well as those utilised by 283 different banking and payment systems [43].

However, their solution relies on a browser plugin, which is built into the Genesium browser, that works only in versions of Chromium below 77. Due to large differences in various browsers and the constant innovation and quick iterations of versions, the hypothesis is that it is almost impossible to 100% correctly spoof another browser or version, due to the fact that using a particular version of a browser introduces unavoidable characteristics. Due to the quick iterations of browser versions, it would be very costly for an attacker to keep updating browser fingerprints based on the version that is used by the infected client. For example, the *loading-lazy-attr* browser feature which allows developers control over when the browser should start loading images & iframes[1]. This attribute is not available for Chrome 75 and earlier, but is available for Chrome 76 and onwards. Extending this to the hypothesis means that if a Chrome 75 browser is used to spoof a Chrome 76 browser it is expected that this feature will not work and therefore give away that it is not truly a Chrome 76 browser.

7.2 APPROACH

To test this hypothesis, feature support from `caniuse.com` was analysed due to the variation in feature support that every browser and browser version introduces.

7.2.1 *Feature support*

"Can I Use" provides up-to-date browser support tables for support of front-end web technologies on desktop and mobile web browsers [52]. Each browser and version has particular features that it supports which might differ from a previous version or another browser. This information can therefore hypo-

1 53, https://caniuse.com/loading-lazy-attr.

thetically be used in order to identify browser and browser version based on the features that are supported. The information from the "Can I Use" database is therefore regarded as the *ground truth*. The testing page `tests.caniuse.com` provides 1244 tests which can either be executed:

- **Auto** Test containing JavaScript can be executed automatically
- **Visual** Test requires visual confirmation to check the result
- **Visual-Square** Test must create a visual square
- **Interactive** Test requires interaction to confirm support

In order to speed up the testing process only the automatic tests are considered in this study, which can be executed without user interaction. In total 588 tests remain. For each test is determined whether the test is supported on unsupported, additionally it can also occur that the feature is unreported.

7.2.2 *Sequences*

Based on the test each browser will result in an unique *sequence*, which is basically an array containing information on each attribute whether it is supported, unsupported or unreported. Hypothetically this sequence can be used to uniquely identify a browser version, similar to DNA sequencing that can be used to uniquely identify a person.

7.2.3 *Similarity scoring*

In order to characterise and compare each test similarity scoring was applied on the feature support sequences from Section 7.2.1. Each sequence was compared on similarity with all available browser from "Can I Use" and versions running from the current version up to 10 versions back. Note that for not all browsers there are 10 versions back, for example Internet explorer only has 7 versions, 5.5 through 11. If this is the case, all available versions from this browser are considered.

7.2.3.1 *Pairwise distance*

In order to calculate similarity between two browsers, the *Hamming distance* is used. The data contains binary information on the supported features, and the *Hamming Distance* is a suitable metric to compare two binary strings of equal length [30]. The Hamming distance between two sequences is calculated, where the distance is increased for each feature that does not match within the sequence.

7.2.3.2 *Weighted distance*

Due to the fact that developers of the Genesis Market are using Chromium based browsers for their fingerprint spoofing plugin, in this situation it

might be extra valuable to 'weigh' the features that have changed in recent versions of chrome heavier than other features. In Chrome version 72 up to version 83 in total 9 attributes changed their supported status, the weight was empirically set to 8. To calculate a weighted distance the *Weighted Hamming Distance* is used, this is similar to regular Hamming distance except that for certain weighted features their distance is multiplied by the weight.

7.3 RESULTS

Similar to the previous chapter, in total the test was executed 30 times using different combinations of browsers, versions and operating systems. Each test was scored on similarity with the most recent versions of each browser and within the 30 results, for a full overview see Table 6.1. The results are structured in the follow fashion, the numbers from the enumeration correspond with the subsections in this section:

1. For each test a sequence is visualised which browser features are supported and unsupported, the results are shown in Figure 7.1.
2. The results from (1) are validated with the database from "Can I Use", here is verified whether the feature is correctly reported as supported or unsupported, the results are shown in Figure 7.2.
3. Subsequently the similarity score between a sequence from the tests and a sequence from the "Can I Use" database can be computed. The similarity between the tests sequences and all sequences of the most recent browser versions is visualised, the results are shown in Figure 7.3.
4. Due to the similarities from (3) across different browsers a baseline figure was created in which the sequences from all of the most recent browser versions are pairwise scored on similarity, the results are shown in Figure 7.4.
5. Due to the inconsistencies in the results from (2) the sequences from the tests were also pairwise scored on similarity, the results are shown in Figure 7.5.

Now, the above points will be discussed in detail in the subsections below.

7.3.1 *Features*

The test consisted of automatic tests whether certain browser features are supported such as the *loading-lazy-attr* as described in Section 7.1. The full list is available in Appendix D, note that for this study it is not important what the functionality of the feature is: it only matters whether it can be tested automatically. For each browser 588 automatic feature support tests are executed. Because some attributes are tested multiple times, in total 274 unique attributes remain. The result of a single test is either `True` or `False`. If

any test for an attribute reports a `True` value, the feature is reported as *supported*. Other options are either *unsupported* or *unreported*. In total "Can I Use" provides information on 501 attributes after running the test 274 attributes contain at least some information. All attributes that are *unreported* for every browser are left out of the results. In Figure 7.1 is shown per browser which features are supported (in green), which features are unsupported (in red) and when the feature support is not reported (in orange). For every browser the column consisting of supported and unsupported attributes forms a distinctive *sequence*. The separated columns correspond to the groups of Table 6.1.

It can be observed that the sequence of each browser differs in most cases, however some browsers have the exact same sequence. These browsers can be grouped into subgroups in the exact same manner as Section 6.3.3. The full grouping of matching sequences is displayed in the last row of Figure 6.1 and displayed again in detail in Table 7.1.

7.3.2 *Verify Reported feature*

In Figure 7.1 is shown which features are supported in green and which are unsupported in red. Now these results can be validated with the "Can I Use" database in order to verify whether their value is actually correctly reported as *Supported* or *Unsupported*. This way it can be determined whether the browser 'lied' about any attributes. In Figure 7.2 can be seen that most attributes are correctly reported (in green), some attributes are incorrectly reported (in red). Finally for some attributes it could not be established whether the attribute was correctly reported due to partial support or the attribute was not reported at all. It can be seen that there are attributes incorrectly reported (in red) even though no spoofing is attempted (Group 1 and 2). The separated columns correspond to the groups of Table 6.1.

7.3.3 *Reported features pairwise scored on similarity*

The sequences of each test, as visualised in Figure 7.1, can now be scored on similarity with the supported features from all browsers. This way a profile is built in which it can be compared which real browser most closely corresponds to the browsers of the test. The correlation matrix in Figure 7.3 is the result from scoring the browser feature support sequences on similarity from each test (on the x axis) with those from the "Can I Use" database (on the y axis). The weighted Hamming distance is used to calculate the similarity score. The lighter the square is, the stronger the higher the similarity is between a browser from the test and a particular version from the database. In total all of the 30 testing browsers have been compared on similarity with all of the available browser from dating maximum 10 versions back from the "Can I Use" database, resulting in 30 x 103 similarity scores. For the unweighted correlation matrix, see Figure E.2.

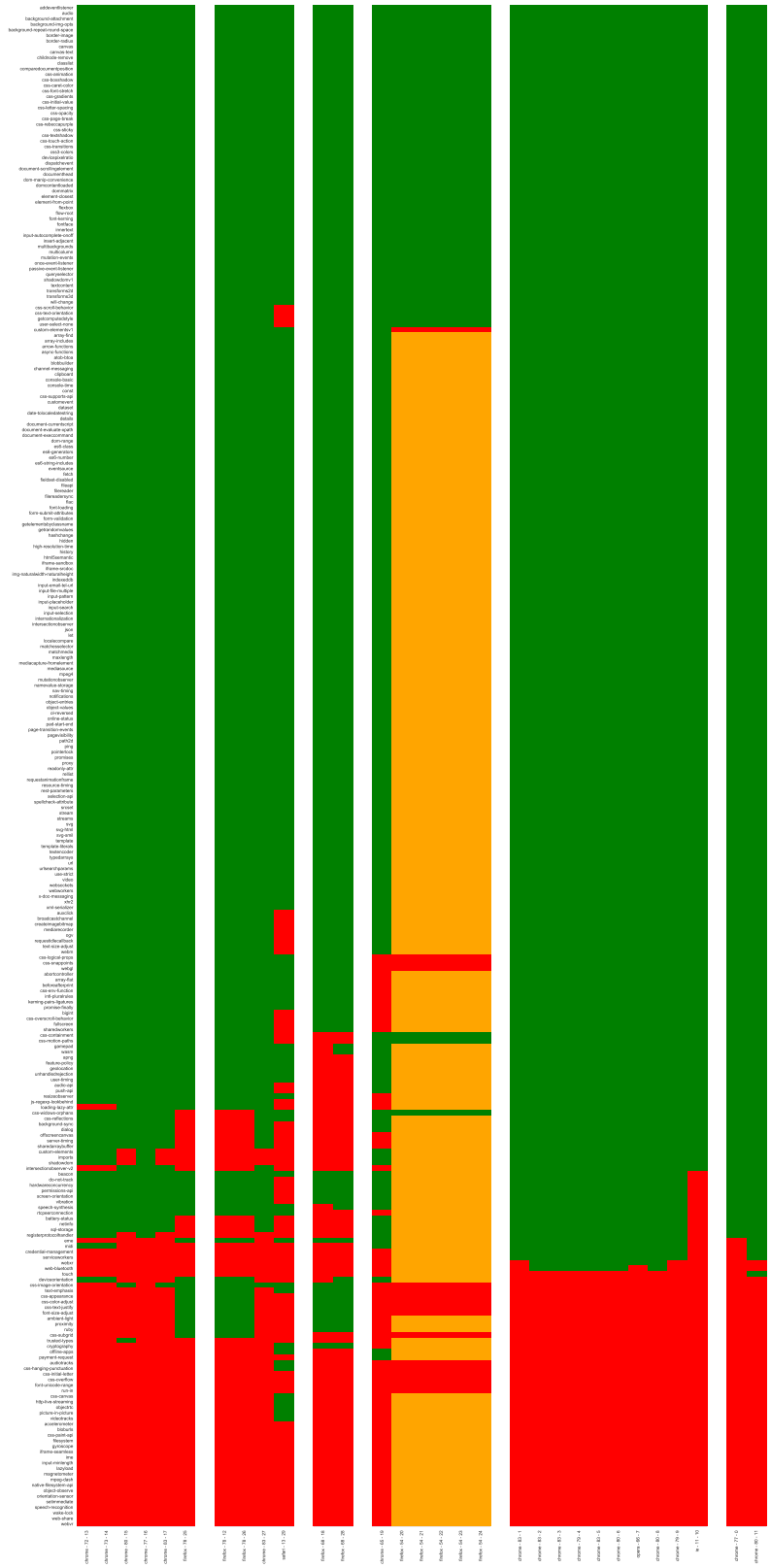

Figure 7.1: *Sequences* of supported features per browser:
Feature is supported
Feature is not supported
Feature is not reported
This figures shows the variations in sequences, however it can also be observed that some browsers have the exact same sequence. For an overview of which sequences match, see Table 7.1.

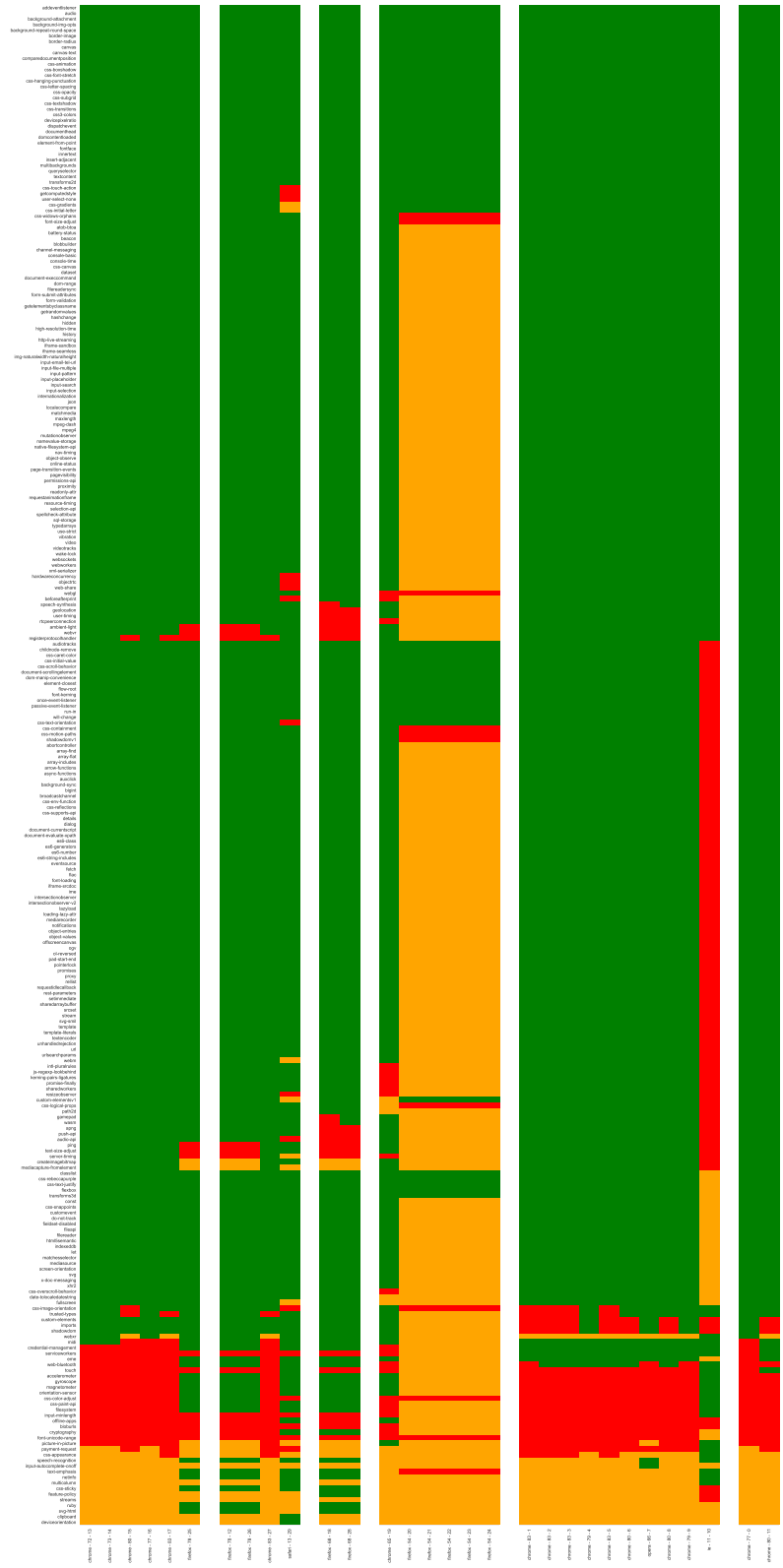

Figure 7.2: List of correctness of reported feature support per browser:
- Feature is reported as supported and supported
- Feature is reported as unsupported and unsupported
- Feature is reported as supported and unsupported
- Feature is reported as unsupported and supported
- Feature is reported as partially supported
- Feature is not reported

Most features are correctly reported, however some features are incorrectly reported even when no spoofing is attempted (Group 1 and 2).

It can be seen that the Chrome, Firefox and Safari browsers from group 1, 2 and 3 show a high resemblance to their respective browser from the "Can I Use" database. The result from Group 4 are very unreliable and the results from Group 5 and 6 show the highest resemblance to Chrome 77 regardless of their reported identified browser version.

7.3.4 *Browser comparison*

Based on the previous Section 7.3.3, it can be seen that Chrome browser from the test not only show a high resemblance to the Chrome browsers from the database but also to the Latest Opera and Edge browsers. In order to provide a baseline, first it is determined how the browsers from the "Can I Use" database score on similarity when compared with each other. For this the pairwise distance is calculated, based on the unweighted Hamming distance. In total all of the available browser from dating maximum 10 versions back from the "Can I Use" database have been compared on similarity with each other, resulting in 103 x 103 similarity scores.

The resulting correlation matrix, see Figure 7.4, shows that the latest versions of Edge, Chrome and Opera have a relatively high similarity score. Confirming that it is to be expected to observe similarities between Chrome, Opera and Edge in Figure 7.3 as well.

7.3.5 *Tests pairwise comparison*

From Figure 7.2 it is visible, in red, that even in the non spoofed groups 1 and 2 some features are wrongly reported as supported or unsupported. Meaning that the sequence reported by the test differs from the sequence reported by the "Can I Use" database. This discrepancy led to another visualisation where the tests are pairwise scored on similarity within the same set, resulting in a 30 x 30 correlation matrix.

Additionally it becomes clear from Figure 7.2 that group 4, the results from the Sphere browser, are very different from the rest of the results. Therefore group 4 has been left out the pairwise comparison, for the results including group 4 see Figure E.1. The similarity scores were calculated using the weighted Hamming distance. For the unweighted correlation matrix, see Figure E.3.

In Figure 7.5 can be seen that tests 0-11 from group 5 and 6 show a strong similarity to Chrome 77 from Group 1. It also becomes clear that the Safari browser from group 2 and Firefox (Tor) browsers from group 3 show the least similarity with any of the other browsers.

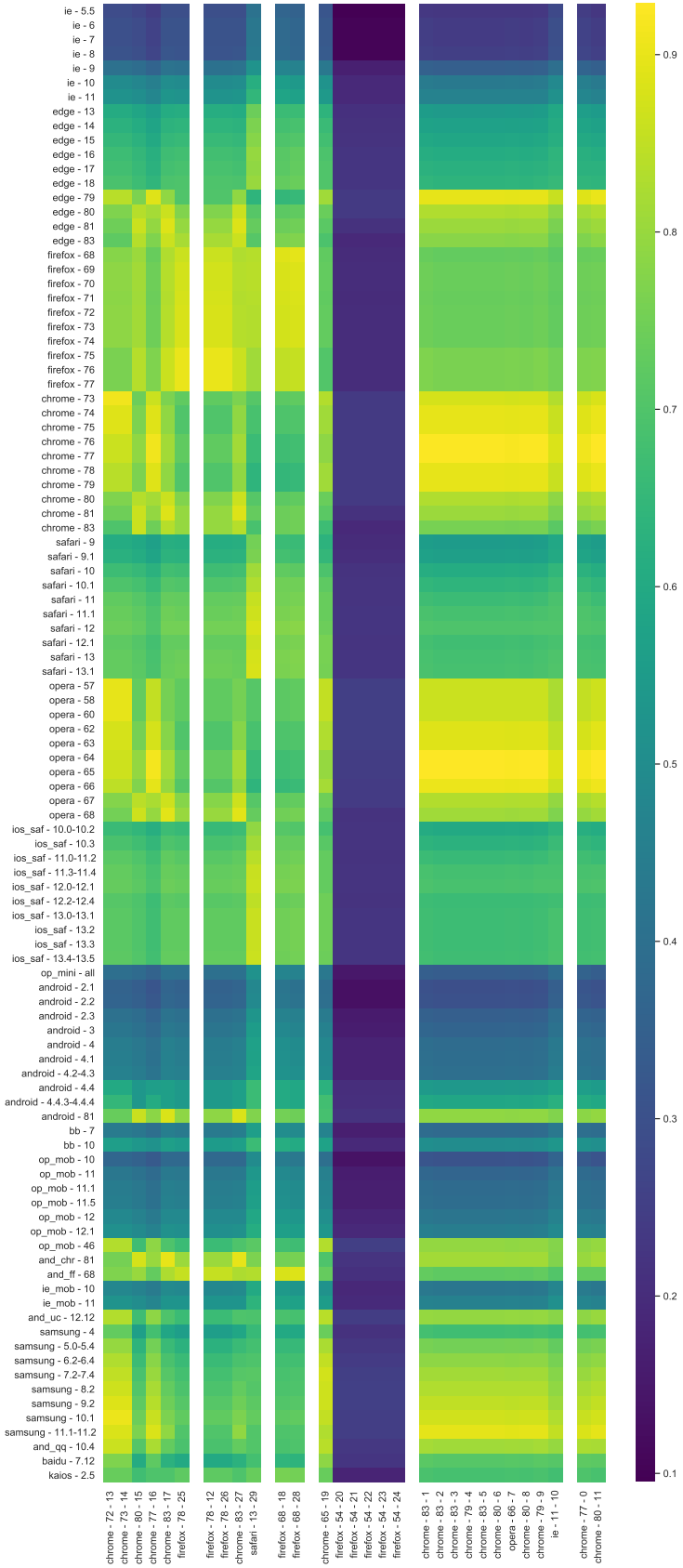

Figure 7.3: Correlation matrix with 30 browser test scored on similarity with 103 most recent browsers from the "Can I Use" database. Each square displays the similarity between those, the brighter the color the higher the similarity score is. To calculate the similarity score the *Weighted Hamming Distance* is used. Particular interest is taken in the fact that Group 5 and 6 (ID 0 - 11) show that all versions (regardless of their reported version), resemble Chrome 77 closer than their *identified browser*.

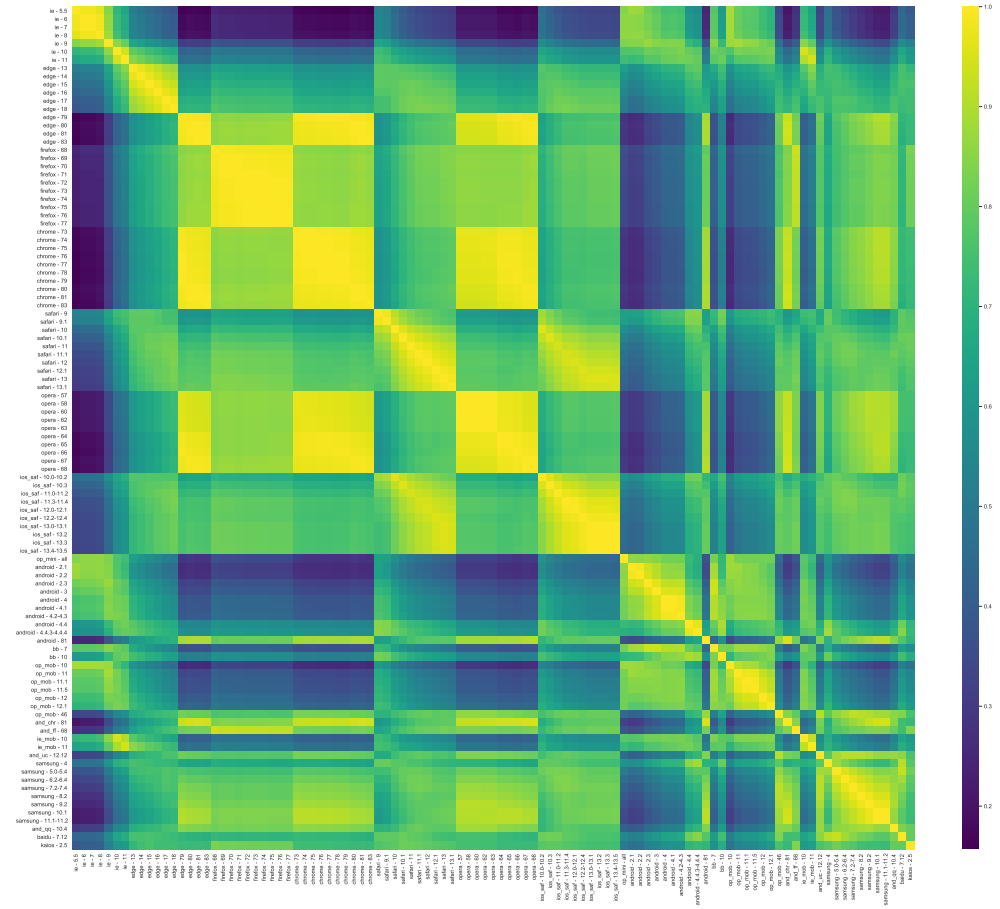

Figure 7.4: Correlation matrix which displays how similar specific browsers and versions are based on their feature support. The 103 browser from the "Can I Use" database have been pairwise scored on similarity. This provides a baseline on which browser versions are the most similar. Identical to Figure 7.3 the similarity score is calculated using the *Hamming Distance*. It can be observed that especially the latest versions of Edge, Chrome and Opera show a high resemblance.

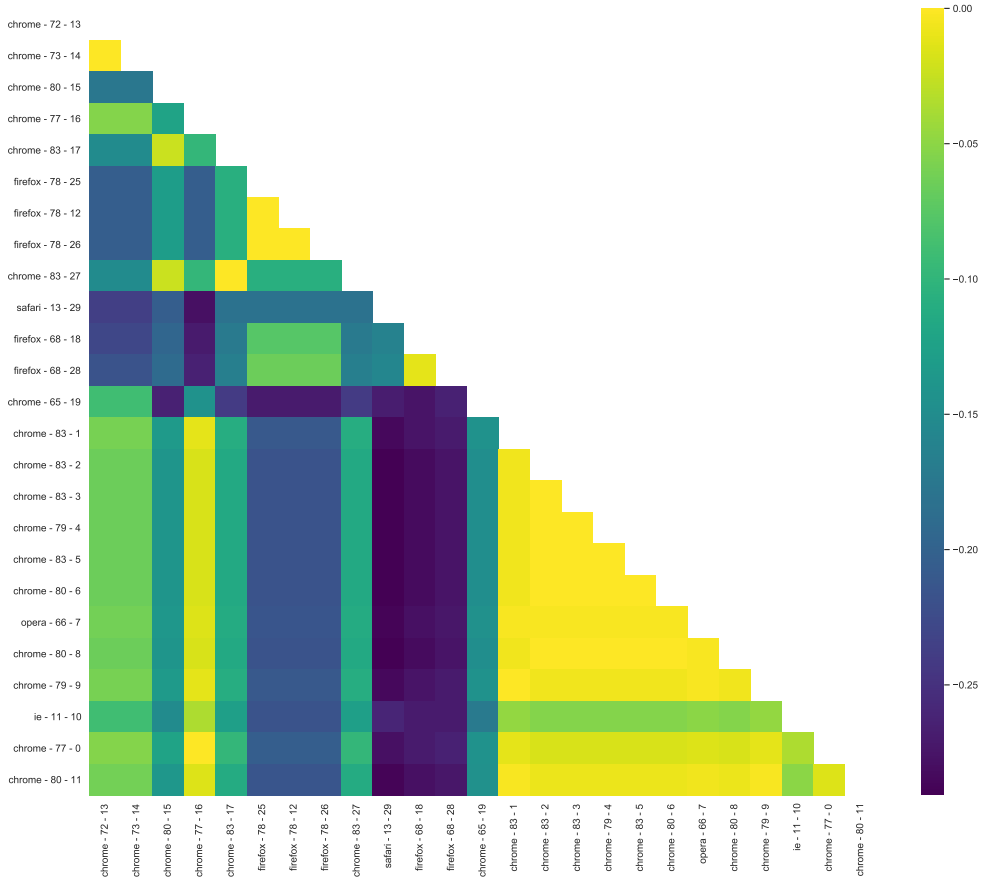

Figure 7.5: Correlation matrix which displays the similarities scores from the 30 test results compared pairwise. The *Weighted Hamming Distance* similarity scoring is used. Special attention can be focused on the column with ID 16 (Chrome 77) and rows with ID 0-11 (Group 5 and 6), this seems to indicate a high similarity with Chrome 77 (Group 1).

7.4 DISCUSSION

By observing the figures of Section 7.3, the following observations can be made:

- All Chromium based browsers are very similar in terms of feature support.
- Browsers are mostly very similar to the previous and next version of the same browser.
- The "Can I Use" database does not always provide correct information about browser feature support, and is therefore not suitable as a ground truth. When using a non spoofed browser there are still 'red squares' in Figure 7.2. Meaning that the feature was incorrectly reported as (un)supported compared to the information from the database.
- The spoofed browsers from group 5 and 6 are more similar to specific version from Chrome and Opera than they are to their own version.
- Pairwise comparison between browsers does seem to highlight the fact that the spoofed browsers from group 5 and 6 are actually Chrome 77.
- Feature support is exactly the same for Chrome 83 on both Windows and Mac, feature support for Chrome 77 is exactly the same as for the UnGoogled browser as for the Genesium browser.

Each column from from Figure 7.2 can be regarded as an unique identifiable sequence, similar to DNA sequencing. It can be seen that the sequence of Chrome 83 from group 1 and 2 are 100% similar and thus a match. The complete overview can be found in Table 7.1. It never occurs that browser from group 5 and 6 have the same supported feature sequence as the other browsers, except for subgroup 0 where Chrome 77 corresponds 100% to the unmodified Genesium browser which is based on Chromium 77. Even though the identified browser version is the same as one of the browsers for which a sequence was built, for example Chrome 83. It can be concluded that the Genesium browser fails to produce a correct sequence, and even produces the same result for many tests.

7.4.1 *Limitations*

Due to the incorrectly reported values of the "Can I Use" database compared to the 'real situation', this approach gives inaccurate results when scoring the browsers from the tests to the database on similarity. This is why the database from "Can I Use" cannot be used for 100% match scoring. Due to the fact that the differences between browser versions that are close to each other are very small, combined with the inaccurate results from the "Can I Use" database, similarity scoring the results does not give a strong visual cue which browser is which version exactly.

SUBGROUP	IDENTIFIED BROWSER - IDENTIFIER					
	Group 1	**Group 2**	**Group 3**	**Group 4**	**Group 5**	**Group 6**
Subgroup 0	Chrome 77 - 0					Chrome 77 - 16
Subgroup 1					Chrome 83 - 1 Chrome 79 - 9	
Subgroup 2					Chrome 83 - 2 Chrome 83 - 3 Chrome 79 - 4 Chrome 83 - 5 Chrome 80 - 6 Chrome 80 - 8	
Subgroup 3					Opera 66 - 7	
Subgroup 4					IE 11 - 10	
Subgroup 5						Chrome 80 - 11
Subgroup 6	Firefox 78 - 25	Firefox 78 - 12 Firefox 78 - 26				
Subgroup 7	Chrome 72 - 13 Chrome 73 - 14					
Subgroup 8	Chrome 80 - 15					
Subgroup 9	Chrome 83 - 17	Chrome 83 - 27				
Subgroup 10			Firefox 68 - 18			
Subgroup 11				Chrome 65 - 19		
Subgroup 12				Firefox 54 - 20 Firefox 54 - 21 Firefox 54 - 22 Firefox 54 - 23 Firefox 54 - 24		
Subgroup 13			Firefox 68 - 28			
Subgroup 14		Safari 13 - 29				

Table 7.1: Browsers categorised into groups and subgroups. Belonging to the same subgroups means that the compatibility sequence from Figure 7.1, is exactly the same for those browsers.

As shown in Figure 7.3 a non spoofed version of Chrome 83 in group 1 seems to closer resemble Chrome 81 and even closer Opera 68. However the spoofed version of Chrome 83 in the 5th group resembles Chrome 76 or 77 the closest or Opera 65, 66 or 67. This visible 'mismatch' might give an indication that the spoofed version is indeed not the version that it is claimed to be.

The close resemblance to both Chrome and Opera are due to the fact that those are both Chromium based browser which is visible in Figure 7.4.

Due to the many attributes that can be identified, to properly differentiate between the groups the *Curse of Dimensionality* arises. To correctly apply pattern recognition the rule of thumb is that for every dimension at least 5 training examples should be available [47]. However, with 274 attributes this leads to a high amount of required tests while performing one test is already a labour intense and costly operation.

Table 7.1 also shows that it is not possible to uniquely identify a browser and version based on the sequence alone, in subgroup 7 it can be seen that both Chrome 72 en Chrome 73 have the same sequence. Therefore, these versions can be seen as a digital twin and cannot be uniquely identified by their sequence. The sequence does seem to be unique for many versions, however the sample size is very limited to draw strong conclusions. What is also interesting to note is that Firefox 68 with identifier 18 and 28, which are both the Tor browser, do not belong to the same subgroup. Therefore the conclusion can be drawn that these browsers utilise spoofing, which is indeed the case [62].

7.4.2 *Suggestions*

In order to overcome the first limitation as described above the results have also been correlated within the testing set as shown in Figure 7.5. Here the stronger resemblance with Chrome 77 from group 1 is highlighted in comparison with other browsers which have been manually tested. However to make this assumption stronger non spoofing tests must be performed with all browser versions in order to be able to built an image that is closer to Figure 7.3. The database from "Can I Use" must be rebuilt for a complete and real situation.

Finally, it does strongly highlight that spoofing IE 11 is much easier to differentiate from other spoofing attempts, this was also very clear from Figure 7.2.

7.5 CONCLUSION

The hypothesis is that it is almost impossible to 100% correctly spoof another browser or version, this is due to the fact that browsers introduce some unique characteristics in each version with regards to feature support. It would be very costly for an attacker to keep updating browser fingerprints based on the version that is used by the infected client. This study indicates

that spoofing the browser feature support 100% correctly is impossible without using the same browser version as the one of the infected client the attacker tries to spoof.

Based on the results mainly from Figure 7.2 it shows that none of the sequences from group 5 are exactly the same as the non spoofed browser even though the version matches, so one could argue that spoofing failed because there is not a 100% match with the 'actual version'. In comparison to other sequences which do have the 100% expected match, based on this information it can be concluded that a fail to match the sequence means that there is some attribute spoofing.

Referring back to the research question 'How can we achieve protection from malicious use of anti-fingerprinting?' the method of additionally creating a sequence based on browser feature support is very promising. Unfortunately, due to inaccuracies in the "Can I Use" database a defensive mechanism could not be constructed immediately because this mechanism would rely on a 100% match. Additionally, due to the high similarities between some browser as reported in Section 7.3.3 and Section 7.3.4. A 100% match might occur when for example a Chrome browser spoofs a Chrome, Edge or Opera browser with a comparable version, this will lead to falsely suggesting that the browser is not spoofing (false negative) even though it is actually spoofing. However, more experiments are required to verify this: the sequences from all available browser versions must be compared in order to support these claims.

In order to be able to correctly tell apart a spoofed browser based on browser feature support more test results are required both with spoofed browsers as with non spoofed browsers. This way the "Can I Use" database can be verified and rebuilt. Using browser feature support in order to defend against anti-fingerprinting can be a part of a larger detection system which makes it very hard to correctly spoof different browser versions. If *real* sequences from all browser are available a defensive mechanism could be constructed based on a mechanism that verifies the sequence with the database and only allows it when there is a 100% match. Currently that is not possible because the sequences that match all 'non spoofed browsers' are not available. In conclusion the hypothesis from Section 7.1 cannot be proven to be incorrect, none of the spoofed instances was able to 100% accurately spoof another browser version. Additionally, protection can be achieved against anti-fingerprinting by allowing only 100% correct matches which results in a false negative in the worst case.

8

CONCLUSION

Browser fingerprinting has evolved into a way to perform illegitimate cross domain tracking on which legitimate anti-fingerprinting to preserve anonymity and privacy was the answer. Legitimate fingerprinting, as an extra step for identification and verification, was developed in order to combat fraud. This was answered by cybercriminals with illegitimate anti-fingerprinting, which is geared towards impersonation and fraud. Referring back to the main research questions 'To what extent do anti-fingerprinting products pose a threat within the cyber criminal landscape?', it can be concluded that it poses a threat to a great extent. The anti-fingerprinting product market is a sophisticated quick growing threat that threatens a large amount of victims from all over the world. The Genesis Market has sold millions of profiles over the course of the years, allowing cybercriminals to circumvent browser fingerprinting. As a proposed counter measure, this study concludes that by implementing additional fingerprinting there are ways to defend against these cybercriminal techniques. The method which verifies the browsers' feature sequence, gives a strong indication of spoofing. This can in turn improve state of the art browser fingerprint protection systems.

8.1 SUMMARY

In this section all conclusions from previous parts will be summarised and will answer all of the research questions. For a more in detail conclusion for each part please refer to the conclusion section of Chapter 4, Chapter 5, Chapter 6 and Chapter 7.

8.1.1 *State of the Market*

> *What is the current state of the anti-fingerprinting market, which products are offered and how do they compare?*

Through history multiple anti-fingerprinting services have risen in popularity, each of them aiming to increase the ease of use to impersonate the digital identity of their victims. The products differ in functionality and pricing model. The latest innovation is market places full of credentials which facilitate quick impersonation. Account Takeover as a Service (ATaaS) is born.

8.1.2 *Threat Analysis*

> *How large is the upcoming threat of anti-fingerprinting? Does it enable mass cybercrime or targeted attacks?*

The upcoming threat of anti-fingerprint has strongly increased past year. The Genesis Market grew to over 300.000 available profiles, because of the high turnover it is estimated that the Genesis Market affected between 1 to 3 million victims. The information that is stolen enables cybercriminals to quickly take control of someone's digital identity which enables fraud and impersonation. This is certainly suitable for targeted attacks.

8.1.3 *Anti Fingerprinting*

How do anti-fingerprinting solutions successfully circumvent browser fingerprinting?

While it is not possible to establish how anti-fraud systems which utilise fingerprinting are successfully circumvented, it can be established that the anti-fingerprinting solution Genesis Market successfully alters the fingerprint. While using the same browser, it is able to utilise different stolen digital identities. Because of to the high volume of digital identities traded on this platform the Genesis Market appears to be successful in circumventing browser fingerprinting, which allows their users to digitally impersonation their victims and enables them to commit fraud on their victims behalf.

8.1.4 *Protection*

How can we achieve protection from malicious use of anti-fingerprinting?

Based on 'feature support sequencing' spoofing browsers can be picked out because they fail to match the sequence of a specific version from a real browser. This approach leverages the fact that the developers of spoofing browsers cannot keep up with the continuous development of consumer browsers. Based on this research alone it is not possible to achieve complete protection from anti-fingerprinting since there are not enough sequences available. However, attribute sequencing shows real potential and can play a role in defending against anti-fingerprinting browsers by utilising the fact that each browser generates an almost unique sequence.

8.2 LIMITATIONS

Due to the criminal nature of anti-fingerprinting solutions it can be difficult to gather data and get insights into the inner workings. Therefore, this study is subject to several limitations.

- Getting insight into competitors of the Genesis Market: already great effort was performed in order to infiltrate the Genesis Market. To study one of the competitors, more or less the same operation must be performed to gain access to any other competitor as well.
- Accurately estimating the number of victims: all numbers in this study have been manually collected from time to time. By performing auto-

matic analysis a more accurate estimation could be provided. However, automated analysis might alarm the administrators and as a result access might be lost.

- Tracking what kind of crimes are committed with the anti-fingerprinting services: based upon articles, offered profiles and user activity on the forum an estimation can be made. Ideally, in cooperation with companies which are targeted by anti-fingerprinting an overview could be created of kind of crimes are committed. Based on conversations with experts we learned that companies are not very keen on sharing this type of information.
- Analysing the workings of the software: because the Genesis Market wants to protect their software from being inspected the software they provide is heavily obfuscated and cannot easily be reversed.
- Subjecting many anti-fingerprinting systems to the test system: similar to the limitation to investigate competitors it is difficult to gain access to all of the various anti-fingerprinting systems that are offered.
- Subjecting existing fingerprinting systems to the stolen fingerprints from the Genesis Market: this involves an ethical debate whether it should be allowed to use the stolen information from multiple victims in order to test how effective the anti-fingerprinting systems bypass existing fingerprint systems. And it would require cooperation from targeted services as well. However, this would be a very effective way to establish the posed threat.
- Comparing existing fingerprinting systems to a fingerprinting system that additionally extracts attribute sequences: because the sequences provided by "Can I Use" were not accurate enough to provide a 100% match, the proposed defence mechanism cannot be applied before a new database has been established with an updated attribute sequence for every browser that must be identified.

Concerning the protection against anti-fingerprinting fraud, based on expert opinions customer satisfaction and business continuity for companies such as banks, online merchandise and payment providers is more important then it is to block all fraud. If the defensive system does not work 99.9% of the time these companies rather tolerates some fraud than it blocks a true customer. Based on the limited experiments in this study this makes the defensive approach of Chapter 7 extra suitable, since it introduces a false negative in the worst case. Therefore not impacting business continuity as severe as would be the case with false positives. Luckily, many of these high profile targets have 2FA by default, rendering the ability to bypass fingerprint protection useless. However this does not mean all, therefore some companies are still vulnerable to this form of cybercrime.

8.3 FUTURE SUGGESTIONS

The browser feature support sequencing shows real potential to combat anti-fingerprinting browsers. In order to built a complete protection system, attribute sequences must be collected on all available browsers. Additionally, existing fingerprinting systems could be tested and compared to a system that additionally extracts the attribute sequence. Ideally tested against anti-fingerprinting browsers, that offer real digital identities of victims.

Part II

APPENDIX

A

EXAMPLES OF PRODUCTS AVAILABLE ON THE MARKET

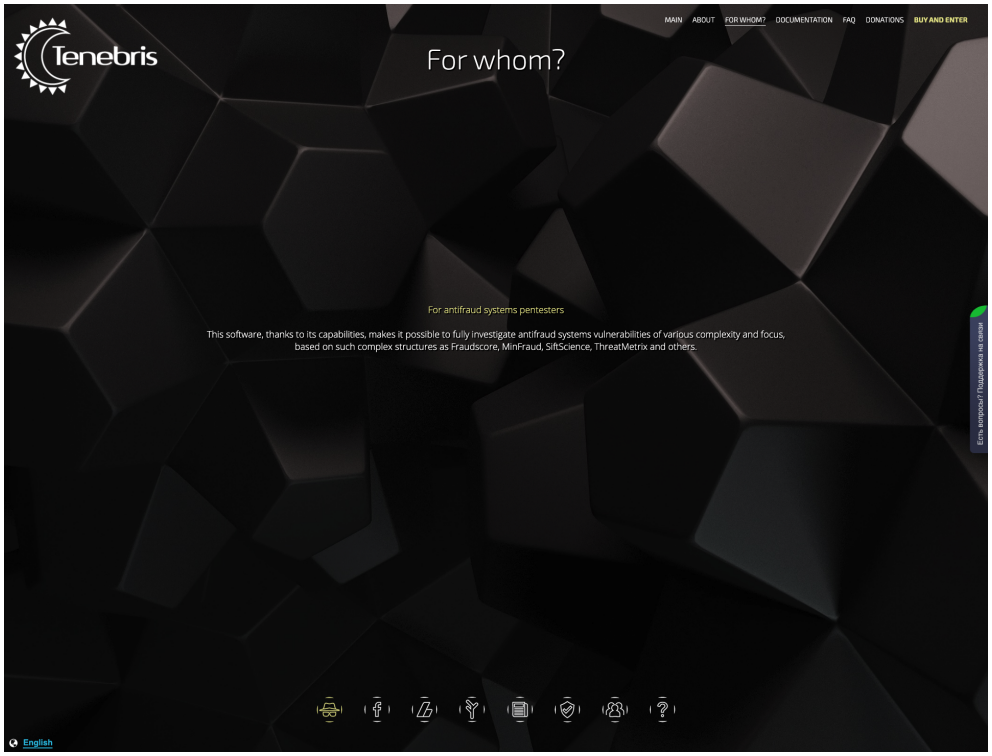

Figure A.1: Linken Sphere promotes ability to investigate antifraud systems.

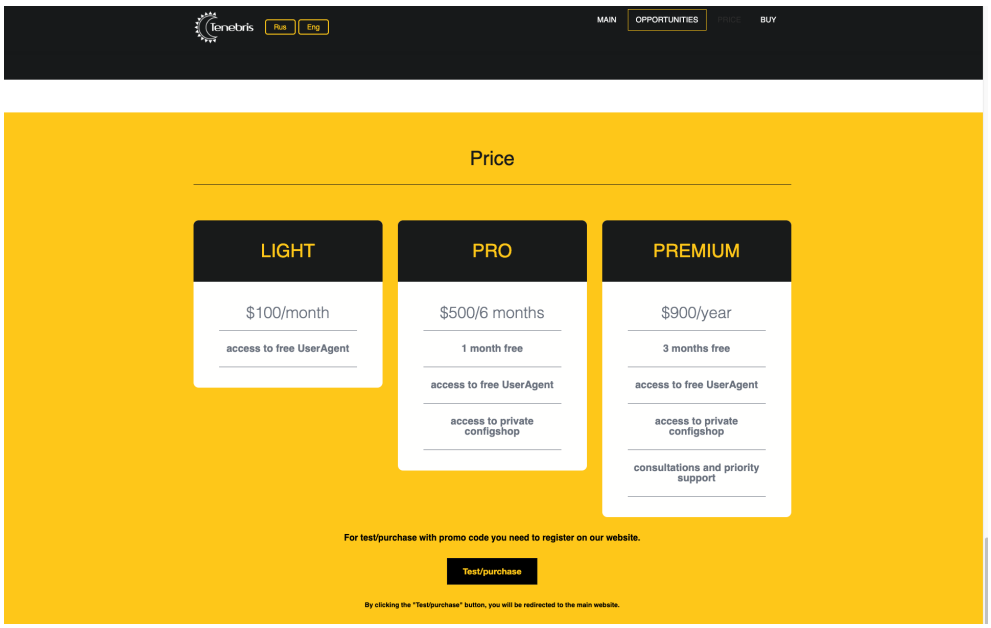

Figure A.2: Marketed at $100 per month, with no up front payment.

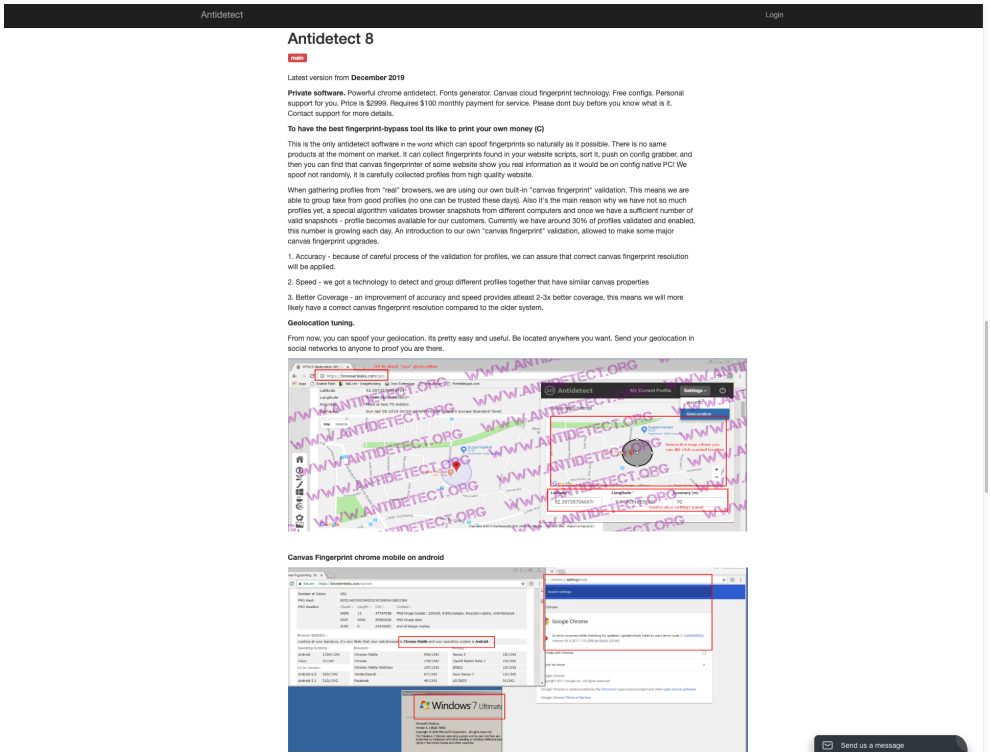

Figure A.3: Promotes the ability to 'Print your own money', marketed at $100 per month with an $2999 up front payment.

B

GENESIS MARKET ADVERTISEMENT

Interesting snippets of an advertisement of the Genesis Market. The text was found on a Russian darkweb forum [16], and translated from Russian. There are some mistakes in the translation, but still gives a good impression.

Note: due to the criminal intent of the Dedik forum be advised to never visit the url without protecting your identity. Using the Tor browser or an anonymous VPN + VM is highly recommended.

B.1 TRANSLATION

What prices?
All bots can be divided into 3 categories:
1. Bots with fingerprinting, cups and actresses - from $51
2. Bots with fingerprinting, no cook, no accounts - $50
3. Bots with accounts but without fingerprinting and cook - $15 - Temporarily removed from sale
The price of one finger-printer is $50, for the second fingerprinting of the same bot, from $10, and for the third, from $5, everything
the following, if available, free!

What does a plagine know?
For those interested in technical details, and in order to answer a number of questions beforehand,
to change the plagine, enumerating only the main points:
1. Screen! brown! and complete change
2. Navigator! brown! and complete change
3. Window! brown! the change
4. Document! brown! the change
5. WebGL
6. IE components & ActiveX
7. Permissions
8. Fonts - Only our replacement is technically literate and correct
9. Geolocation - Even with the same temporary delay as the holder
10. Cookies
11. Canvas
12. Web Audio - Who knows, will understand
13. WebRTC
14. Headers - Correct replacement of hikers!
15. Security Headers
16. JS versions
17. CSS @media
18. Navigator Cores
19. hidden*
20. hidden*

21. hidden*
22. hidden*
23. hidden*
24. hidden*
25. hidden*

* This is not randomized paragraphs; we mean clear points that no one has on the market. We would be very happy to tell the most important chips of our hidden system, which allow you to work unnoticed and precisely as the detective systems need... but not to do so at this time

C

FULL LIST OF FINGERPRINTED ATTRIBUTES

webdriver
language
colorDepth
deviceMemory
pixelRatio
hardwareConcurrency
screenResolution
availableScreenResolution
timezoneOffset
timezone
sessionStorage
localStorage
indexedDb
addBehavior
openDatabase
cpuClass
platform
doNotTrack
canvas
webgl
webglVendorAndRenderer
adBlock
hasLiedLanguages
hasLiedResolution
hasLiedOs
hasLiedBrowser
touchSupport
enumerateDevices
plugins
fonts
audio

FULL LIST OF FEATURE SUPPORT

abortcontroller
accelerometer
addeventlistener
ambient-light
apng
array-find
array-flat
array-includes
arrow-functions
async-functions
atob-btoa
audio
audio-api
audiotracks
auxclick
background-attachment
background-img-opts
background-repeat-round-space
background-sync
battery-status
beacon
beforeafterprint
bigint
blobbuilder
bloburls
border-image
border-radius
broadcastchannel
canvas
canvas-text
channel-messaging
childnode-remove
classlist
clipboard
comparedocumentposition
console-basic
console-time
const
createimagebitmap
credential-management
cryptography
documenthead
dom-manip-convenience
dom-range
domcontentloaded
dommatrix
element-closest
element-from-point
eme
es6-class
es6-generators
es6-number
es6-string-includes
eventsource
feature-policy
fetch
fieldset-disabled
fileapi
filereader
filereadersync
filesystem
flac
flexbox
flow-root
font-kerning
font-loading
font-size-adjust
font-unicode-range
fontface
form-submit-attributes
form-validation
fullscreen
gamepad
geolocation
getcomputedstyle
getelementsbyclassname
getrandomvalues
gyroscope
hardwareconcurrency
hashchange
hidden
high-resolution-time
object-observe
object-values
objectrtc
offline-apps
offscreencanvas
ogv
ol-reversed
once-event-listener
online-status
orientation-sensor
pad-start-end
page-transition-events
pagevisibility
passive-event-listener
path2d
payment-request
permissions-api
picture-in-picture
ping
pointerlock
promise-finally
promises
proximity
proxy
push-api
queryselector
readonly-attr
registerprotocolhandler
rellist
requestanimationframe
requestidlecallback
resizeobserver
resource-timing
rest-parameters
rtcpeerconnection
ruby
run-in
screen-orientation
selection-api
server-timing
serviceworkers

Continued on the next page...

css-animation
css-appearance
css-boxshadow
css-canvas
css-caret-color
css-color-adjust
css-containment
css-env-function
css-font-stretch
css-gradients
css-hanging-punctuation
css-image-orientation
css-initial-letter
css-initial-value
css-letter-spacing
css-logical-props
css-motion-paths
css-opacity
css-overflow
css-overscroll-behavior
css-page-break
css-paint-api
css-rebeccapurple
css-reflections
css-scroll-behavior
css-snappoints
css-sticky
css-subgrid
css-supports-api
css-text-justify
css-text-orientation
css-textshadow
css-touch-action
css-transitions
css-widows-orphans
css3-colors
custom-elements
custom-elementsv1
customevent
dataset
date-tolocaledatestring
details
deviceorientation
devicepixelratio
dialog
dispatchevent
do-not-track
document-currentscript
document-evaluate-xpath
document-execcommand
document-scrollingelement
history
html5semantic
http-live-streaming
iframe-sandbox
iframe-seamless
iframe-srcdoc
ime
img-naturalwidth-naturalheight
imports
indexeddb
innertext
input-autocomplete-onoff
input-email-tel-url
input-file-multiple
input-minlength
input-pattern
input-placeholder
input-search
input-selection
insert-adjacent
internationalization
intersectionobserver
intersectionobserver-v2
intl-pluralrules
js-regexp-lookbehind
json
kerning-pairs-ligatures
lazyload
let
loading-lazy-attr
localecompare
magnetometer
matchesselector
matchmedia
maxlength
mediacapture-fromelement
mediarecorder
mediasource
midi
mpeg-dash
mpeg4
multibackgrounds
multicolumn
mutation-events
mutationobserver
namevalue-storage
native-filesystem-api
nav-timing
netinfo
notifications
object-entries
setimmediate
shadowdom
shadowdomv1
sharedarraybuffer
sharedworkers
speech-recognition
speech-synthesis
spellcheck-attribute
sql-storage
srcset
stream
streams
svg
svg-html
svg-smil
template
template-literals
text-emphasis
text-size-adjust
textcontent
textencoder
touch
transforms2d
transforms3d
trusted-types
typedarrays
unhandledrejection
url
urlsearchparams
use-strict
user-select-none
user-timing
vibration
video
videotracks
wake-lock
wasm
web-bluetooth
web-share
webgl
webm
websockets
webvr
webworkers
webxr
will-change
x-doc-messaging
xhr2
xml-serializer

E

ADDITIONAL CORRELATION FIGURES

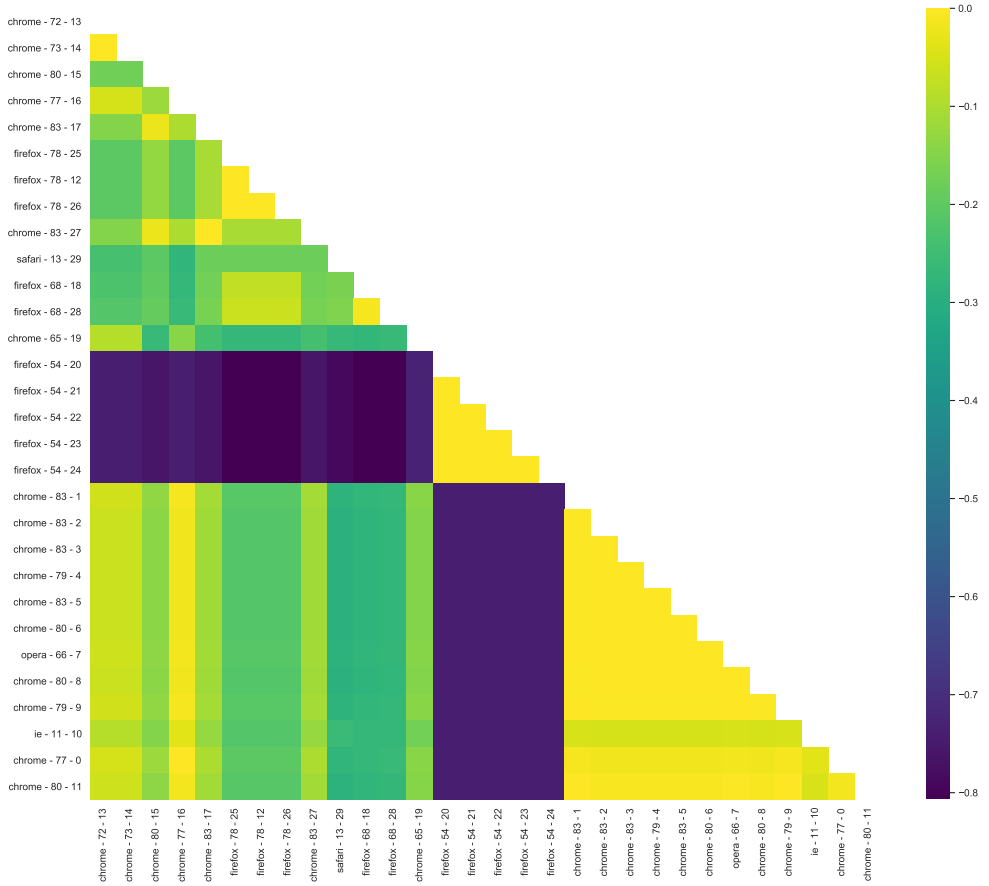

Figure E.1: Correlation matrix which displays the similarities scores from the 30 test results compared pairwise. The *Weighted Hamming Distance* similarity scoring is used. Special attention can be focused on the column with ID 16 (Chrome 77) and rows with ID 0-11 (Group 5 and 6), this seems to indicate a high similarity with Chrome 77 (Group 1). However, the results from Group 4 'pollute' the results. The similarity described above is still visible however it is less clear than in Figure 7.5.

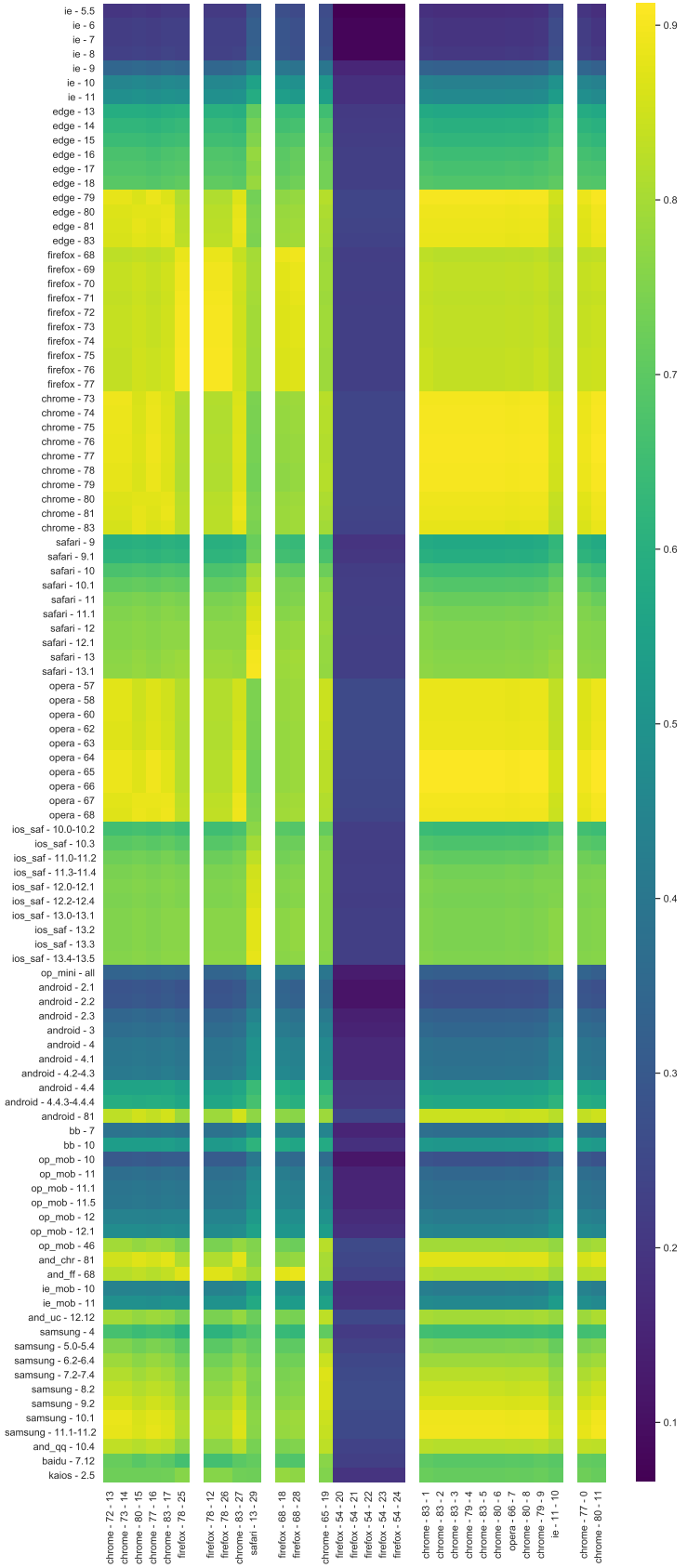

Figure E.2: Correlation matrix with 30 browser test scored on similarity with 103 most recent browsers from the "Can I Use" database. Each square displays the similarity between those, the brighter the color the higher the similarity score is. To calculate the similarity score the *Hamming Distance* is used. Particular interest is taken in the fact that Group 5 and 6 (ID 0 - 11) show that all versions (regardless of their reported version), resemble Chrome 77 closer than their *identified browser*. However, due to the fact that the unweighted distance method is used here this resemblance is visually less strong than it is in Figure 7.3.

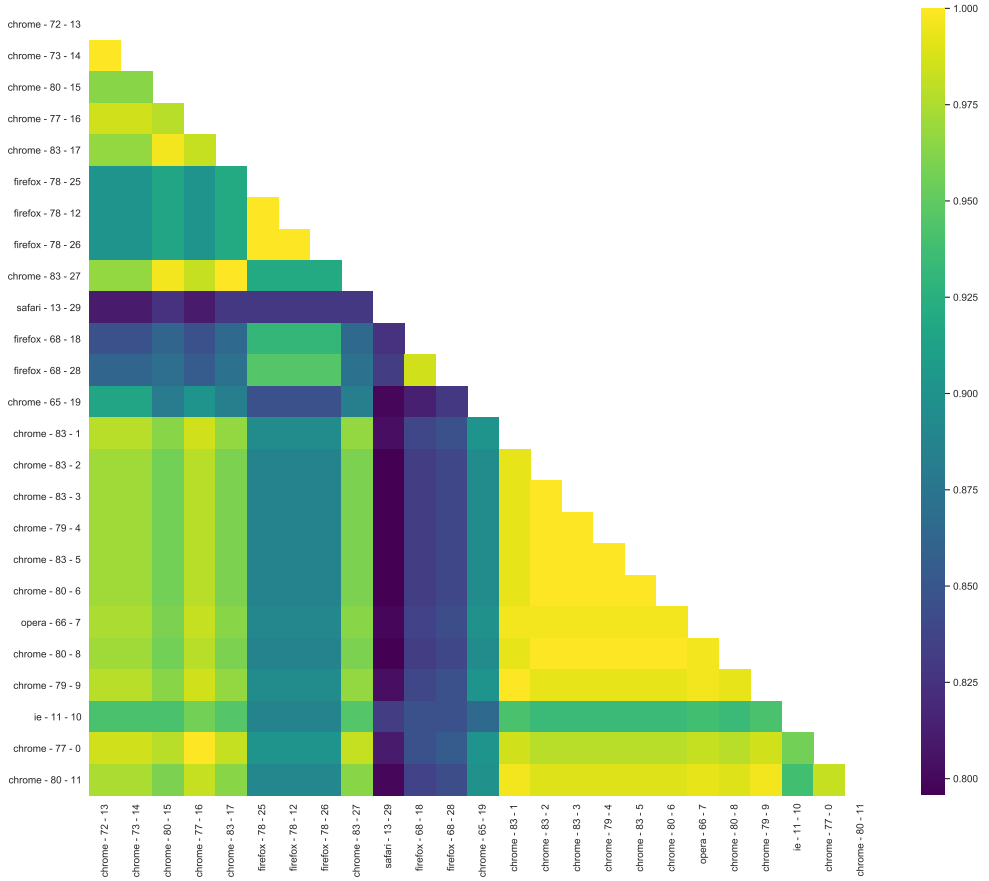

Figure E.3: Correlation matrix which displays the similarities scores from the 30 test results compared pairwise. The *Unweighted Hamming Distance* similarity scoring is used. Special attention can be focused on the column with ID 16 (Chrome 77) and rows with ID 0-11 (Group 5 and 6), this seems to indicate a higher similarity with Chrome 77 (Group 1). However, due to the fact that the unweighted distance method is used here this resemblance is visually less strong than it is in Figure 7.5.

BIBLIOGRAPHY

[1] *Aanwijzing opsporingsbevoegdheden*. Accessed: 15.07.2020. Sept. 2014. URL: https://wetten.overheid.nl/jci1.3:c:BWBR0035498.

[2] Gunes Acar, Marc Juarez, Nick Nikiforakis, Claudia Diaz, Seda Gürses, Frank Piessens, and Bart Preneel. "FPDetective: dusting the web for fingerprinters." In: *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*. 2013, pp. 1129–1140.

[3] Ariel Ainhoren. *Digital Browser Identities: The Hottest New Black Market Good*. 2019.

[4] N.M. Al-Fannah, W. Li, and C.J. Mitchell. "Beyond Cookie Monster Amnesia: Real World Persistent Online Tracking." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11060 LNCS (2018). cited By 0, pp. 481–501. DOI: 10.1007/978-3-319-99136-8_26. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053994382&doi=10.1007%2f978-3-319-99136-8_26&partnerID=40&md5=4ce0fcbe04bd6da0873e97620c82e835.

[5] P. Baumann, S. Katzenbeisser, M. Stopczynski, and E. Tews. "Disguised chromium browser: Robust browser, flash and canvas fingerprinting protection." In: cited By 4. 2016, pp. 37–46. DOI: 10.1145/2994620.2994621. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84998704828&doi=10.1145%2f2994620.2994621&partnerID=40&md5=ce36ec1455af02449e0d36a7366f70d8.

[6] Károly Boda, Ádám Máté Földes, Gábor György Gulyás, and Sándor Imre. "User tracking on the web via cross-browser fingerprinting." In: *Nordic conference on secure it systems*. Springer. 2011, pp. 31–46.

[7] S. Bodoarca and M.-L. Pura. "Assuring privacy in surfing the internet." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 11359 LNCS (2019). cited By 0, pp. 185–203. DOI: 10.1007/978-3-030-12942-2_15. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85062975452&doi=10.1007%2f978-3-030-12942-2_15&partnerID=40&md5=d160cac83c80e94788f1f3d5b0f4b9f0.

[8] Under The Breach. *Genesis Market 2020 overview, a bazaar for buying data out of compromised computers.* Accessed: 25.02.2020. Feb. 2020. URL: https://medium.com/@underthebreach/genesis-market-2020-overview-a-bazaar-for-buying-data-out-of-compromised-computers-85b581b903ec.

[9] The CyberWire. *Inside Magecart and Genesis.* [Audio podcast]. Dec. 2019. URL: https://thecyberwire.com/podcasts/research-saturday/115/notes.

[10] A. Datta, J. Lu, and M.C. Tschantz. "Evaluating anti-fingerprinting privacy enhancing technologies." In: cited By 0. 2019, pp. 351–362. DOI: 10.1145/3308558.3313703. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066915370&doi=10.1145%2f3308558.3313703&partnerID=40&md5=12d6fd222f4d366bab7ec9d276d8e758.

[11] P. Eckersley. "How unique is your web browser?" In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 6205 LNCS (2010). cited By 229, pp. 1–18. DOI: 10.1007/978-3-642-14527-8_1. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-77955457549&doi=10.1007%2f978-3-642-14527-8_1&partnerID=40&md5=3a72caee41c2bd7bb71b49a6bb34b8fb.

[12] S. Englehardt and A. Narayanan. "Online tracking: A 1-million-site measurement and analysis." In: vol. 24-28-October-2016. cited By 115. 2016, pp. 1388–1401. DOI: 10.1145/2976749.2978313. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84995395759&doi=10.1145%2f2976749.2978313&partnerID=40&md5=04dde1dff554ccaf425cf6b32fb70efd.

[13] Amin FaizKhademi, Mohammad Zulkernine, and Komminist Weldemariam. "FPGuard: Detection and prevention of browser fingerprinting." In: *IFIP Annual Conference on Data and Applications Security and Privacy*. Springer. 2015, pp. 293–308.

[14] Richard Feynman. *The Relation of Physics to Other Sciences*. volume I, lecture 3.

[15] FingerprintJS. *Github of FingerprintJS2*. Accessed: 09.03.2020. 2020. URL: https://github.com/fingerprintjs/fingerprintjs2.

[16] GenesisStore. *Genesis Store | Sell bots | #Fingerprints#Cookie#Logs*. Aug. 2019. URL: hxxp://dedik[.]cc/topic/6143-%D0%BC%D0%B0%D0%B3%D0%B0%D0%B7%D0%B8%D0%BD-genesis-storesell-botsfingerprintscookielogs/?tab=comments#comment-37078.

[17] A. Gómez-Boix, D. Frey, Y.-D. Bromberg, and B. Baudry. "A collaborative strategy for mitigating tracking through browser fingerprinting." In: cited By 0. 2019, pp. 67–78. DOI: 10.1145/3338468.3356828. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85076095599&doi=10.1145%2f3338468.3356828&partnerID=40&md5=51dec3ed18056ae8e4bbe6f4518d588f.

[18] ONR Haalstra. "Browser Fingerprinting and the application ofCountermeasures." Research Topics. University of Twente, 2020.

[19] Y. Haga, Y. Takata, M. Akiyama, and T. Mori. "Building a scalable web tracking detection system: Implementation and the empirical study." In: *IEICE Transactions on Information and Systems* E100D.8 (2017). cited By 0, pp. 1663–1670. DOI: 10.1587/transinf.2016ICP0020. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85026509498&

doi=10.1587%2ftransinf.2016ICP0020&partnerID=40&md5=0ce5902a9c0340ae259652d4ee335827.

[20] Satori Threat Intelligence and Research Team. *The Cybercrime Starter Kit: Inside Anti-Detection Browsers and Account Takeovers*. Feb. 2020. URL: https://www.whiteops.com/blog/the-cybercrime-starter-kit-inside-anti-detection-browsers.

[21] Kaspersky. *Financial cyberthreats in 2020*. Accessed: 27.02.2020. Jan. 2020. URL: https://www.brighttalk.com/webcast/15591/381190?utm_source=twitter&utm_medium=social&utm_campaign=us_webinar_zt0106&utm_content=sm-post&utm_term=us_twitter__zt0106_sm-post_social_webinar.

[22] SecureList Kaspersky. *Digital Doppelgangers: Cybercriminals cash out money using stolen digital identities*. Accessed: 25.02.2020. Apr. 2019. URL: https://securelist.com/digital-doppelgangers/90378/.

[23] Jerimy Kirk. *This tool may make it easier for thieves to empty bank accounts*. Accessed: 10.09.2020. Jan. 2015. URL: https://www.computerworld.com/article/2871926/this-tool-may-make-it-easier-for-thieves-to-empty-bank-accounts.html.

[24] Anna Kobusinska, Jerzy Brzezinski, and Kamil Pawulczuk. "Device Fingerprinting: Analysis of Chosen Fingerprinting Methods." In: *IoTBDS*. 2017, pp. 167–177.

[25] A. Kobusińska, K. Pawluczuk, and J. Brzeziński. "Big Data fingerprinting information analytics for sustainability." In: *Future Generation Computer Systems* 86 (2018). cited By 1, pp. 1321–1337. DOI: 10.1016/j.future.2017.12.061. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85040440837&doi=10.1016%2fj.future.2017.12.061&partnerID=40&md5=0896e12d2a1002112cee209750af57d5.

[26] Leon Kurolapnik. *What's Dead May Never Die: AZORult Infostealer Decommissioned Again*. Feb. 2020. URL: https://ke-la.com/whats-dead-may-never-die-azorult-infostealer-decommissioned-again/.

[27] Raveed Laeb. *Exploring the Genesis Supply Chain for Fun and Profit*. Feb. 2020. URL: https://ke-la.com/exploring-the-genesis-supply-chain-for-fun-and-profit/.

[28] John Leyden. *Number of stolen credentials on cybercrime marketplaces quadruples in just two years*. July 2020. URL: https://portswigger.net/daily-swig/number-of-stolen-credentials-on-cybercrime-marketplaces-quadruples-in-just-two-years.

[29] Sakchan Luangmaneerote, Ed Zaluska, and Leslie Carr. "Survey of existing fingerprint countermeasures." In: *2016 International Conference on Information Society (i-Society)*. IEEE. 2016, pp. 137–141.

[30] David JC MacKay and David JC Mac Kay. *Information theory, inference and learning algorithms*. Cambridge university press, 2003.

[31] SC Magazine. *Russian cybercriminals' most popular anonymization tools include Linken Sphere, Whatleaks*. Accessed: 08.01.2020. Oct. 2017. URL: https://www.scmagazine.com/home/security-news/cybercrime/russian-cybercriminals-most-popular-anonymization-tools-include-linken-sphere-whatleaks/.

[32] Paul Marks. *Dark web's doppelgängers aim to dupe antifraud systems*. 2020.

[33] Vereniging van waterbedrijven in Nederland. *Drinkwater voorzieningsgebieden per bedrijf*. Accessed: 10.03.2020. URL: https://www.vewin.nl/_layouts/15/vewin/DetailPageApplication/index.html.

[34] N. Nikiforakis, W. Joosen, and B. Livshits. "PriVaricator: Deceiving Fingerprinters with little white lies." In: cited By 43. 2015, pp. 820–830. DOI: 10.1145/2736277.2741090. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84968765165&doi=10.1145%2f2736277.2741090&partnerID=40&md5=3769c3afce0008eea1d23cfdbcc3c2bc.

[35] NucleairNederland. *Overzicht nucleaire bedrijven*. Accessed: 12.03.2020. URL: https://www.nucleairnederland.nl/over-ons/.

[36] Lindsey O'Donnell. *SAS 2019: Genesis Marketplace Peddles 60K Stolen Digital Identities*. Apr. 2019. URL: https://threatpost.com/genesis-marketplace-digital-identities/143558/.

[37] Ł. Olejnik, G. Acar, C. Castelluccia, and C. Diaz. "The leaking battery: A privacy analysis of the HTML5 battery status API." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9481 (2016). cited By 10, pp. 254–263. DOI: 10.1007/978-3-319-29883-2_18. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84961151977&doi=10.1007%2f978-3-319-29883-2_18&partnerID=40&md5=8785608f540fa835a1091be6637dbb6a.

[38] Jarrod Overson. *The State of Credential Stuffing and the Future of Account Takeovers.* Accessed: 25.02.2020. Sept. 2019. URL: https://www.slideshare.net/JarrodOverson/the-state-of-credential-stuffing-and-the-future-of-account-takeovers.

[39] Redacted. *Redacted*. Redacted.

[40] Assaf Dahan & Lior Rochberger. *Hunting Raccoon: The New Masked Bandit on the Block*. Oct. 2019. URL: https://www.cybereason.com/blog/hunting-raccoon-stealer-the-new-masked-bandit-on-the-block.

[41] Sector035. *Certificates: The OSINT Gift that Keeps on Giving...* Accessed: 19.03.2020. Mar. 2019. URL: https://osintcurio.us/2019/03/12/certificates-the-osint-gift-that-keeps-on-giving/.

[42] Bank Info Security. *For Sale on Cybercrime Markets: Real 'Digital Fingerprints'*. Accessed: 25.02.2020. Aug. 2019. URL: https://www.bankinfosecurity.com/for-sale-on-cybercrime-markets-real-digital-fingerprints-a-12943.

[43] Bank Info Security. *Analyst Insights on Genesis Market*. Accessed: 15.07.2020. June 2020. URL: https://www.bankinfosecurity.com/analyst-insights-on-genesis-market-a-14504.

[44] Krebs on Security. *'AntiDetect' Helps Thieves Hide Digital Fingerprints*. Accessed: 22.12.2019. Mar. 2015. URL: https://krebsonsecurity.com/2015/03/antidetect-helps-thieves-hide-digital-fingerprints/.

[45] Alexander Sukharenko. *Russian ITC Security Policy and Cybercrime*. Accessed: 01.09.2020. July 2019. URL: https://www.ponarseurasia.org/memo/russian-itc-security-policy-and-cybercrime.

[46] TenderNed. *Overzicht aanbestedingen Voorraadvorming Aardolieproducten*. Accessed: 13.03.2020. URL: https://www.tenderned.nl/tenderned-tap/aankondigingen/142590;section=1#detail-publicatie:linkS5.

[47] Sergios Theodoridis and Konstantinos Koutroumbas. "Pattern recognition. 2003." In: *Google Scholar Google Scholar Digital Library Digital Library* (2009).

[48] Daniel R Thomas, Sergio Pastrana, Alice Hutchings, Richard Clayton, and Alastair R Beresford. "Ethical issues in research using datasets of illicit origin." In: *Proceedings of the 2017 Internet Measurement Conference*. 2017, pp. 445–462.

[49] C.F. Torres, H. Jonker, and S. Mauw. "FP-block: Usable web privacy by controlling browser fingerprinting." In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9327 (2015). cited By 13, pp. 3–19. DOI: 10.1007/978-3-319-24177-7_1. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84951810025&doi=10.1007%2f978-3-319-24177-7_1&partnerID=40&md5=78e7ed1b1a6602d29bf7eeb8a76dfe2f.

[50] Bill Toulas. *Digital Browser Identities On Richlogs is the Hottest Thing On Dark Web Right Now*. Aug. 2019. URL: https://www.technadu.com/digital-browser-identities-richlogs-dark-web/77309/.

[51] Unkown. *Genesis Market*. Accessed:06.04.2020. Nov. 2017. URL: hxxps://genesis[.]market.

[52] Can I Use. *"Can I use" provides up-to-date browser support tables for support of front-end web technologies on desktop and mobile web browsers.* Accessed: 09.03.2020. 2014. URL: https://caniuse.com.

[53] Can I Use. *"Lazy loading via attribute for images & iframes.* Accessed: 30.09.2020. 2014. URL: https://caniuse.com/loading-lazy-attr.

[54] A. Vastel, P. Laperdrix, W. Rudametkin, and R. Rouvoy. "FP-STALKER: Tracking Browser Fingerprint Evolutions." In: vol. 2018-May. cited By 10. 2018, pp. 728–741. DOI: 10.1109/SP.2018.00008. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85050906822&doi=10.1109%2fSP.2018.00008&partnerID=40&md5=f3d042c8a81e930e321921ce63d1008c.

[55] A. Vastel, W. Rudametkin, and R. Rouvoy. "FP-TESTER : Automated Testing of Browser Fingerprint Resilience." In: cited By 0. 2018, pp. 103–107. DOI: 10.1109/EuroSPW.2018.00020. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85050922459&doi=10.1109%2fEuroSPW.2018.00020&partnerID=40&md5=7ffa93af757e0e954e2e6b285b3557c5.

[56] Nationaal Coördinator Terrorismebestrijding en Veiligheid. *Overzicht vitale processen*. Accessed: 09.03.2020. URL: https://www.nctv.nl/onderwerpen/vitale-infrastructuur/overzicht-vitale-processen.

[57] Jai Vijayan. *Study Finds 15 Billion Stolen, Exposed Credentials in Criminal Markets*. July 2020. URL: https://www.darkreading.com/attacks-breaches/study-finds-15-billion-stolen-exposed-credentials-in-criminal-markets/d/d-id/1338309.

[58] *Wetboek van Strafrecht*. Accessed: 30.05.2020. Jan. 2020. URL: https://wetten.overheid.nl/jci1.3:c:BWBR0001854.

[59] *Wetboek van Strafvordering*. Accessed: 15.07.2020. Sept. 2014. URL: https://wetten.overheid.nl/jci1.3:c:BWBR0001903.

[60] T. Yamada, T. Saito, K. Takasu, and N. Takei. "Robust Identification of Browser Fingerprint Comparison Using Edit Distance." In: cited By 2. 2015, pp. 107–113. DOI: 10.1109/BWCCA.2015.106. URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84964925978&doi=10.1109%2fBWCCA.2015.106&partnerID=40&md5=008ad66fbe749ebfc1f6f88c97e69577.

[61] energieleveranciers.nl. *Overzicht Netbeheerders stroom en gas in Nederland*. Accessed: 11.03.2020. URL: https://www.energieleveranciers.nl/netbeheerders/overzicht-netbeheerders.

[62] gk. *Browser Fingerprinting: An Introduction and the Challenges Ahead*. Sept. 2019. URL: https://blog.torproject.org/browser-fingerprinting-introduction-and-challenges-ahead.