Risky Business

Analysing the security behaviour of cybercriminals active on a darknet market

J.W. van de Laarschot

Technische Universiteit Delft

TU Delft

Risky Business

Analysing the security behaviour of cybercriminals active on a darknet market

by

J.W. van de Laarschot

Master thesis submitted to Delft University of Technology

in partial fulfilment of the requirements for the degree of

Master of Science

in **Complex Systems Engineering & Management**

to be defended publicly on Monday September 14, 2020 at 3:00 PM.

Student number:	4227530	
Thesis committee:	Prof. dr. M.J.G. van Eeten,	TU Delft, chair
	Msc. R. S. van Wegberg,	TU Delft, first supervisor
	Dr. A.M.G. Zuiderwijk-van Eijk,	TU Delft, second supervisor
	Dr. G.J. van Hardeveld,	FIOD, external advisor

TU Delft

Abstract

Cybercrime thrives and online anonymous markets, or darknet markets, play an important role in the cybercriminal ecosystem. Vendors active on darknet markets invest in security mechanisms to compromise the availability or usefulness of evidence to Law Enforcement Agencies. Therefore, difficulties arise in linking identities or machines to cybercrimes facilitated by darknet markets. As a result, many cybercrime investigations are ineffective.

This thesis consists of an exploratory case study based on the full administration of the Hansa Market. The Hansa Market (2015-2017) was infiltrated and eventually taken over and shut down by the Dutch Police. The data used in this thesis originates from the server that hosted the market and is made available by the Dutch National High Tech Crime Unit (NHTCU) and the Fiscal Information and Investigation Service (FIOD). This data is used to answer the research question: "*Which factors influence the security behaviour of darknet market vendors active on Hansa Market?*".

To answer this question, vendors that are similar regarding a) their experience, b) the activity on other markets, c) the amount of physical items sold, e.g. drugs and d) the amount of digital items sold, e.g. stolen credit card information, are clustered into five 'vendor types' using Latent Profile Analysis. It is researched whether these clusters of vendors differ in terms of the following security behaviours observed: a) authentication related security practices (password strength, password uniqueness and two-factor authentication usage), b) encryption of communication, in the form of PGP-adoption and PGP-key strengths used, c) the linkability of a vendors' pseudonym through PGP-key matching and d) a vendors' choice of Online Financial Service Providers within the bitcoin ecosystem, measured by querying a service that provides contextual information on bitcoin transactions and addresses.

The findings indicate that approximately causal relationships may be inferred between on the one hand vendor types, that represent a combination of business success in terms of physical and digital sales, experience and activity on other markets and on the other hand security behaviour. Vendors offering digital items tend to behave less securely than vendors selling large amounts of drugs. This thesis explains the observed (differences in) suboptimal security behaviours by arguing that vendors on Hansa Market conduct *subjective risk assessments*. This implies that the probability of being targeted by LEA and the value of the vendors' assets that are at stake (e.g. informational assets containing incriminating evidence or 'years of freedom') are of influence on security behaviour.

Lastly, recommendations to Law Enforcement Agencies include to exploit the subjectiveness in cybercriminals' risk assessments and to consider focusing on vendors transacting digital items. Academics are recommended to extend this research by investigating cybercriminal security behaviours through improved measurement methodologies in larger and more recent datasets.

Preface

Dear reader,

With this thesis, I conclude the Master Complex Systems Engineering & Management at the Delft University of Technology. This work on the security practices of cybercriminals active on a darknet market has been carried out at the Fiscal Information and Investigation Service (FIOD). I would like to thank all colleagues from the Financial Advanced Cybercrime Team for their friendliness, helpfulness and opportunities to learn. The technical assistance and insights greatly contributed to my thesis.

I also would like to express my gratitude to the members of my academic committee. Starting with Prof. Dr. Michel van Eeten, who as chair of the committee made sure to steer my thesis in the right direction during the kick-off and advised me on research approaches most fitting to this kind of research. Secondly, I would like to thank Dr. Anneke van Zuiderwijk-van Eijk for her very specific and structured feedback. This helped me in improving my research step-by-step. Thirdly, I would like to thank Rolf van Wegberg. Thank you for the excellent supervision of my thesis process. I greatly appreciate your willingness to frequently discuss my thesis project, your enthusiasm and positivity. Lastly, I'm thankful to dr. Gert Jan van Hardeveld. In the role of external advisor, you provided me with great insights on how the world of cybercrime works in practice. This helped me throughout the whole thesis process.

Lastly, I would like to thank my family, friends and (the primary victim of me writing a thesis) Tessa, for their support and patience during this process.

J.W. van de Laarschot
Rotterdam, September 2020

Contents

List of Figures

List of Tables

1

Introduction

"The dark nets are getting scary these days - f*** I hope this isn't the beginning of the end !?"

A pseudonymous cybercriminal on Reddit

A response to the seizure of AlphaBay, as documented by Bradley (2019, p.176)

Cybercrime is one of the biggest challenges of law enforcement agencies (Zhang, Xiao, Ghaboosi, Zhang, & Deng, 2012). Regarding the prosecution of cybercrime in The Netherlands, Van de Sandt (2019) even speaks of an *effectiveness crisis* in police investigations. An inadequacy to prosecute, is said to erode the willingness of citizens to report cybercrimes (Huisman, Princen, Klerks, & Kop, 2016, p. 58).

Online anonymous marketplaces are prominently placed in today's cybercrime ecosystem (Hartel & Van Wegberg, 2019, p. 67). The first successful online anonymous marketplace (or: *darknet market*) was Silk Road, which opened its doors in early 2011 (Soska & Christin, 2015). Predominantly hard- soft- and prescription drugs were transacted on this marketplace (Christin, 2013). The market enabled pseudonymous trading through a platform only accessible via TOR[1] on which solely cryptocurrencies were accepted as payment. By the end of 2013, Silk Road was shut down by Law Enforcement Agencies (LEA). In the short period of time that Silk Road was active, other initiatives copied the business model (Soska & Christin, 2015). Half a decade later, the yearly estimated revenue of all darknet markets combined is more than $790 million worth of cryptocurrencies (Chainalysis, 2020).

Roughly two types of products are sold on darknet markets: products that have to be shipped physically and items that can be transacted digitally. The physical goods are mostly drugs, e.g. cocaine, cannabis, heroin or other psychoactive substances (Dolliver, Ericson, & Love, 2018). Drug trade via darknet markets is a global phenomenon. When demand is high, supply facilitated by darknet markets follows (Dittus, Wright, & Graham, 2018). A broad selection of digital items is offered on darknet markets. These include: credit card details, hacked PayPal accounts, gift cards, login credentials for adult websites and streaming services such as Netflix and Spotify, fake identities, pirated software, botnet related items, databases of e-mail addresses, exploits and malware (Van Wegberg, Tajalizadehkhoob, et al., 2018). Lastly, 'guides to free money' -

[1] **T**he **O**nion **R**outer, a software package used for pseudonymous communication over the Onion Network, through which encrypted messages are sent over various network nodes. Onion Routing provides the sender a high level of security and anonymity (Goldschlag, Reed, & Syverson, 1999).

which are money mule[2] recruitment ads - are numerous (Van Wegberg, Tajalizadehkhoob, et al., 2018).

After taking down Silk Road, LEA have successfully shut down other darknet markets through large scale, internationally coordinated operations. According to Bradley (2019, pp. 228-230), these interventions impacted the capability of cybercriminals to trade in the darknet market ecosystem. In contrast, recent analyses of cryptocurrency transactions show that, despite law enforcement scrutiny, the transaction volume and the number of active markets in the ecosystem have an upward trend since 2014 (Chainalysis, 2020). Besides, the number of convictions remains relatively low (Bradley, 2019).

The security mechanisms of cybercriminals hinder investigators in attributing cybercrime (Eurojust and Europol, 2019). Differently put, these mechanisms thwart LEA in linking a cybercriminal or its machine to an identity or location (Wheeler & Larsen, 2003). In this thesis, it is posited that the 'security' of darknet market users is safeguarded through compromising the availability or usefulness of evidence to Law Enforcement Agencies (Harris, 2006). These safeguards can be found on a platform level, such as reputation mechanisms, verified accounts (Aldridge & Décary-Hétu, 2014; Hardy & Norgaard, 2016), rules, policies, content moderation (Wehinger, 2011) and user banning (Holt, Smirnova, Chua, & Copes, 2015). Such mechanisms constitute a form of 'extra-legal governance' (Dixit, 2011) and contribute to a more secure trading environment (Lusthaus, 2012). The security mechanisms on platform level provide users a level of *pseudonymity* - not full anonymity. History shows that LEA successfully infiltrated, took over and shut down darknet markets and that these events resulted in a number of arrests (Bradley & Stringhini, 2019). From this follows that users of darknet marketplaces do have to invest in security themselves as well. Indeed, Bradley (2019) shows that actionable knowledge on security mechanisms for cybercriminals is repeatedly being shared on discussion fora. Stronger still: users that receive letters or visits from LEA are being mocked by the community for having bad security practices (Bradley, 2019, pp.163-169).

Cybercriminals do not always achieve maximum security. Parallelly to the legitimate online world, security in the illegitimate world comes at a cost (*cf.* Bauer & van Eeten, 2009). Darknet market users need to invest in maintaining their knowledge, skills, equipment and secure work routines. This conception is in line with the reasoning of Van de Sandt (2019), who argues that cybercriminals perform security risk assessments and have certain risk appetites[3]. Additionally, cybercriminals can be subjected to certain behavioural biases that impede their security (Van Hardeveld, 2018). Van Hardeveld shows evidence that cybercriminals do not keep up to date with the latest security practices, due to a behavioural pitfall referred to as a 'status-quo bias' (Samuelson & Zeckhauser, 1988). This entails that people might prefer to keep things (e.g. security practices) the way they are.

Hence, explanations exist why cybercriminals might exhibit less secure behaviour, despite the risky environment they operate in. This thesis aims to provide insights into the security behaviour of darknet market users. To this end, data of the full administration of the Hansa Market is used. This market was active from June 2015 to July 2017. Hansa was infiltrated and eventually taken over and shut down by the Dutch Police. The Hansa Market data used in this thesis originates from the server that hosted the market. It has been made available to the Fiscal Information and Investigation Service (FIOD) by the Dutch National High Tech Crime Unit (NHTCU).

[2]Money mules are intermediaries that (often unknowingly) aid criminals in laundering money by transferring funds between accounts.
[3]The level of risk that an individual accepts, balancing costs and security threats.

This thesis is the first academic effort that extensively explores the security behaviour of darknet market users using quantitative data. In the work presented here, the security behaviour of Hansa Market vendors is related with their characteristics, e.g. how long they are active on the market or the amount and type of sales they made. Using these characteristics, similar vendors are grouped into five criminal types. It is researched how these criminal types perform relatively to each other in terms of security behaviour. From the differences in security behaviour between the groups of darknet market users, approximately causal relationships are inferred that explain the security behaviour of cybercriminals.

1.1. Knowledge Gaps

Van de Sandt (2019, p.231) demonstrates the necessity for a new academic field of study demystifying the security practices of cybercriminals. This field of study awaits major contributions from the socio-technical disciplines known for combining social sciences with computer science research. The current academic works that arguably make a first contribution to shaping this field of study are discussed in this section. From these, two knowledge gaps are inferred. A more comprehensive review of literature relevant to this research is found in chapter 3.

- *A limited conceptual insight in the security behaviour of darknet market users*

  Research on cybercriminals' security practices is scarce: only the works of Van Hardeveld (2018), Van de Sandt (2019) and Van Wegberg and Verburgh (2018) are aimed at researching security behaviour of cybercriminals specifically. Van Hardeveld (2018) elaborates on the decision-making of carders[4]. The author examines technical security mechanisms found in online carding tutorials and discusses cognitive biases that lead to suboptimal security. Expert interviews provided evidence that some of these biases apply to carders. Van de Sandt (2019) lays a mostly theoretical foundation of how cybercriminals deploy technical computer security controls that aim to protect the criminal and the crimes he or she commits (Van de Sandt, 2019, p.7). While Van de Sandt has a strong focus on conceptualising security practices of cybercriminals and his findings are predominantly of qualitative form, his research is not aimed at darknet market users specifically nor does it provide explicit definitions of security behaviours that might be observed on darknet markets. He does acknowledge that approaching the research on cybercriminal security behaviour in a quantitative manner will produce more granular insights (Van de Sandt, 2019, p.232). Lastly, the research of Van Wegberg and Verburgh (2018) revolves around a single and specific security behaviour. The authors show that vendors attempt to reduce the linkability of their pseudonyms when migrating from one market to another. The security mechanism analysed is whether vendors stick with their PGP-key and/or username when switching markets.

  Apart from research on the security behaviour of individuals, a significant amount of work focuses at understanding the security mechanisms of entities that facilitate the illegal transactions. These are efforts on a *platform-level*. For example, the self-regulation on darknet markets (Wehinger, 2011) through its reputation mechanisms (Aldridge & Décary-Hétu, 2014; Hardy & Norgaard, 2016), the interaction between the popularity of anonymity enhancing cryptocurrencies and darknet markets (Foley, Karlsen, & Putniņš, 2019; Janze, 2017) or the forensic challenges and opportunities that the popularity of cryptocurrencies results in (Tziakouris, 2018). While important to the field of research regarding security practices of cybercriminals, these studies fail to generate insights on what additional security measures individuals on darknet markets take to safeguard their security.

[4]Carders trade stolen credit card and bank account details.

- *It is unknown to what extend suboptimal security behaviour can be observed on darknet markets, among what types of darknet market users suboptimal security behaviour is most prevalent and what might cause this suboptimal security behaviour.*

There are numerous indications throughout literature that (cyber)criminals not always achieve maximum security. First, any criminal is economically incentivised, resulting in a trade-off between enhanced security and improved efficiency of operations (Morselli, Giguère, & Petit, 2007). Second, Holt et al. (2015) observe that users actively trade-off between risks and rewards of a transaction on forums facilitating the trade of stolen data. Third, Van de Sandt (2019) argues that 'perfect security' is not economically viable for cybercriminals. Fourth, in a study on online underground forums, Sundaresan, McCoy, Afroz, and Paxson (2016) show that vendors do not consistently use VPN services to hide their likely geolocation and that they are prone to use less secure communication methods. Fifth, Bradley (2019, p. 195) observes that users on a darknet marketplaces use the auto-encryption features of the market, even though this feature poses a security risk.

Obviously, the arrests of darknet market users by LEA is also indisputable evidence that some cybercriminals exhibit suboptimal security. Law Enforcement Agencies even have been successful in arresting key players and operators of (amongst others) the Silk Road, AlphaBay, Hansa, Valhalla and Wallstreet darknet markets (Broadhurst, Ball, & Trivedi, 2020). The arrests of these key players are significant blows to the darknet market ecosystem (Bradley, 2019) and show that even the most notorious cybercriminals make mistakes in their security. Nonetheless, the sparse data on arrests of darknet market users do not provide insight in the security behaviour of the darknet market population at large.

Three academic works that analyse the security behaviour of larger populations of darknet market users exist. Next to the before-mentioned Van Wegberg and Verburgh (2018), who analysed the evasion measures of vendors upon switching markets, Soska and Christin (2015) measure how the use of encrypted communication methods increases over time on darknet marketplaces. Décary-Hétu, Paquet-Clouston, and Aldridge (2016) make an effort in measuring security risk-taking behaviour of vendors. The authors operationalise 'security risk' in a very limited way by only taking into account the willingness of vendors to ship internationally. Moreover, their findings are based on analysis of outdated and incomplete data. These three works tend to focus at only one aspect of security behaviour and do not consider characteristics of the vendors when drafting conclusions about their security behaviour.

1.2. Scope of this thesis

The delineation of this thesis is further clarified through a problem statement, research objective and the research questions. Additionally, the link to the study programme of Complex Systems Engineering and Management is elaborated upon.

1.2.1. Problem Statement

A significant amount of cybercrime is facilitated by darknet marketplaces. The complexity of attributing these crimes is fuelled by darknet market users deploying security mechanisms in addition to the security mechanisms provided the platforms.

Research on the security practices of cybercriminals is still in its infancy (Van de Sandt, 2019). Currently, there is a limited conceptual understanding of the security practices of darknet market users, it remains unknown how prevalent less secure behaviour on darknet markets is and insights based on empirical and quantitative analysis regarding the causes of such behaviour are

scarce. Furthermore, because the security practices of darknet market users are not structurally measured, it is unknown to what extend this behaviour evolves over time.

The complexity of cybercrime attribution and the lack of insights in security behaviour of cybercriminals, make cybercrime investigations cost intensive and time consuming. LEA try to maximise their impact by focussing on key players in the darknet market ecosystem. Although this resulted in numerous successful interventions, the volume of illegally transacted items has been on the rise and the amount of convictions remains relatively low.

1.2.2. Research Objective

Taking into account the academic knowledge gap and the problem statement, the research objective is formulated as:

> *To develop a deeper understanding of security behaviour of cybercriminals by measuring and comparing the security behaviour of different types of users that are active on Hansa Market. This deeper understanding should provide LEA insights into opportunities for more effective cybercrime investigations.*

1.2.3. Research Questions

In order to address the academic knowledge gaps and to achieve the objective of the research, this thesis answers the main research question:

> *Which factors influence the security behaviour of darknet market vendors active on Hansa Market?*

To approach this research question in a structured manner, the following sub-questions are answered.

- **SQ1**: *What security behaviours can potentially be observed on darknet marketplaces?*

  This sub-question is answered through literature review in section 3.1. Here, a theoretical framework of security behaviours that compromise the availability or usefulness of evidence to the investigative process is presented.

  In chapter 2 it is explained that approximately causal relationships between individuals and their security behaviour are inferred through analysing and comparing the security behaviour of different groups of darknet market users. To this end, this thesis distinguishes a) buyers from vendors and b) within the group of vendors, subgroups of *vendor types*. These are created by grouping similar vendors together. This yields sub-question 2:

- **SQ2**: *What characteristics of vendors are relevant to include when distinguishing between different types of vendors?*

  Sub-question 2 is answered through literature review in section 3.2. Here, characteristics of vendors are selected that are hypothesised to be related with security behaviour. Based on these characteristics, similar vendors are grouped together.

- **SQ3**: *What types of vendors can be distinguished on Hansa Market?*

  This sub-question is answered through a Latent Profile Analysis in section 5.1. The selected vendor characteristics are used to group similar vendors. Each group of similar vendors is assigned a vendor type.

- **SQ4**: *How do the vendor types compare relatively to each other in terms of the security behaviours analysed?*

This sub-question is answered throughout chapter 6. Here, it is shown to what extend vendor types statistically significantly differ from each other in terms of each of the security behaviours identified.

- **SQ5**: *How do the vendor types compare relatively to each other when all security behaviours are considered jointly?*

  This sub-question is answered in section 6.6, in which a scoring method is applied to visualise how the overall security behaviour of vendors compares.

- **SQ6**: *How can these differences in security behaviour be explained?*

  Within chapter 6, the analysis of each security behaviour is completed with a conclusion and discussion on how the result impacts LEA. Here, it is attempted to generalise the findings and to discuss underlying causes. These findings are further discussed and summarised in chapter 7.

1.2.4. Societal relevance of this study

Underground market places and hiding technologies of cybercriminals are two out of the seven 'grand challenges' of cybercrime (Koops, 2016). Indeed, Van de Steur (2016) acknowledges that cybercrime is thriving and that LEA are not able to intervene effectively. This undermines democracy (Van de Steur, 2016). Fighting cybercrime is high on the political agenda (Grapperhaus, 2018). The current Dutch Minister of Security and Justice introduced research programs to develop deepened insights into cybercriminal behaviour. Such academic research contributes to effective policy-making (Grapperhaus, 2018).

Dutch LEA advocate evidence based policing (Huisman et al., 2016), in which gathering, sharing and analysing information is leveraged to enable rational decision-making on effectively allocating LEAs resources (Bokhorst, Steeg, & de Poot, 2011). The use of *residual information* is an essential part of evidence based policing. Residual information entails any information acquired that needs further processing or cannot be directly used in ongoing investigations (Van Wijk & Scholten, 2006, p.9). The research presented here involves the use of gathered and shared residual information and facilitates decision-making towards effective cybercrime investigations through thorough analysis.

Lastly, both academic research into cybercrime and evidence based policing prevent an 'incident driven' approach. Such an approach may lead to disproportional countermeasures that are at odds with human rights. A challenge of fighting cybercrime is rebalancing the enforcement of the law and the protection of the right to privacy (Koops, 2016). Due to the ineffective prosecution of cybercriminals, politicians started to debate the legality of encryption and argue for backdoors in encryption protocols (Barr, 2020; Grapperhaus, 2020). This thesis evaluates which tools of anonymity on a darknet market are used most and where, within current mandates and existing regulatory frameworks, opportunities for law enforcement are to be found.

1.2.5. Link to the CoSEM Study Programme

This research is conducted as part of the study programme Complex Systems Engineering and Management at the Delft University of Technology. Three links with this programme are important to note. First, this research embraces the interrelatedness of the behavioural and technical aspects of cybercriminal security behaviour. This means that this thesis does not aim to discover new technical vulnerabilities in the security enhancing mechanisms deployed by cybercriminals. Rather, it focusses on the factors that influence how such tools are used. The latter is only possible when the technical qualities of security enhancing mechanisms *and* human behaviour are studied jointly. Secondly, institutional considerations are important to this thesis. Insights into regulatory frameworks, the mandate of law enforcement and international institutions that combat cybercrime are needed to understand cybercriminal security behaviour. Some security mechanisms rely more on exploiting institutional loopholes than technical qualities to increase security. Third, LEA may develop new ways to intervene in the darknet market ecosystem using the insights generated in this thesis.

1.3. Structure of this thesis

The remainder of this thesis is as follows. First, the case study research approach and how approximately causal relationships are inferred are discussed in chapter 2. Chapter 3 consists of a review of academic literature and results in two products: a theoretical framework of behaviours influencing 'security' and an overview of characteristics of darknet market vendors that are relevant to consider. Chapter 4 elaborates upon the methodology used to cluster similar vendors into vendor types and explains how the identified security behaviours are measured. Chapter 5 presents the results of analysis of vendor characteristics and the clustered vendor types. Chapter 6 entails descriptive and statistical analyses of the security behaviours of the Hansa Market users. Each analysis of the security behaviours is completed with a discussion, in which the results are related to more general cybercriminal behaviour. All results and practical insights of the separate analyses are summarised in the last section of this chapter (section 6.6). This thesis concludes by answering the research questions (section 8.1), elaborating upon future work & recommendations (section 8.3) and reflecting on ethical considerations (section 8.4).

2

Research Approach

> "Remaining markets, vendors and buyers will tighten their opsec. Unless they have a tor exploit it will get much more difficult for LE to take down dnm's"
>
> *A pseudonymous cybercriminal on Reddit*
>
> Hypothesising about security after the intervention on AlphaBay and Hansa Market, as documented by Bradley (2019, p.187)

For the simple reason that a particular darknet market is observed, this thesis entails a *case study*. The Hansa Market case is selected rather pragmatically, since detailed data of this market has been made available to this research. As a consequence, a sample bias is present (Collier & Mahoney, 1996). As Seawright and Gerring (2008) rightfully note, even when cases are selected for pragmatic reasons, it is essential to reflect on how the properties of the case itself (i.e. Hansa Market) relate to the rest of the population (any other darknet market). This is reflected upon in section 8.3.

Within the Hansa Market case, the security behaviours of different types of vendors are studied. Thus, using the terminology of Yin (1993), an embedded single case study is performed in which: within the *context* of cybercriminals trading illicit goods on the dark web, the *case* of security behaviour on Hansa market is studied through comparing different embedded *units of analysis* (Figure 2.1). On defining units of analysis, Ragin, Becker, et al. (1992) discuss four approaches to 'cutting whole units out of reality'[1]. The first approach is of an inductive and ethnographic kind. Units of analysis are formulated without the use of any theory and based on empirical observations only (Ragin, Becker, et al., 1992, pp. 139-157). Such an approach is not feasible since the population is large and little contextual information on all individuals belonging to this population is available for ethnographic descriptions. The second approach is to make use of formally defined objects, e.g. 'nation states', 'families' or 'organisations' (Ragin, Becker, et al., 1992, pp. 173-180). Clearly, this approach is applied when differentiating between *buyers* and *vendors* (see section 4.1). However, formal definitions that aid in distinguishing different types of vendors within the vendor population are not available in a darknet market context. Thirdly, researchers may resort to using generally accepted conventions (Ragin, Becker, et al., 1992, p.10). Again, such conventions do not exist in darknet market research. Lastly, the authors suggest that 'whole units' can be formed by *making* them. Based on common characteristics, the units

[1] While the authors use the terminology *case* instead of *unit of analysis*, the discussion is very relevant to this research. The authors debate how to meaningfully cut 'a whole unit' from reality, when this unit has ambiguously defined temporal or spacial boundaries (Ragin, Becker, et al., 1992, p.166).

of analysis are gradually imposed on the empirical evidence (Ragin, Becker, et al., 1992, p.10). The last approach is selected to distinguish vendor types within the vendor population. Thus, the units of analysis are defined through deriving vendor types empirically from the data based on theoretically justifiable common characteristics.

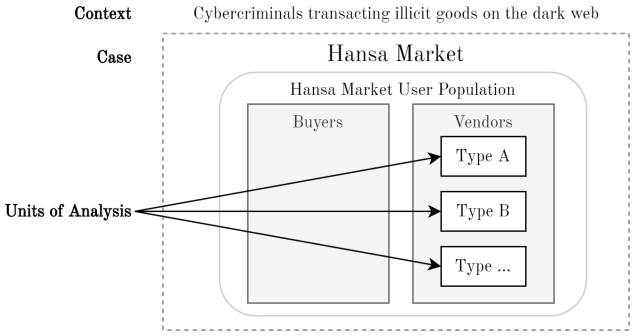


Figure 2.1: Case study design

In this thesis, most differences in security behaviour are observed on group level (vendor types), but inferences are made that apply to the individual. Being, that when a certain vendor type on Hansa displays relatively poor security behaviour, it is inferred that a Hansa Market vendor belonging to this group is likely to have poor security practices (compared to other Hansa Market vendors not belonging to this group). Such reasoning is valid when security behaviour exhibited within vendor types is somewhat homogeneous. In that case, it is safe to assume that micro-level behaviours are linked with macro-level observations (Alexander, 1987) and *downward cross-level inferences* could be made that infer individual level relationships from a higher level of analysis (Mossholder & Bedeian, 1983).

When homogeneity within groups is absent, aggregation (especially using group means) introduces biasses (Hammond, 1973). The within-group variation is not captured by this high level measure as a consequence (James, 1982). Because this thesis is of exploratory nature, it is more of a 'plausibility probe' or a 'theory building exercise' rather than a confirmation of well researched theories (Eckstein, 2000). Therefore, the concerns of James and Hammond are justified: it is certainly plausible that the variance in security behaviour is high within the defined vendor types. Mossholder and Bedeian (1983) suggests to reduce this risk by controlling the relationship between the aggregated measure and the dependent variable of interest for the effect of individual-level measurements using common analytic approaches such as regression analysis and analysis of covariance (ANCOVA).

However, Ragin, Becker, et al. (1992, pp. 105-118) contest the relevance of this family of statistical procedures for exploratory case studies. It would require selecting arguable factors to control for in the research design. Additionally, it forces the researcher to discuss variables having a (somewhat) measurable effect on the variance of another variable, instead of 'discussing the things individuals actually do' (Ragin, Becker, et al., 1992, p.206). The combination of descriptive analyses and logical inferences (as opposed to probabilistic statistical techniques) is put centrally by Ragin (1987). It is not attempted to follow Ragins approach completely. Still, similar to Ragins work, approximate causal inferences are made through counting how many times certain behaviour is (not) observed. Additionally, to address the concerns of aggregation biasses a) robustness is introduced into the analysis by considering multiple security behaviours. When an inferred relationship holds for multiple security behaviours analysed, this may be regarded as proximal evidence of causality (Gerring, 2006) and b) it is chosen to visualise the variance in the security behaviours within vendor type whenever possible.

3

Literature Review

"Jupiter has made what Whannahave and Bob Marley asked him to, but apparently it doesn't work properly"

A pseudonymous cybercriminal

Excerpt from an intercepted chat between cybercriminals, as documented by Odinot, Verhoeven, Pool, and De Poot (2017)

This chapter presents relevant previous academic work in two sections. The first section is tailored to discussing the security behaviour of cybercriminals. In this section, security behaviours that can be observed on darknet markets are identified and related to more general classes of cybercriminal security behaviour found in literature. The second section elaborates on characteristics that differentiate darknet market vendors. This section concludes with a short discussion how these characteristics are expected to relate with security behaviour.

The search strategy for this literature review can be described as follows. A snowballing approach is used, because few works are aimed specifically at researching security practices of cybercriminals. This process started with the dissertations of Van de Sandt (2019) and Van Hardeveld (2018). Within Van de Sandt (2019), works belonging to the field of digital anti-forensics were cited. These provide definitions of what individuals do to increase their security. Existing taxonomies of these security practices were found that group certain types of behaviours together. Parallelly to this process, a snowballing approach was used to explore (empirical) darknet market literature. Within darknet market research, some academics focus on analysing large datasets created through 'scraping' darknet markets. Two influential works (Soska and Christin (2015) and Christin (2013)) were used to find similar academic efforts that make behavioural inferences based on quantitative data. The academic knowledge gaps resulting from this literature search have been synthesised in section 1.1. Furthermore, international security standards, information security behaviour related literature and cryptography related research is consulted to establish what practices are considered to be secure. In the section below, all security practices found through this literature search are discussed, related to each other and presented in a theoretical framework.

3.1. Observable security behaviour on darknet markets

Research on the cybersecurity behaviour of cybercriminals is still in its infancy (Van de Sandt, 2019). On the security behaviour of darknet market users specifically, research is nearly absent. Because of these limitations, this literature review draws from various research areas with the goal to identify and define concepts that describe cybercriminal security behaviour. In this review, existing taxonomies of cybercriminal security behaviour are connected to security behaviour that is observable on darknet markets. The benefits of doing so are twofold. Firstly, the existing taxonomies are extended into further detail to accommodate security behaviour specific to darknet market users. And secondly, the security behaviour of darknet market users can be interpreted in the context of well defined taxonomies.

3.1.1. Defining 'security behaviour of cybercriminals'

In the cybercriminal community, security practices are named 'OPSEC' (Bradley, 2019), which stands for *operational security*. The term that originates from the US Military. In the context of cybercriminals, it is used to describe the criminals' efforts that enable them to conduct their illegal business safely (Van Hardeveld, 2018, p.12) or *"the process of securing transactions on the darknet"* (Bancroft & Scott Reid, 2017). This definition does not fully satisfy the properties of a complete definition of 'security' (see Baldwin, 1997) since it is not indicated 'from what' security is needed.

Another definition, which does satisfy Baldwins requirements, is found in Van de Sandt (2019). Here, the author introduces the term 'deviant security' to describe *"all technical computer security controls of natural and legal persons who are criminally liable for the commission of crime, in order to protect the criminal and his/her crimes"* (Van de Sandt, 2019, p.7). This definition is rather broad and not specific enough for the purpose of this thesis, i.e. measuring the security behaviour of darknet market users. Additionally, darknet market users rely on numerous financial controls (Van Hardeveld, Webber, & O'Hara, 2017), of which it could be debated whether they can be regarded as technical computer security controls.

A significant share of research on cybercriminal security behaviour stems from the field of digital anti-forensics research. Harris (2006) defines digital anti-forensics as *"any attempts to compromise the availability or usefulness of evidence to the forensic process"*. Sremack and Antonov (2007) add that although digital anti-forensics is technical in nature, these acts may be expressed physically as well. The field of digital anti-forensics proved to be a valuable source for concepts that describe cybercriminal security behaviour. This field is relatively well researched and it distinguishes between types of clearly defined security behaviours. From this point on, cybercriminal security behaviour will be regarded from the perspective of Harris (2006). Because not all evidence is gathered through forensic methods, 'security' of cybercriminals is generated through: any attempts to compromise the availability or usefulness of evidence to the investigative process.

Security behaviour can be divided into subclasses of data hiding, trail obfuscation, data destruction and data minimisation. It was found that some behaviours in the financial domain transcend these subclasses. Therefore the additional subclass 'choice of Online Financial Service Providers' (OFSPs) is added. All security behaviour subclasses are discussed in the context of darknet markets below.

3.1.2. Data Hiding

Data hiding is *"the act of removing evidence from the view so that it is less likely to be incorporated into the forensic process"* (Harris, 2006). Similar definitions are given by Conlan, Baggili, and Breitinger (2016), Rogers (2006). It is stressed by Peron and Legary (2005) that compared to other forms of digital anti-forensics, data hiding techniques allow the data to only be accessed and used by those who hid the data. In general, literature from the field of anti-forensics distinguishes between data hiding through either applying steganographic or cryptographic techniques (Garfinkel, 2007). Steganography - meaning 'covered writing' - refers to concealing a message in another message. It aims to hide that certain data is exchanged at all (Cabaj et al., 2018), by for example hiding text in image files (Hetzl & Mutzel, 2005). Steganography differs from encryption, given that encrypted information is easily recognisable through its extremely high entropy and specific tags, headers or signatures (Garfinkel, 2007). This causes the exchange of encrypted data to be easily observed. The content of the exchanged data cannot be accessed without a secret key (Cabaj et al., 2018). No indicators were found that steganographic approaches are used by criminals trading on darknet markets. Cryptographic security measures however, are widely deployed and thus are further explored in this thesis.

Next to cryptographic approaches, also other security measures that fit the definition of data hiding are deployed by darknet market users. Security behaviour that refers to authentication, crossing jurisdictions and the unlinkability of pseudonyms can be observed as well.

3.1.2.1. Encryption of communication

The use of encrypted communication is security behaviour that can be observed easily. It is known that among cybercriminals active on darknet markets, PGP is the most used encryption protocol for secure communication (Cox, 2016). A distinctive public PGP-key is publicly listed by the vendor, such that a buyer can encrypt information that only can be decrypted by the vendor: the vendor is the only one in possession of the secret private key. In a longitudinal study, Soska and Christin (2015) researched the adoption of PGP among vendors on darknet markets. The authors estimated that the PGP-adoption was about 25% in 2012 and in 2015 it reached levels of 90% on some markets. The authors merely observe the fact whether a PGP-key has been mentioned by the vendor. Thus, the authors opted for a rather inclusive view of PGP-deployment: *mentioning* a key does not imply consistent encryption of messages.

Some markets offer an 'auto-encryption' functionality. This functionality allows darknet market users to PGP-encrypt their communication with the click of a button, without the need to go through the cumbersome PGP-installation procedure. The PGP setup procedure is infamously known to be difficult to understand for the layman (Ruoti, Andersen, Zappala, & Seamons, 2015; Whitten & Tygar, 1999). To make the procedure easier, tutorials for PGP-encryption are widely available in the cybercriminal community (Van Hardeveld et al., 2017). Compared to the full PGP-installation, the auto-encryption functionality is easy to use but less secure. During past undercover operations, Law Enforcement was able to switch off the auto-encryption and consequently, it could read along communication that cybercriminals presumed to be encrypted (Van Hardeveld, 2018, p.132). The security flaw of auto-encryption became known in the cybercriminal community, as analysis of Bradley (2019, p.194) shows through an excerpt of a Reddit discussion forum. Here, users recommend to personally encrypt communication instead of using auto-encryption functionalities because of security risks.

Next to or instead of PGP encrypted communications, users on darknet markets also list other ways through which they can be contacted. Researchers have found that a.o. Jabber, Skype, PrivNote, ICQ, Wickr and Exploit.im are used as ways to communicate (Aldridge & Askew, 2017;

Sundaresan et al., 2016; Van Wegberg et al., 2020). It is remarkable that Skype is included in this list, since it has a security flaw through which the likely geo-location of its users can be obtained (Sundaresan et al., 2016). These alternative communication platforms are referred to as 'other communication services'.

3.1.2.2. Authentication

Authentication is the process of confirming the identity of a user. On a darknet market, a user is proving that he is who he says he is by entering a secret password that matches his darknet market username. After a successful login, i.e. when the user is authenticated, data that the user is authorised to see is shown ('unhidden'). In the case of a vendor active on a darknet market, these include communications with other users, transaction overviews and his or her own listings. Despite the fact that password authentication has been around for decades, its use still comes with a significant amount of bad practice (Furnell, 2011). Improving and understanding password hygiene is well researched in the field of information security behaviour (Stanton, Stam, Mastrangelo, & Jolton, 2005). Results show that even forcing, or nudging (Furnell, Esmael, Yang, Li, et al., 2018; Kankane, DiRusso, & Buckley, 2018), users to adhere to strict password requirements does not withhold users from picking predictable and easy-to-hack passwords (Komanduri et al., 2011; Shay et al., 2015). When plaintext passwords are be obtained, password strength can be determined.

Research suggests that in the legitimate world, password reuse is very common (Golla et al., 2018). Even when people are aware that strong passwords are important, these strong passwords are often reused over different websites (Wash, Rader, Berman, & Wellmer, 2016). The authors stipulate that when people choose a complex password for a website they frequently log in to, this password is easily memorised because of its frequent use. After memorisation, it is tempting to use the complex password on other websites as well. Password reuse is very non-secure behaviour since data breaches in which passwords are obtained are plentiful. Good password cracking software makes use of such databases with leaked passwords. A theoretically complex password can be easily breached, when it is reused on a website which has poor security practices (Ives, Walsh, & Schneider, 2004). The National Institute of Standards and Technology (NIST) develops information security guidelines. Because password reuse bears security risks, NIST advises organisations to assess passwords uniqueness in their most recent recommendations. When a user registers an account, organisations should make sure the users' password is not identical to any password in leaked password databases (Barker, Barker, Burr, Polk, Smid, et al., 2020).

Next to logging in with only a password, some markets allow users to enable two-factor authentication (2FA) (p.146 Van Hardeveld, 2018). The 2FA-enabled login works via PGP, meaning that the user is required to have PGP set up properly (Zhou, Zhuge, Fan, Du, & Lu, 2020). The user is presented a text that is encrypted with their public key. The user is challenged to obtain and submit the original unencrypted text, which only can be done using their secret private key (Carr et al., 2019).

3.1.2.3. (Un)linkability of darknet market pseudonyms

Van de Sandt (2019, p.153) states that cybercriminal activities leave behind fragments of information. These fragments are decentrally stored in a variety of databases. To ensure that these fragments are not included in the investigative process (*cf.* 'data hiding'), it is in the cybercriminals' interest to keep these fragments dispersed. Consequently, unlinkability from a security perspective is an attribute of confidentiality (Pfitzmann & Hansen, 2010).

The unlinkability of darknet market pseudonyms refers to the inability of Law Enforcement to link two or more usernames to the same real world identity. Formally put: *"two or more items of interests are unlinkable if one cannot sufficiently distinguish whether these items are related or not"* (Pfitzmann & Hansen, 2010) and a pseudonym is *"an identifier of a subject other than one of the subject's real names"* (Pfitzmann & Hansen, 2010).

Acts of linking, 'matching' (Tai, Soska, & Christin, 2019), 'record linkage' (Christen, 2012) or 'sybil account detection' (Kumar et al., 2020) describe finding the pseudonyms that presumably refer to the same real-world entity. When multiple pseudonyms belonging to a single cybercriminal can be connected, a security risk for this criminal is created. LEA may accumulate advanced knowledge on a persons behaviour and identity, which may result in bringing this person to justice (Ho & Ng, 2016). Also Van de Sandt (2019) describes how Law Enforcement Agencies build profiles on real-world identities by connecting disconnected fragments of information. A famous example is how Law Enforcement de-anonymised the alias 'Dread Pirate Roberts', the founder of the original Silk Road, by linking different aliases across message boards (Popper, 2015). Linking vendor accounts is also done by academics to accurately estimate sales volume (Kumar et al., 2020) or to track migration patterns of vendors over different markets (Van Wegberg & Verburgh, 2018).

In this literature review a distinction is found between two types of security behaviour that relate to linkability. The first is that vendors may actively increase the linkability of their pseudonyms. These vendors allow that their pseudonyms can be linked through simple reasoning, or by connecting some dots that say, a mainstream buyer, is able connect. The second type of linkabilty is that vendors may 'passively allow' linking through advanced techniques, from which it is safe to assume that the average buyer would not apply them.

Actively increasing linkability Interestingly enough, vendors on darknet marketplaces cope with a perverse incentive with regard to hiding the links between various pseudonyms that are owned by them. Next to being a security risk, having a clear link between multiple user accounts owned by the same entity is believed to increase business success (Van Wegberg & Verburgh, 2018). Vendors active on (multiple) darknet markets may knowingly increase the linkability of their darknet market pseudonyms with the goal to increase their sales. By having this clear link between user accounts, valuable reputations can be transferred to other markets. Van Wegberg and Verburgh (2018) reason that in the pseudonymous world of darknet marketplaces, reputation distinguishes frauds from high quality vendors. This makes 'reputation' an important asset to the cybercriminal. Because darknet market users trade pseudonimously, usernames (instead of real names) come with a certain reputation (Décary-Hétu & Leppänen, 2016). They represent a brand (Lusthaus, 2012) and are signals of trust (Holt, Smirnova, & Hutchings, 2016). Because vendors have the incentive of reusing a username over different markets, a variety research is done on matching usernames across markets. Ranging from obtaining an exact match (Soska & Christin, 2015) to more elaborate techniques where similar but not identical usernames are matched (Décary-Hétu & Giommoni, 2017; Tai et al., 2019; Van Buskirk et al., 2017).

Not only the username signals trust and is tied to a reputation. This also goes for the public PGP-key listed by a vendor (Van Wegberg & Verburgh, 2018). PGP-keys are suitable for signalling trustworthiness, because their legitimacy can be verified by asking the signalling party to decrypt a text (Tai et al., 2019). In theory, PGP-keys thus have a high 'cost-to-fake'. This is explained using signalling theory (Gambetta, 2009; Holt et al., 2016). Signalling theory departs from the observation that criminals are unable to adhere to formal institutions to settle disputes. Consequently, the quality of goods that are transacted may be reduced. Criminals therefore

aim to communicate their trustworthiness through the use of signals, which are 'intentionally displayed and clearly observable features' (Gambetta, 2009). These signals however, also have a defect: they can be manipulated through imitation and forgery. Because of this, signalling theory takes the *costs to fake principle* into account. Signals that are costly to fake but 'cheap to emit' are most valuable. Unfortunately for cybercriminals, PGP-keys are not fully mimic proof in practice. An imposer still can use a PGP-key to signal trustworthiness and if the other party does not go through the trouble of verifying the PGP-key, the imposer is still successful (Tai et al., 2019).

So vendors have incentives to register the same PGP-key over different markets to signal trustworthiness. This feature has been successfully used in practice and in academic literature to link pseudonyms with (Broséus et al., 2016; Soska & Christin, 2015; Tai et al., 2019; Van Wegberg & Verburgh, 2018). Still, it must be taken into account that vendors may choose to use more than one key as a) it is an evasive strategy (Van Wegberg & Verburgh, 2018), b) keys can expire (Broséus et al., 2016)[1], or c) private keys can be lost (Tai et al., 2019).

The online darknet market search engine service *Grams* used username and PGP-key matching to offer linkability insights to a large audience (Branwen, 2020). This implies that when a vendor chooses a similar username or reuses a PGP-key over multiple markets, he actively increases the linkability of his or her pseudonyms. This is in contrast with the other type of linkability, which only can be achieved through more advanced linking techniques.

Passively allowing linkability When the available content produced under pseudonyms (e.g. posts, listings or feedbacks) is analysed, pseudonyms may be linked. On unlinkability of two anonymously sent messages Pfitzmann and Hansen (2010) state: *"Please note that unlinkability of two (or more) messages of course may depend on whether their content is protected against the attacker considered* [...] *with access to their content the attacker can notice certain characteristics which link them together - e.g. similarities in structure, style, use of some words or phrases, consistent appearance of some grammatical errors* [...] *the content of messages may leak some information on their linkability".* Afroz, Islam, Stolerman, Greenstadt, and McCoy (2014) refers to the analysis of content produced by criminals with 'authorship analysis'.

The following authorship analysis techniques to discover vendors that have multiple darknet market pseudonyms but choose not to reveal are found. Spitters, Klaver, Koot, and van Staalduinen (2015) analyse written text in listings to link vendors. Outcomes of such stylometric approaches can even be used in court as evidence (Van de Sandt, 2019). A limitation of this technique is that the substantial amount of l33t-speak, or hacker jargon, troubles the results (Afroz et al., 2014). To match different usernames, also photos and photography styles are analysed. Using the aid of machine learning techniques, Wang, Peng, Wang, and Wang (2018) show that this technique can outperform conventional methods of stylometry. The authors mention that a limitation of this approach is that (publicly available) photos are reused by different real world identities. Tai et al. (2019) add that cybercriminals can easily counter image linking techniques by altering or normalising photos. This would be a form of *anti* anti-forensics, as described by Van de Sandt (2019).

[1] Although an expired key can still be used!

3.1.2.4. Crossing jurisdictions

Security behaviour of cybercriminals is heavily shaped by the (absence) of laws and regulations (Van de Sandt, 2019, pp.76-81, p.198). Cybercriminals create information asymmetries between key players (offenders, victims, law enforcement agencies) by distributing evidence across multiple jurisdictions (Van de Sandt, 2019, p.120). In this section, such behaviour is described as 'crossing jurisdictions'.

Crossing jurisdictions refers to *"limiting what evidence can be captured due to inability to access data in one or more jurisdictions* (Sremack & Antonov, 2007). Van de Sandt (2019, pp.168-173) regards this type of security behaviour as a 'distribution countermeasure'. The author, referring to all types of cybercrimes, argues that three geographical locations give LEA jurisdiction: the location where the attack originates from, where the victims of the attack are located and the location of infrastructures that support the attack.

In a darknet market context, the country of residence of vendors can be observed since these preferences are often denoted in the listing of vendors. Vendors often specify to which countries they are willing to ship (Décary-Hétu et al., 2016). Van Hout and Bingham (2014) and Van Buskirk, Naicker, Roxburgh, Bruno, and Burns (2016) argue that this information given by the vendors can be assumed to be correct, because incorrect information leads to negative feedbacks. Consequently, most vendors that ship physical goods are also truthful when listing the country from which their packages ships from. It is however difficult to relate shipping preferences with security preferences. Non-security related factors that influence the willingness to sell or buy internationally vary among countries and are numerous: geographic isolation, domestic prices, domestic supply and demand and the proximity to producing countries (Décary-Hétu et al., 2016; Van Buskirk, Naicker, Roxburgh, et al., 2016). This is underlined by (Dittus et al., 2018), who conclude that the geographical routes of drug trade on darknet markets is primarily driven by customer demand. Additionally, Van de Sandt (2019) argues that marketing products in another jurisdiction may decrease rather than increase security risks (Van de Sandt, 2019).

Not only the origin and the destination of a product are related with a certain jurisdiction, the location of supporting infrastructures is also of importance. The bitcoin ecosystem is such an infrastructure. In this decentralised payment system, there are a few intermediaries that have the characteristics of a central authority (Moore & Christin, 2013). Bitcoin exchanges facilitate the conversion of cryptocurrencies to fiat money, they can be subjected to regulation and subpoenaed for information on their clients (Meiklejohn et al., 2013; Moore & Christin, 2013). It is decided to list the choice of such intermediaries as separate subclass of security behaviour under 'choice of Online Financial Service Providers' (OFSPs, section 3.1.5). A cybercriminals' decision to user certain OFSPs may not only be related to 'crossing jurisdictions', it also overlaps with trail obfuscation (section 3.1.3) and data minimisation (section 3.1.4) techniques.

3.1.3. Trail obfuscation

Trail obfuscation is defined by Conlan et al. (2016) as: *"the deliberate activity to disorient and divert a forensic investigation"*. Rogers (2006) defines it as *"adding misdirection to the evidence"* which is very similar to the 'evidence counterfeiting' of Sremack and Antonov (2007). Both refer to the creation of false or misleading evidence.

In the darknet market ecosystem, obfuscation techniques are applied to obfuscate money streams in the bitcoin blockchain. Bitcoins (and other cryptocurrencies) facilitate transactions among cybercriminals (Janze, 2017; Kethineni, Cao, & Dodge, 2018). Estimations of the share of bitcoin transactions that are linked with illicit activities range from 1.1% (Chainalysis, 2020), to 10%-30% (Sun Yin & Vatrapu, 2017) or even 50% (Foley et al., 2019). Cybercriminals obfuscate

these financial trails through 'mixing' or 'tumbling' their bitcoins (Van Hardeveld, 2018, p.128; Van Wegberg, Oerlemans, Deventer, et al., 2018). This is a type of 'cooperative obfuscation', which consists of mixing the funds of various users (Narayanan & Möser, 2017). In return for a transaction fee, these services generate a stream of transactions that turn investigations into the money stream into highly complex procedures (Moore & Rid, 2016). The mixing of criminal proceeds obtained through darknet market transactions, is not an irregularity. Janze (2017) shows that usage of transaction obfuscation services is related to the amount of sales on darknet markets. Stronger still, the first Silk Road included mixing functionalities on their platform (Cox, 2016).

Next to mixing services, online gambling sites that accept bitcoins also receive high proportions of illicit bitcoins (Ermilov, Panov, & Yanovich, 2017; Fanusie & Robinson, 2018). The highly popular gambling sites receive huge amounts of relatively small transactions (Athey, Parashkevov, Sarukkai, & Xia, 2016; Lischke & Fabian, 2016) and in the years 2013-2016, much darknet market revenue was sent to these gambling sites (Fanusie & Robinson, 2018). These scholars and others, e.g. Paquet-Clouston, Haslhofer, and Dupont (2019) therefore assume that gambling services are used to obfuscate money trails. However, Meiklejohn et al. (2013) stipulates that - at least in 2013 - gambling sites such as SatoshiDice are not mixing bitcoin effectively. According to the authors, the addresses belonging to SatoshiDice are publicly known and users have to specify a payout address. This makes an permanent link between the bitcoins that are placed as bets and the bitcoins that are paid out.

3.1.4. Data Destruction & Minimisation

With data destruction is referred to the act of *"destroying evidence, either completely or to unanalysable state"* (Sremack & Antonov, 2007). Others name these practices 'artefact wiping' (Van de Sandt, 2019), which is *"the deliberate destruction of data that could be used as evidence"* (Conlan et al., 2016). Often, specialised software is used to permanently destroy files, disks, logs, metadata and removable media using multiple overwrites (Kessler, 2007). Data destruction differs from data minimisation. It entails the neutralisation of evidentiary sources (Harris, 2006). When the 'data footprint' is minimised, there is less data for law enforcement to analyse (Garfinkel, 2007). Because there is no need to destroy evidence when it's not created, it is more of a preventive measure compared to data destruction (Van de Sandt, 2019).

While data destruction and minimisation behaviour is not directly observable on darknet marketplaces, Aldridge and Askew (2017) mention that in the listing descriptions, data handling procedures may be mentioned. Examples given include that vendors may state that addresses are immediately deleted after dispatch or that message logs are not stored on their computers. Additionally, the vendors mention security practices on profile pages, feedbacks and terms & conditions to increase trust Rhumorbarbe, Staehli, Broséus, Rossy, and Esseiva (2016), Tzanetakis, Kamphausen, Werse, and von Laufenberg (2016).

3.1.5. Obfuscation, minimisation & data hiding: choice of OFSPs

As described in section 3.1.2.4, a few intermediaries in the bitcoin ecosystem can subjected to regulation (Moore & Christin, 2013). A bitcoin exchange is an intermediary that functions as a digital currency exchange office. At these exchanges, bitcoins can be traded in for fiat currencies (hard currencies such as euros and dollars). When a cybercriminal wants to convert the bitcoins earned with criminal activities to spendable money at scale, using an exchange at some point is unavoidable (Meiklejohn et al., 2013). Because of the public and transparent nature

of bitcoin transactions, it can be observed to what exchanges cybercriminals transact to (Fanusie & Robinson, 2018).

Bitcoin exchanges can be subjected to regulation and thus, from the perspective of a cybercriminal, form a security risk. For years however, *universal* regulation of these online financial service providers was absent (Reynolds & Irwin, 2017). Only in June 2019, the Financial Action Task Force (FATF) released an update on their recommendations addressing the Anti Money Laundering (AML) challenges associated with cryptocurrencies such as bitcoins (FATF, 2019c). Rules of the FATF are not binding, however the FATF has the authority to sanction countries that do not comply with their directives. The updated 16$^{\text{th}}$ recommendation now specifies that all beneficiaries of all transfers of digital funds are obliged to exchange identifying information (FATF, 2019a, 2019b). Member states are ought to implement the new regulations before June 2020, so the effect of these regulation still remains to be seen.

The current situation is that large inconsistencies in identity verification and monitoring transactions exist, which give way to fraudulent behaviour (Campbell-Verduyn, 2018). For years, bitcoin exchanges did not have to comply to any *universal* anti-money laundering (AML) regulations (Reynolds & Irwin, 2017). AML efforts are dispersed and gaps in global governance exist (Campbell-Verduyn, 2018). The authors show that by 2016, some bitcoin exchanges have adopted self-regulated policies to combat money laundering. These range from voluntary registration, mandatory identification for raising 'suspicion', to providing certified photo ID's and proofs of addresses of each user. In short, the self-regulated firms vary in their AML efforts (Campbell-Verduyn, 2018). Without universally binding rules and the limited self-regulation, national anti-money laundering efforts are ineffective in solving the bigger problem. Bitcoin exchanges relocate their offices to jurisdictions that have less stringent AML requirements, as Van Valkenburgh (2015) and del Castillo (2015) show. Thus, bitcoin exchanges located in these jurisdictions remain capable of laundering money originating from criminal activities (Boxerman & Schwerin, 2016; Möser, Böhme, & Breuker, 2013). While numerous exchanges require identification and apply 'know your customer' (KYC) principles correctly, others do not with the goal of serving clients that prefer anonymity (Moore & Christin, 2013).

A type of exchange that raises the concerns of the FATF is peer-to-peer (P2P), or decentrally organised exchanges. P2P exchanges facilitate transactions between peers directly. In practice, this means that one party transacts bitcoins to a counterparty without storing the bitcoins at central authority first. The counterparty then transfers the agreed amount of fiat currency back to the first party. P2P-exchanges may facilitate a meet-up in person, such that bitcoins can be traded for money in the form of physical currencies (cash). Because no central authority is involved in this transaction, enforcing identity verification is challenging (FATF, 2015). When cybercriminals make use of P2P-exchanges, the usefulness of monitoring their transactions stored in the blockchain is limited (FATF, 2015).

To conclude this section: it becomes clear that cybercriminals may perform data minimisation and trail obfuscation techniques by choosing certain online financial service providers. Some centrally organised exchanges may ask for proof of identity, while others do not. When no proof of identity is required, falsified information can be submitted to the exchange. Decentrally organised P2P-exchanges are suitable for money laundering, because identity checks are weakly enforced and links between illicitly obtained bitcoins and spendable (cash) money are hidden from the investigator.

3.1.6. Theoretical Framework

The security behaviours identified in section 3.1 and how these relate to 'security' as defined in this chapter are summarised in the theoretical framework presented in Figure 3.1. In this figure, the '+' indicates that a behaviour positively influences security, the '−' indicates that a behaviour negatively influences security and the '?' indicates that no or contradicting evidence has been found regarding the relation between security and the behaviour identified. The framework is shortly elaborated upon below.

The acts of **data hiding** found in literature are the following. *Encryption of communication* is influenced by PGP-usage, PGP-key length, usage of the auto-encryption functionality and using other communication services. The latter two security mechanisms may have a positive, or negative influence on security. Auto-encryption increases security of persons who would not use PGP-encryption otherwise. On the other hand, the functionality might become a significant security risk if LEA are able to gain control over the market. Whether the use of alternative communication platform increases or decreases security, depends on how secure these platforms are. *Authentication security* is influenced by the usage of 2FA, password strength and password uniqueness. These security practices all contribute to secure trading on darknet markets. To what extend *crossing jurisdictions* in the form of cross-border shipping increases security is not clear from literature. Academics make contradicting statements regarding this practice. *Actively increasing linkability* and *passively allowing linkability* both have a negative impact on security of the cybercriminal. **Data minimisation & destruction** behaviour is not directly observable in darknet market data. Some vendors mention these security practices in their listings, profile descriptions and terms & conditions. It is unknown whether these practices are consistently followed by vendors that mention them. **Trail obfuscation** behaviour that increases security consists of the absence of any clear links to central exchanges and the use of mixing services. Contradicting statements are found regarding the effectiveness of obfuscating money trails by sending bitcoins through gambling services. Lastly, practices that consist of a **combination of behaviours** (data hiding, minimisation & trail obfuscation) are identified. The first practice identified is using exchanges with weak AML/KYC controls. These exchanges are located in jurisdictions that do not take part in international agreements. When AML/KYC controls are weak, identification is not needed or falsified information may be used for this cause. The second practice identified is making use of P2P-exchanges. These exchanges are not directly involved with the financial transactions and often do not have strong AML/KYC controls. Exchanging cryptocurrencies to spendable money using transactions facilitated by P2P-exchanges, results in hidden links between payment systems and opportunities for not having to register (correct) personally identifiable data.

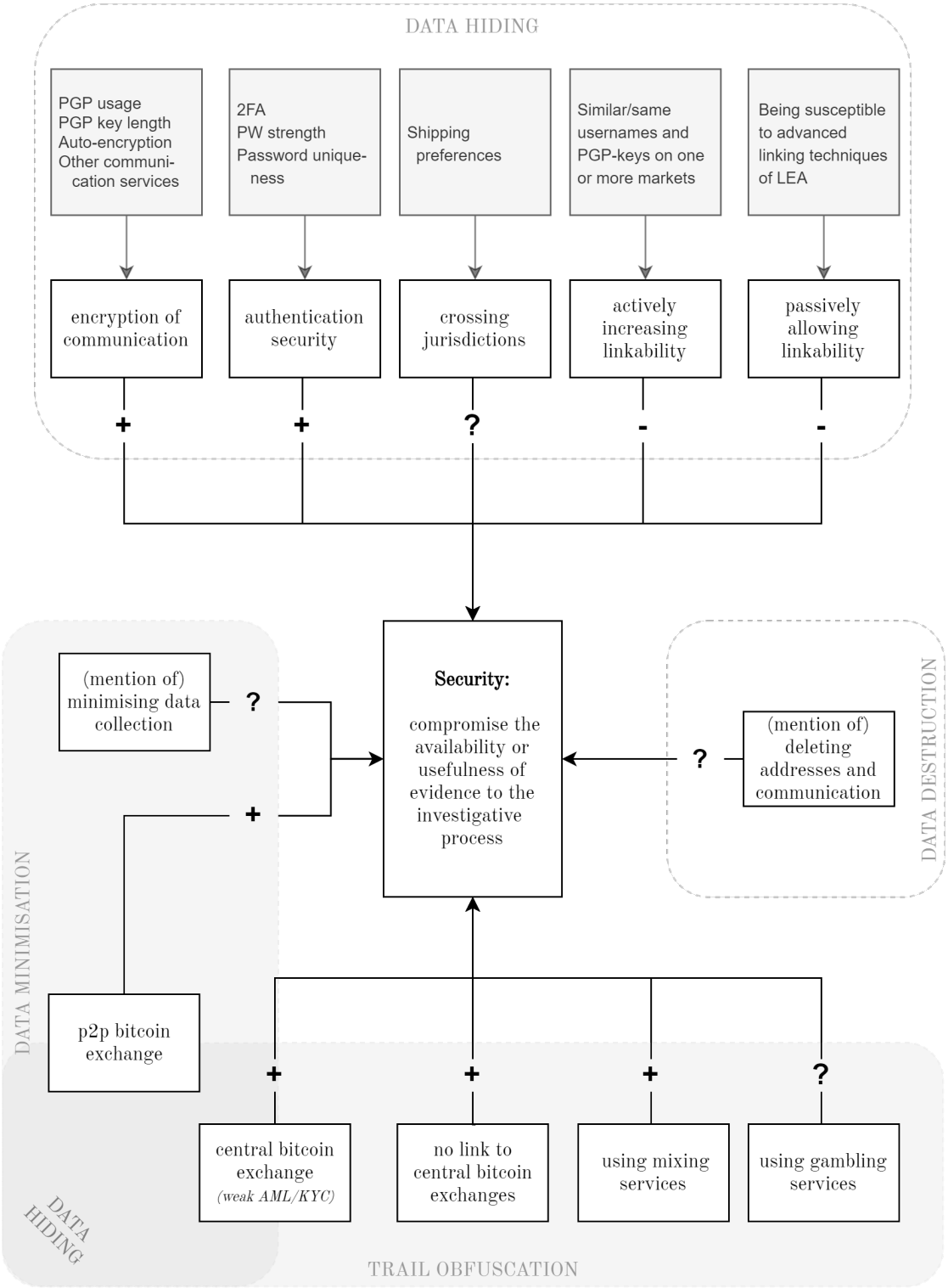


Figure 3.1: Synthesis of security behaviours into a theoretical framework

3.2. Vendor characteristics

This thesis aims to research to what extend characteristics of vendors relate to security behaviour. So far, only a distinction is made between 'buyers' and 'vendors'. Within the vendor population, vendor types are distinguished by grouping vendors together that are similar regarding certain characteristics (chapter 2). This section elaborates on what characteristics can be observed in darknet market data. The findings below heavily draw from both Grapperhaus (2019) and Van Wegberg et al. (2020). These works represent similar efforts of clustering darknet market vendors into vendor types using characteristics of vendors.

Experience Experience entails for how long a criminal has been involved in a criminal market (Décary-Hétu & Leppänen, 2016). Both in darknet market context (Grapperhaus, 2019; Paquet-Clouston, Décary-Hétu, & Morselli, 2018; Van Wegberg et al., 2020) as in conventional crime networks (Morselli & Tremblay, 2004) the duration of criminal involvement is a distinctive factor. Some works assess the experience of a darknet market vendor by subtracting the date of a vendors' first sale from the date of their last sale (e.g. Van Wegberg et al., 2020), others do so by counting the months in which vendors had at least one sale (Christin, 2013; Grapperhaus, 2019). Experience can also be gained on other markets. When a vendor is active on other markets, scholars refer to it as a 'multi-homing vendor' (Calis, 2018; Grapperhaus, 2019).

Loyalty Décary-Hétu and Quessy-Doré (2017) research the loyalty of buyers to vendors on darknet markets. Loyalty can be described as buyers showing repeated buying behaviour while having other opportunities for buying another product (Anderson & Srinivasan, 2003). In a competitive online market, creating loyalty is essential to remain successful (Gefen, 2002). On digital marketplaces, trust, satisfaction and price fairness are determinants of loyalty (Doong, Wang, & Shih, 2008). Grapperhaus (2019) regards loyalty as a 'passive' vendor characteristic, which cannot be directly influenced by the vendor itself.

Reputation Reputation is expressed as the ratio of positive/negative feedbacks. It is seen as an important asset to the cybercriminal because it signals trustworthiness (Lusthaus, 2012). Research has repeatedly shown that on darknet markets (Grapperhaus, 2019; Van Wegberg et al., 2020) and on online fora facilitating illicit trade (Décary-Hétu & Leppänen, 2016; Holt et al., 2016), high reputations correlate with high numbers of sales.

Digital & Physical sales Roughly two types of products are sold on darknet markets. Firstly, physical products of which most of are drugs such as cocaine, cannabis, heroin or other psychoactive substances (Dolliver et al., 2018). Drug trade facilitated by darknet markets is a global phenomenon primarily driven by demand (Dittus et al., 2018). The authors argue that for most drug trade, traditional supply chains remain intact. Next to physical products, digital products are transacted as well. Van Wegberg, Tajalizadehkhoob, et al. (2018) provide a first insight in what products are transacted on darknet markets. The authors differentiate between business-to-business cybercrime offerings, such as botnet related items (tutorials, source codes, DDoS services), databases of e-mail addresses and personal details to be used for spear-phishing campaigns, exploits and malware such as ransomware or keyloggers. The largest category of digital items sold are 'cash-out offerings'. These include credit card details, or listings that aim to recruit money mules. While not explicitly mentioned as a vendor characteristic in other academic work, current research tends to focus on either vendors that offer digital goods, i.e. cybercrime offerings such as hacking tools, malware and cash-out solutions (Van Wegberg et al., 2020; Van Wegberg, Tajalizadehkhoob, et al., 2018) or physical goods such as drugs (e.g.

Broséus, Morelato, Tahtouh, & Roux, 2017; Van Buskirk, Naicker, Bruno, Breen, & Roxburgh, 2016). The criminal context of drugs and cybercrime offerings cannot be compared directly. As Van Hardeveld (2018) indicates, cybercriminals that are digitally proficient might make mistakes on the physical security which conventional criminals are not prone to make. Vice versa also holds: cybercriminals stemming from a traditional criminal background might make mistakes on the technical part of their security (p.154 Van Hardeveld, 2018).

Diversity Diversity refers to the amount of categories a vendor has listings in. Three ways to express diversity are found throughout literature. A method to calculate a 'diversity index' (Grapperhaus, 2019; Paquet-Clouston et al., 2018), a count of the number of categories a vendor is active in (Décary-Hétu & Quessy-Doré, 2017; Grapperhaus, 2019) and a a dichotomous variable whether a vendor is active in multiple categories or not (Van Wegberg et al., 2020). The diversity index correlates highly with the number of categories a vendor is active in (Grapperhaus, 2019, p.46).

3.2.1. Vendor characteristics & Security

Firstly, it is expected that there is a relation between 'experience' and 'security'. Van de Sandt (2019, p.96) argues that the security practices of cybercriminals are positively influenced by the cybercriminals' experience. However, Van Hardeveld (2018, p.161) argues that vulnerabilities, software updates or new security developments may be ignored by cybercriminals because of a 'status-quo bias'. The status-quo bias describes that one has the tendency to let things (in this case, security) remain the same (Kahneman, Knetsch, & Thaler, 1991). The term is first coined by Samuelson and Zeckhauser (1988), who argue that a status-quo bias is also present when an individual is not aware of the other, updated, options available to him. The behavioural pitfall of remaining the satus-quo with regard to security measures, might become a security risk which consequently may lead to de-anonymisation (Van Hardeveld, 2018, p.42). Secondly, it is expected that a relation between the number of sales and security exists. Van de Sandt (2019) hypothesises that a) investments in security are costly and b) that increasing profits result in higher risks to security, which translates to a need of better security practices to protect the cybercriminal. From this could be inferred that vendors with more business success are expected to have higher security standards.

4

Methodology

> "I just checked the PGP key of a vendor that I had ordered from once. Turns out the encryption subkey of their key only has a strength of 512 bits"
>
> *A pseudonymous cybercriminal on Dread*
>
> Taken from a dark web discussion forum (Dread Forums, 2019)

In the first section of the previous chapter (section 3.1), a collection of security behaviours that can be observed on darknet markets has been identified. These behaviours increase or decrease the availability or usefulness of evidence to the investigative process. This is summarised in the theoretical framework depicted in Figure 3.1. In the second section of the previous chapter (section 3.2), characteristics of darknet market vendors were elaborated upon.

In this chapter, a methodology is presented to further investigate the relationship between vendors and their security behaviour. This chapter is structured as follows. Section 4.1 first discusses the vendor characteristics that are considered in this thesis. Then, it is explained how these characteristics are measured in the Hansa Market data. Hereafter, it is set out how the vendor characteristics are used to cluster similar vendors into vendor types. Section 4.2 revolves around how the security behaviour of vendors is measured. First, it is explained which security behaviours from the theoretical framework are included for further investigation. Next, it is elaborated upon how the selected security behaviours are measured in the Hansa Market dataset. And lastly, section 4.3 summarises the steps described above in the conceptual model. This models shows the relations between concepts that are analysed in this thesis.

4.1. Vendors

In this thesis, the criterium for being a 'vendor' is having at least one sale that is marked as a 'finalised transaction' *or* to have a feedback rating given by a buyer. The latter criterium is necessary due to the absence of any transactional data for certain months. By only regarding finalised transactions, uncompleted orders in which for example no payment has been done or no goods have been shipped are not considered. Additionally, only sales made between 25-09-2015 and 20-06-2017 are included in this research. Before September 25$^{\text{th}}$, the data includes mostly test transactions of administrators and the market was not truly active. After the 20$^{\text{th}}$ of June 2017, the market was taken over by the Dutch Police (Greenberg, 2018). The market remained active for precisely one more moth, but now under full control by the Dutch Police, who had garnered administrator rights (Greenberg, 2018). The reason for discarding the data created

after June 20th, is that as part of Operation Bayonet, the closing of Alphabay resulted in a huge influx in users. These users cannot be typified easily using characteristics based on only Hansa Market data. While only active for a number of weeks on Hansa, they would seem as newcomers, whereas in fact they might be very experienced vendors. Needless to say, it remains possible that vendors included in the analysis gained experience elsewhere on other markets. However, no large scale migrations resulting from Law Enforcement interventions occurred during the selected time frame. Additionally, when analysing the vendors in the Hansa data, it is considered whether the reputation import functionality has been used. This variable accounts for the whether a vendor is active on another market or not. Given these observations, the analysis of vendors based on solely Hansa Market data is expected to be fairly reliable.

4.1.1. Vendor characteristics

Ample academic research on the relation between vendor characteristics and security behaviour is available, making this research of the exploratory kind. In section 3.2.1 indications that the amount of experience of cybercriminals might influence security behaviour are discussed. Therefore, *experience* and *active on other markets* have been included in the conceptual model. Similar to Van Wegberg et al. (2020), experience is defined as the timespan between the first and last transaction, expressed in days. *Active on other markets* is a dichotomous variable and refers to whether vendors made use of the reputation import functionality of Hansa Market. This functionality enables vendors to transfer their reputation gained on other markets to their Hansa Market account. It may be assumed that the Hansa administrators verified whether the account of which the reputation was imported belongs to the same real world entity as the account importing the reputation. Thus, a vendor with an imported reputation is or has been active on other markets. *Active on other markets* is set to '1' for vendors who have an imported reputation.

Next to variables related to the vendors' experience, the relation between business success and security is also of interest. Previously, it is shown that business success is expected to relate to security behaviour (section 3.2.1). To measure business success, three variables are constructed and observed in the data.

First, *revenue* represents an estimation of the vendors' turnover in US dollars. Drawing upon the works of Grapperhaus (2019), bitcoin exchange rates are queried and for about half of the orders the date and time of purchase had to be estimated (see Grapperhaus, 2019). Because this only entails an estimation, other business performance metrics are specified as well.

In this thesis, a distinction is made between sales that need to be physically shipped and those that can be digitally transferred. Vendors transacting digital items are inherently different from those that sell physical goods. The number of physical and digital sales per vendor is counted. Because large gaps data exist, sales between June 2016 and September 2016, December 2016 and February 2017 and the first three weeks of April 2017 are estimated via the number of feedbacks given by buyers. The number of feedbacks are considered to be fairly accurate proxies for the number of sales (Soska & Christin, 2015; Van Wegberg, Tajalizadehkhoob, et al., 2018). For this, the parsed data set by Grapperhaus (2019) has been used.

In the Hansa data, a column indicating whether the advertised product entails an item that must be physically shipped or can be digitally transacted is found[1]. A benefit of using said column, is that the categories in which listings are posted are fairly inaccurate. Physical items are advertised in categories in which only digital items are expected and vice-versa. The amount

[1]While the name of this column does not reveal the function of the column, extensive manual checks resulted in the conclusion that the indicator is completely accurate in distinguishing physical items from digital items.

of physical and digital sales per vendor are then counted in the data and included as the variables *physical sales* and *digital sales*.

4.1.2. Clustering vendors

Using the vendor characteristics *experience, physical sales, digital sales* and *active on other markets* described above, vendors are clustered and assigned 'vendor types'. The estimated *revenue* variable is not reliable enough to cluster vendors with. This is described in appendix A. Therefore *revenue* is not included in the clustering algorithm. After the vendors have been clustered into vendor types, the revenue distribution within clusters is compared between clusters to verify whether the clustering produced results that make sense.

The reasons for clustering vendors into vendor types are threefold. Firstly, clustering is a parsimonious way to capture the multidimensional characteristics of vendors. Simply put, vendor types are easier to interpret than sets of individual characteristics that all have varying parameters. This enhances the interpretability of this research. Secondly, previous work shows that models using 'criminal profiles' perform as good (Grapperhaus, 2019) or even better (Van Wegberg et al., 2020) compared to their counterparts that only include individual characteristics. Thirdly, discussing vendor types instead of individual vendors help in generalising the findings.

In this thesis, it is chosen to perform a Latent Profile Analysis (LPA) using the LatentGOLD statistical package of Vermunt and Magidson (2013). Through LPA, latent profiles based on the vendor characteristics are constructed. Thus, profiles are not defined upfront (the units of analysis are 'made', see chapter 2). A latent profile model maximises homogeneity within clusters and heterogeneity between clusters (Magidson & Vermunt, 2004). Because of four reasons this clustering method is applied in this thesis. 1) *experience, physical* and *digital* are count data and *active on other markets* is of the nominal type. LPA works on any measurement scale: nominal, ordinal, continuous and counts (Magidson & Vermunt, 2004), 2) LPA models are probabilistic in nature. This means that no biases are introduced because of deterministic assignment of cluster centres. For example, this is the case in k-means clustering (Magidson & Vermunt, 2002), 3) statistical tests exist to determine the optimal number of classes and 4) Latent profile analysis has been successfully applied in similar research by Van Wegberg et al. (2020) and has been suggested as further research by Grapperhaus (2019).

For determining the number of resulting profiles four criteria are traded-off. Firstly, the global fit of the model as assessed through the Bayesian Information Criterium (BIC). The BIC tends to favour parsimonious (but underfitting) models (Dziak, Coffman, Lanza, Li, & Jermiin, 2020). Secondly, the local fit through the non-significance of the bivariate residuals (BVR) is assessed. Formally, a BVR is significant when the residual is below the critical χ^2 with one degree of freedom ($\chi^2(1) = 3.84$ with $p = 0.05$). However, since BVR statistics are sensitive to large sample sizes, scholars also work with a reduction of the BVRs relative to the 1-cluster model. A rule of thumb is to require a reduction of 90% instead of striving for non-significance of the BVRs (Notelaers, Einarsen, De Witte, & Vermunt, 2006). Thirdly, whether all variables contribute significantly to the chosen clusters and whether the clusters differ significantly in terms of the vendor characteristics (Van Wegberg et al., 2020). And fourth, to what extend the resulting clusters are easy to interpret in the context of this research and whether the sample sizes remain sufficiently large (Masyn, 2013; Meeus, van de Schoot, Klimstra, & Branje, 2011).

4.2. Security

Departing from the theoretical framework depicted in Figure 3.1, this section discusses why the decision is made not to include all identified security behaviours into the conceptual model. Then, it is explained how the selected security behaviours are measured in the Hansa Market data.

4.2.1. Selecting security behaviours

Security behaviours of which it is not clear whether they certainly benefit or impede security are not further analysed. From the behaviours identified in the theoretical framework (Figure 3.1), the following are excluded from the conceptual model:

- *Auto-encryption usage.* Section 3.1.2.1 highlights that auto-encryption increases security of those that would not use PGP-encryption otherwise, while it also constitutes a significant risk to security in case of Law Enforcement interventions. Hansa users were made aware by the administrators that auto-encryption could not be regarded as being fully secure. Because the use of auto-encryption could be interpreted as either enhancing or diminishing security, the usage of this functionality is excluded from the conceptual model.
- *Other communication services.* In section 3.1.2.1 and 3.1.6 it is reflected upon that other communication platforms may be secure or not secure. Therefore, mentioning alternative communication services in listings or profiles is not included in the conceptual model.
- *Mentioning minimising or destructing data.* As highlighted in section 3.1.4, some vendors mention in their profile descriptions, listings or feedbacks that data collection is minimised and logs of addresses and communications will be destroyed. No research is found that provides empirical evidence that vendors mentioning these practices, actually apply these security mechanisms as well. Therefore, the mere mentioning cannot be equalled to more secure behaviour and is therefore excluded from the conceptual model.
- *Crossing jurisdictions in terms of shipping behaviour.* This behaviour is excluded because of three reasons. First, there is contradicting evidence whether it is beneficial to security to ship to a different jurisdiction than the sender is located in. Décary-Hétu et al. (2016) equals cross-border shipments to a security risk, while Van de Sandt (2019) argues that doing so creates information asymmetries between jurisdictions, which he believes to benefit security of the cybercriminal. Second, as explained in section 3.1.2.4, there are many factors that influence the decision to ship internationally, of which not all of them are security related. Third, this security behaviour does not apply to vendors that sell digital goods. Van Buskirk, Naicker, Roxburgh, et al. (2016) show that most vendors transacting digital goods select 'worldwide shipping', from which no security related decisions can be drawn.
- *Passively allowing linkability* of darknet market pseudonyms is excluded. It would involve advanced content analysis of darknet market listings and profiles to match vendors, which is out of scope of this thesis.

4.2.2. Measuring security behaviours

This section elaborates on the measurement methodology used to observe the selected security behaviours. First, authentication related behaviours (Password strength, password uniqueness, 2FA usage) are elaborated upon. Then, it is explained how the security of encrypted communication could be assessed. The section continues with how the linkability of darknet market pseudonyms is regarded via matching PGP-keys registered on other markets. Hereafter, it is elaborated upon how vendors' choices regarding online financial service providers are observed through analysing transactions in the bitcoin blockchain. Lastly, a simple scoring method allowing for joint consideration of multiple security behaviours is presented.

4.2.2.1. Authentication

Due to the Dutch Police gaining admin rights on the 20$^{\text{th}}$ of June 2017, Law Enforcement was in the position to alter the configuration of the market. This resulted in the ability to retrieve plaintext passwords of Hansa Market users (Greenberg, 2018).

Password strength Two metrics describe password strength: password length and password complexity in entropy bits. Password length is simply the number of characters a password consists of. Generally, a longer password is more difficult to guess. The amount of guesses it takes on average to crack a password, or 'the uncertainty of a password' is described by the password complexity in entropy bits. The entropy of a password is related with its length. Formally put, the complexity of a *randomly generated password* is calculated by Equation 4.1.

$$H_i = \log_2(R^{n_i}), \quad pw_i = (c_1 + c_2 + c_3 \dots c_n) \tag{4.1}$$

This equation represents the following. If each vendor i has a password pw that consists of n characters and R is the amount of all characters recognised by the system, the entropy H is described by taking the $\log_2$ of R^n. So from Equation 4.1 follows that the complexity of a password is derived from a) the length of the chosen password and b) the character set in use ('all characters recognised by the system'). Often, this R is defined in password policies. For example, when both upper and lower letters of the Roman alphabet and the numbers 0-9 are allowed, R equals 62. When also special characters are allowed, $R = 95$. A character set with $R = 95$ is referred to as 'extended ASCII' or 'ASCII-128'. Roughly, the extended ASCII set represents all characters that can be created using keys and combinations of keys on commonly found on keyboards.

To estimate R of the Hansa Market, the amount of unique characters found in the passwords was counted. Surprisingly, this resulted in 511 unique characters. Hence, the administrators of Hansa allowed any *Unicode* character to be used as a password. As to date, there are many ($n \approx 143,000$) Unicode characters in use. This includes Chinese and Arabic characters, but also emoji and all kinds of symbols and signs. Assuming $R = 143,000$ for all passwords, would result in almost any password having extremely high entropy. Certainly, this would overestimate the security of many users. In this thesis it is assumed that most passwords consist of characters from the extended ASCII set and thus R is assumed to be 95.

It is important to recognise that entropy as per Equation 4.1 does not apply to non-randomly generated passwords. Human-generated passwords follow certain trends, which greatly reduces uncertainty. For example, the characters *a, e, r* and *1* are used more often in human-generated passwords than other characters (Burnett, 2006). Thus, these characters add less entropy ($\approx$ uncertainty) to the password. Additionally, in terms of length and style, people follow heuristics (Burr, Dodson, & Polk, 2006). Think of passwords starting with a capital letter and ending

with special characters or numbers. Many scholars attempted to estimate 'the true' entropy of passwords (e.g. Bonneau (2012), Shay et al. (2010)). Estimating true password strength is a complex task to which no single best solution exists (Galbally, Coisel, & Sanchez, 2016). A constant factor in estimating true entropy is penalising the repetition of characters in some way (see Carnavalet & Mannan, 2015). Therefore, the decision is made approach 'true entropy' using Equation 4.2:

$$H_i = \log_2(R^{\#pw_i}), \quad pw_i = \{c_1, c_2, c_3, \dots c_n\} \tag{4.2}$$

By Equation 4.2, the amount of *unique* characters (i.e. the cardinality of set pw) is used for the entropy calculation. When assumed that human-generated passwords have a tendency of repeating certain characters, these passwords would be assigned slightly lower entropies than their randomly generated counterparts. As a consequence of the heuristic presented in Equation 4.2, the entropy of very lengthy passwords (both human and random) might be underestimated: the probability of repeated characters increases with n. However, since H increases exponentially with n, an underestimation of long passwords is regarded less problematic than overestimating the complexity of short passwords. Any 'long' and not very simplistic password will have a 'high' entropy, whether it is estimated via Equation 4.1 or Equation 4.2.

The analysis of password strength is as follows. First, the distributions of the password strength metrics are visualised for both vendors and buyers. This shows how vendors' password strengths compare to those of buyers. Hereafter, the focus is only on the vendors and their password complexity (entropy bits). The characteristics of vendors are visually correlated with password complexity. Next, statistical tests are performed on the relation between vendor types and password complexity. To statistically determine whether there is any difference between the mean password complexity of vendors grouped per vendor vendor type, a one-way Analysis of Variance (ANOVA) is performed. This is an *omnibus test*. Thus, it only shows whether at least two vendor profiles are significantly different in terms of password complexity. Therefore, an additional post-hoc test is ran to learn how the types of vendors significantly differ from each other. The post-hoc test of choice is the Tukey-Kramer HSD-test, because it is recommended for unbalanced designs, in which the variances are assumed equal and the assumptions for the ANOVA-test have been met (Westfall, Tobias, & Wolfinger, 2011).

Password uniqueness The 'Have I Been Pwned' database includes more than 10 billion leaked passwords, of which 573 million are unique. This extensive database of leaked passwords is used to analyse to what extent Hansa users use non-unique passwords. The SHA1-hashes of the passwords in the PWND database are made available for research. To find matches, the plaintext passwords of Hansa are hashed and compared with the hashes in the PWND database.

A password match is a potential security risk in two ways. First, Law Enforcement may find email addresses and usernames of passwords that have been matched. If indeed a cybercriminal reuses his or her password on websites that suffered from data leaks and the matched password is fairly unique, the criminal may be de-anonymised quickly. Second, matches of SHA1 hashes that occur often in the PWND databse, show that passwords are common. Common passwords are easy to guess, they are not uncertain and thus have a very low 'true entropy'.

The analysis of the password matching is structured as follows. First, the percentage of password matches of buyers is compared to that of vendors. The analysis continues by only regarding the passwords of vendors. Descriptives are generated on how many matches are found. Occurrences of 1-9 matches, 10-99 matches and 100+ matches are counted and visualised. Doing so, creates an insight in to what extend 'fairly unique' passwords are reused on other websites by possibly the same entity as the vendor and to show how often 'common' and 'very common' passwords are used

by vendors. While this grouping provides insight in what kind of security risks are potentially introduced by reusing a password, it also is an arbitrary way of grouping. Considering that any type of password match is a display of non-secure behaviour, it is decided that statistical testing of password reuse behaviour is only performed on whether there is a password match has been found or not. It is statistically determined how vendor types perform relatively to each other, in terms of the proportion of vendors that have non-matched passwords. To achieve this, a χ^2-test of proportions is performed, followed up by a z-test as post-hoc. The z-test is performed pairwise. It tests whether the percentages of matches are statistically different between pairs. Pairwise post-hoc testing is done using the False Discovery Rate (FDR-BH) procedure of Benjamini and Hochberg (1995). This method is preferred over a Bonferroni correction. The Bonferroni method combined with 10 pairwise comparisons yields very conservative results.

Two-factor authentication In the Hansa Market data could be observed which users enabled Two-Factor Authentication (2FA). The analysis regarding this security behaviour is therefore rather straightforward. First, the amount of vendors that use 2FA is compared with the number of buyers that make use of this additional layer of security. Then, the distributions of the characteristics of vendors who enabled 2FA are visually compared with those who did not enable 2FA. Lastly, a χ^2-test of proportions and post-hoc z-tests show how the vendor profiles perform relatively to each other in terms of the proportion of vendors within each profile that enabled 2FA.

4.2.2.2. Encryption of communication

As explained in section 3.1.2, the PGP encryption protocol is widely used on darknet markets to hide the content of the communication between darknet market users. The PGP-keys are retrieved from the Hansa data and analysed with a Python implementation of GNU Privacy Guard (GnuPG). GnuPG is a command line tool used for PGP encryption, signing and key management. All metadata stored in the public PGP-keys found in the Hansa Market data was extracted. This way, it is observed when each key was created, the algorithm and key length used. From preliminary analysis was concluded that the majority of the keys have a key size of 2048 or 4096 bits. Both these key sizes are considered 'secure', according to most recent NIST specifications (Barker & Dang, 2015). A number of keys had peculiar key sizes of 2047 or 4095 bits. A 2047-bit key falls 1 bit short of the NIST recommendation. Appendix C.1 explains that these are in fact no 'mistakes' made by users nor do these aberrant key sizes decrease security significantly. In this thesis, such keys are equalled to the commonly found key sizes (e.g. a 2047-bit key is considered to be a 2048-bit key).

The analysis of the encryption of communication behaviour is structured as follows. First, the adoption of PGP-encryption is compared between vendors and buyers. Then, the algorithms used within the PGP protocol are regarded. Because many key sizes were found, the key sizes are grouped by <2048-bits (below the NIST-threshold), 2048-bits and 2048> bits. The key strengths are compared to the creation dates of the keys to learn whether the PGP-keys increase in cryptographic strength over time. Because only a few vendors use keys that do not meet NIST-recommendations, further analysis only regards two groups: vendors with PGP-keys with length $\leq$ 2048 bits and 2048> bits. In terms of security risk, the difference between a 2048-bit key and 2048+ bit key is negligible (Lenstra, 2004). However, it is interesting to observe which vendors are particularly cautious about their security. Especially in conjunction with the other security behaviours analysed in this thesis, it indicates what types of vendors are more security aware than others. To statistically determine which vendor types have relatively high proportions of vendors with extremely secure key sizes, a χ^2-test and pairwise z-test as post-hoc is performed.

4.2.2.3. Actively increasing linkability of DNM pseudonyms

The security risks darknet market vendors are willing to take by allowing their darknet market pseudonyms to be linked, is assessed by matching their public PGP-keys over different markets. It is chosen to match PGP-keys and not usernames because of two reasons. First, usernames are signals that are more easily faked (see section 3.1.2.3). Second, usernames have been omitted from the Hansa data because of potential breaches of privacy[2].

The database of the currently no longer available *Grams* darknet market search engine is used to regard which vendors can be linked by means of their PGP-key. In this analysis, only vendors that are known to be active on other markets, i.e. those with a successfully imported reputation, are regarded. If a vendors' PGP-key cannot be matched in the Grams data, or the only match is linked to a Hansa account, this indicates that the vendor uses another PGP-key for his or her other vendor account(s).

Because non-matches can also be the result of the Grams database not being complete, the coverage of Grams is estimated by calculating what percentage of all the Hansa PGP-keys can be found in the Grams data.

4.2.2.4. Online Financial Service Providers

As stated in section 3.1.3, bitcoin transactions are traceable in the publicly available blockchain. This is leveraged to investigate what darknet market vendors do after sales have been finalised and revenue has been generated. This section discusses how it can be observed what types of online financial service providers (OFSPs) are popular among vendors. Security-related decisions are inferred from the choice of OFSP. Firstly OFSPs may introduce potential security risks. These central entities in the bitcoin ecosystem can be subpoenaed for information on its users. Secondly, they can also indicate relatively secure behaviour, this is the case if vendors transact to P2P-exchanges or mixing services.

Raw bitcoin blockchain data includes of a log of transactions between bitcoin addresses. It does not show any context to facilitate sense-making of these data (Haslhofer, Karl, & Filtz, 2016). Commercial an non-commercial tools are available that provide this context (Hinteregger & Haslhofer, 2018). These tools apply extensive analyses to the raw bitcoin data to cluster addresses belonging to the same real-world entities. This enables them to track money flows (Haslhofer et al., 2016).

To observe what darknet market vendors use which financial services, the API of *Chainalysis* is queried. Chainalysis provides the much needed contextual information on the transactions performed by each vendor. A custom API-script queries the vendors' payout addresses. The payout addresses are found in the Hansa data, to these addresses the market transfers the revenue generated by sales[3]. The script requests Chainalysis to return which clusters ('real-world entities') are associated with these addresses.

More often than not, Chainalysis returns that it cannot link an OFSP to the address or cluster of addresses. In this case, two explanations are valid. Firstly, the address belongs to an OFSP but is not recognised as such. Some service wallets are not identified by Chainalysis. For example, it is in the best interest of operators of mixing services to design their mixers as such that their service wallets will not be recognised as being part of a mixing service. And secondly, the address

[2] The ethical considerations of this research are discussed in section 8.4

[3] These are *not* the multisignature addresses Hansa created for each new order. The addresses do not start with the leading '3' as is the case with all P2SH-enabled multisignature addresses. In addition, one of these addresses has been used as evidence in court where it was confirmed to be the 'final' payout address of a vendor. Also, unlike for example AlphaBay, Hansa did not create a 'Hansa Wallet' for each new user. Payout addresses on Hansa are vendors' own.

does in fact not belong to an OFSP. As indicated by cybercrime investigators, most vendors do *not* send their criminal proceeds directly to a service wallet. Their earnings are first accumulated on a privately owned (hardware) wallet.

When analysing an address that is not directly linked with a known entity, it is of utmost importance to separate the first scenario (a non-recognised service wallet) from the second scenario (vendor uses a private wallet). When a service wallet is mistaken for a private wallet, transactions of other people than the vendor could be analysed. The clustering heuristics of Chainalysis are not completely transparent. For a large part, Chainalysis makes use of co-spend clustering. Co-spending is when two addresses engage in a single outgoing transaction (Harlev, Sun Yin, Langenheldt, Mukkamala, & Vatrapu, 2018). If this happens, one can assume that these addresses belong to the same real world entity. Still, knowing that two or more addresses belong to the same entity, does not imply any knowledge on who or what this entity represents.

Chainalysis documents that unknown clusters with more than 500 addresses should be interpreted as belonging to services. Because the accuracy of the analysis performed in this research would be greatly reduced if any unrecognised service wallets are interpreted as private wallets, it is decided that only unknown clusters with 1 to 5 addresses are assumed as private wallets. This is on the conservative side, since vendors could well be managing more than 5 bitcoin addresses which have co-spended.

When the number of addresses in the unknown cluster exceeds five, it is stated that it remains unknown whether this cluster is a service wallet or a private wallet. These are labelled 'unknown' and are not further analysed. The outgoing transactions of the assumed private wallets are queried at Chainalysis. The receiving parties of these transactions that are associated with OFSPs are considered for further analysis. The full algorithm is described in pseudo code 1.

The analysis of OFSP is as follows. Descriptives are generated on what services are most transacted to over time. The vendor characteristics are visually correlated with found links to the OFSPs. If a vendor has transacted one or multiple times to a type of OFSP, e.g. a P2P-exchange, the P2P-exchange link is set to '1' for this vendor. As such, it does not matter how many times a vendor has transacted to this type of OFSP. From LEAs perspective, this is a sensible choice. One link is enough to request data at this central entity. As pseudo code 1 shows, it is difficult to determine whether an unknown cluster belongs to an individual vendor or whether it is in fact a service wallet used by multiple people. The observed *direct* links between vendors and OFSPs however (line no. 3-5 in pseudo code 1), are very certain[4]. When the profits are directly transacted to a central exchange, investigators are presented a solid reason to subpoena the exchange. Later on, if the bitcoin address can be tied to an identity, such a direct link can be used as undisputable evidence in court. The non-secure behaviour of directly transacting darknet market payouts to an exchange is analysed to a further extend. First, it is visually depicted how often and at what points in time of their careers vendors makes such mistakes. Then, to statistically determine which vendor profiles have relatively high proportions of vendors that transact directly to exchanges, a χ^2-test and pairwise z-test as post hoc is performed.

[4]Still, the possibility should be taken into account that vendors use intermediaries such as money mules.

Pseudo code 1 Get OFSPs linked with payout addresses

```
1:  procedure CHAINALYSIS(addresses)                    ▷ Query bitcoin payout addresses
2:      for all addr ∈ addresses do
3:          get cluster from API
4:          if cluster ≠ None then
5:              OFSP ← cluster                          ▷ Transacts directly to OFSP
6:          else
7:              get size of cluster from API
8:              if size ≤ 5 then                        ▷ Assume small clusters are priv. wallets
9:                  get exposure of addr from API
10:                 services ← all OFSPs in exposure
11:                 if services ≠ None then
12:                     OFSP ← services                 ▷ Transacts via priv. wallet to OFSP
13:                 else
14:                     OFSP ← "no exposure"
15:                 end if
16:             else
17:                 OFSP ← "unknown"                    ▷ Either a private or service wallet
18:             end if
19:         end if
20:     end for
21:     return (addr, OFSP)
22: end procedure
```

4.2.2.5. Joint analysis of security behaviours

In section 4.2.2 all security behaviours are explored and analysed separately. First, relations between security behaviours are correlated in a correlation table. Then, to analyse all security behaviours jointly, a simple scoring function is calculated. Vendors are awarded one point for each of the satisfied criteria:

- A higher password complexity than median password complexity
- No match in the PWND password database
- 2FA Enabled
- PGP key strength of 2048+ bits
- No direct transactions to central exchanges

Note how the PGP-match in the Grams database is omitted from the scoring method. This security behaviour only applies to vendors that are active on other markets. If this behaviour would be included, the fact whether vendors are active on other markets is measured instead of their actual security behaviour. While this scoring method has its obvious limitations (it assumes that very different security behaviours are equal), it does separate the more security conscious vendors from the vendors that often show relatively non-secure behaviour.

4.3. Conceptual Model

Figure 4.1 represents the conceptual model, in which the concepts of interest and the relations between them are visualised. The vendor characteristics *experience, active on other market places* and *business success* in terms of *physical sales* and *digital sales* are captured in the vendor type concept. By comparing how the vendor types perform relatively to each other in terms of the defined security behaviours, approximately causal relations between vendor type and security behaviour are inferred (chapter 2).

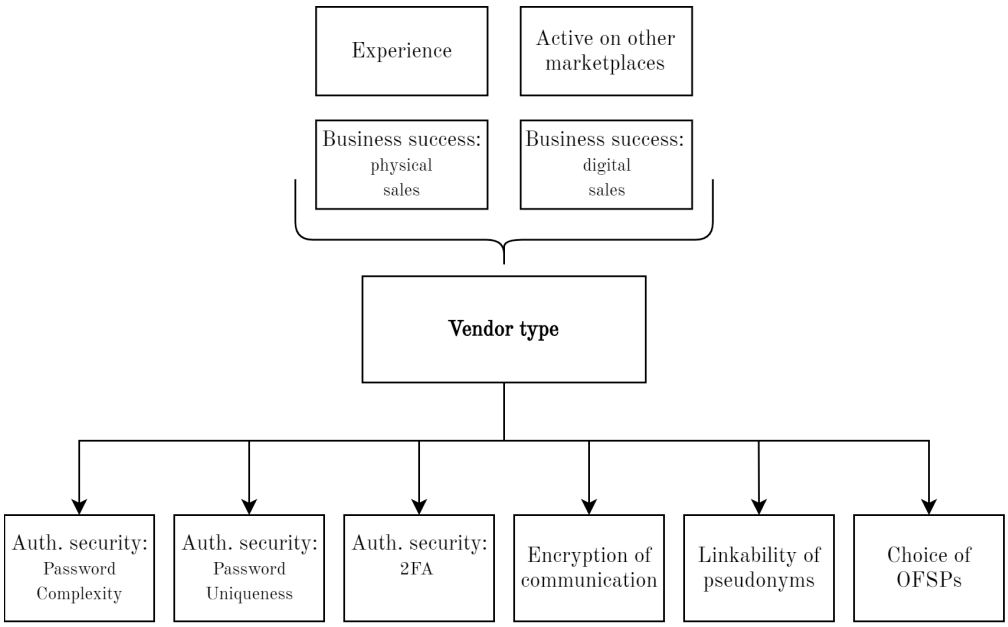


Figure 4.1: Conceptual model

Table 4.1 summarises the concepts that are central to this thesis and how these concepts are measured, as explained in section 4.1 and section 4.2. A distinction is made between the vendor related concepts (upper half) and the security related concepts (lower half).

Table 4.1: Concepts and measurement of concepts

Concept	Measurement of concept	Units	Type
Experience	Transaction date t, $t_{max} - t_{min}$	days	count
Activity on other markets	Reputation import observed in data	yes/no	dichotomous
Business success	Sum of finalised physical sales (or feedbacks)	sales	count
Business success	Sum of finalised digital sales (or feedbacks)	sales	count
Vendor type	Groups of vendors constructed using LPA	types	nominal
Auth. security	Pwd. complexity, see Equation 4.2	entr. bits	continuous
Auth. security	Pwd. uniqueness, SHA1-match in PWND	yes/no	dichotomous
Auth. security	2FA, observed in data	yes/no	dichotomous
Encrypt. of comm.	PGP key strength, $\leq$2048-bits or >2048-bits	yes/no	dichotomous
Linkability of psd.	PGP-match in Grams database	yes/no	dichotomous
Choice of OFSPs	Direct links to central bitcoin exchanges	yes/no	dichotomous

5

Vendor analysis

> "I dont think they would waste their limited resources to go after small fish [...] It would make much more sense to plan to bust big vendors who make millions [...] How would the average tax payer react if they knew LE spent 100+ man hours to bust Mary in Oregon who ordered $40 worth of weed from the internet."
>
> *A pseudonymous cybercriminal on Evolution Forums*
>
> Taken from darknet forum scrapes, made available by Branwen (2020)

5.1. Vendors on Hansa Market

As formulated in section 4.1, the criterium for being a 'vendor' is having at least one sale that is marked as a 'finalised transaction' *or* to have a feedback rating given by a vendor. This results in 1733 users being considered as vendors. Remarkably, this includes 160 vendors that are registered as buyers in the Hansa Market administration. From this is concluded that these users used to be vendors and downgraded their accounts to regular member accounts at a later point of time. The majority of these users are inexperienced and rather low selling vendors. With the data at hand, it is not possible to test whether the switch from being a registered vendor to a regular buyer might be motivated by security considerations. Stronger still, it would be no surprise if for most users retrieving the vendor bond was the main motivator behind their decision. For now, the main takeaway is that the group of 'used to be vendors' ($n = 160$) are regarded as regular vendors.

These 1733 vendors, active between the 25$^{\text{th}}$ of September 2015 and the 20$^{\text{th}}$ of June 2017 (as explained in section 4.1), accumulate to 321,457 finalised sales in total. In this thesis, a distinction is made between sales that need to be physically shipped and those that can be digitally transferred. Physical goods account for 209,411 sales and digital sales for the remaining 112,046 sales. In total, it was estimated that $33.1M of revenue was generated in the selected time period.

A notable growth of both physical and digital sales is observed when the amount of sales and estimated revenue are aggregated over months (Figure 5.1). While the generated revenue through physical sales follows the same trend as the number of sales, digital sales seem to generate less revenue per order.

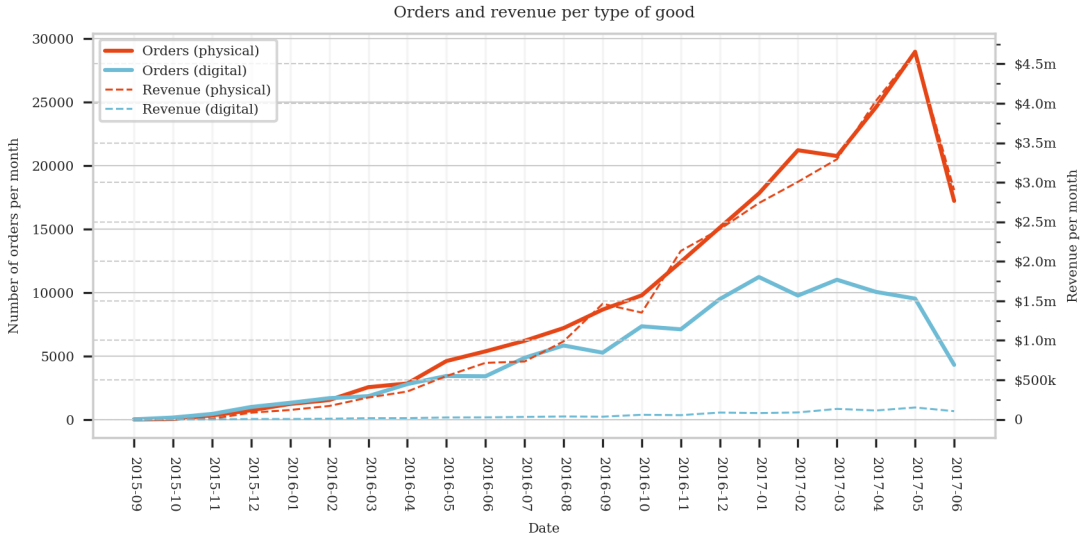


Figure 5.1: Amount of orders and estimated revenue in USD

5.2. Vendor characteristics

5.2.1. Experience & Activity on other markets

The experience of a vendor is defined by the amount of days between a vendors' first and last sale. About 50% of the vendors is active for less than 80 days, while approximately 13% has been active on Hansa Market for longer than a year (Figure 5.2). The peak at the left shows that a large share ($n = 123$) of the vendor population has equal to or less than 10 days of experience. By definition, these include vendors that have made just a single sale in their career. An average of experience of $\mu = 147.63$ days is measured and the distribution of experience has a standard deviation of $\sigma = 159.92$. Of the 1733 vendors 52.3% gained experience on other markets, as indicated by the reputation import variable.

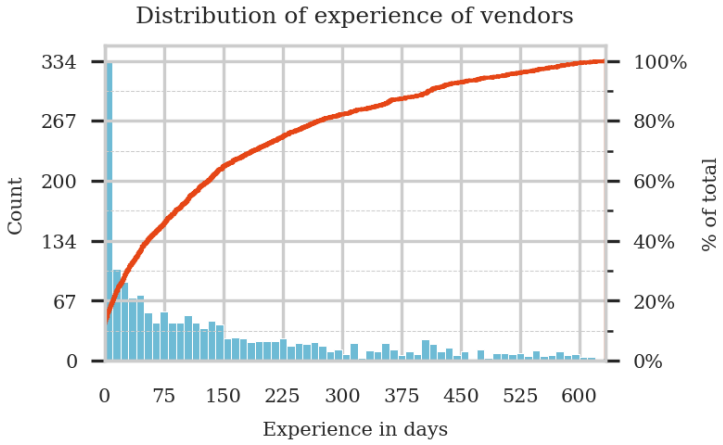


Figure 5.2: Experience of vendors. Horizontal axis grouped per 10 days, left axis shows amount of vendors per 10 days of experience. Right vertical axis shows the empirical cumulative distribution of vendors' experience.

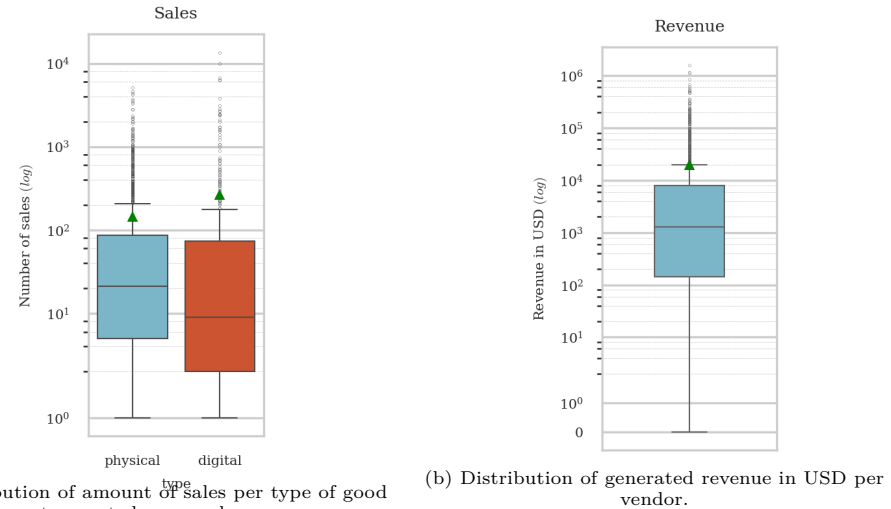
5.2.2. Sales

There are 266 vendors which solely sell digital goods opposed to 1305 vendors that completely focus on transacting physical products. Besides these, also vendors selling both types of goods

exist ($n = 162$). To give a little insight in what types of vendors these are, two examples are presented that were found through manual inspection of the data: a) there are vendors that are specialised in selling drugs (physical) that also sell guides on safe drug use (digital), b) a vendor offering counterfeited identification documents, which are brought to the markets as both physical cards or digital scans. Generally put, it seems that the high selling vendors tend to specialise in either of the product types and happen to have a few sales in the other sale category. The low selling vendors (e.g. <100 sales) have equal amounts of physical and digital sales more often.

After examination of the products offered by the vendors that have both types of sales, it was concluded that this is not an important distinguishing feature of vendors. However, the fact that a fair share of the vendors ($\approx 9.3\%$) sells in both categories does have implications for the analysis. Because of this, it is not possible to estimate 2 mutually exclusive clustering models, i.e. one for vendors transacting physical goods and the other for those selling digital items. Manually assigning a dominant category per vendor is infeasible.

From the distribution of the amount of sales (Figure 5.3a) can be inferred that vendors selling physical items have more sales in general. In Figure 5.3a only vendors that have at least one sale in the physical/digital category are included in the distribution of physical/digital sales: consequently, the minimal amount of sales per vendor per category cannot be equal to zero. At the upper limits, vendors with exceptionally high numbers of sales can be observed. There are 40 vendors that have 1000+ physical sales, this number of digital sales is reached by 26 vendors. The highest amount of sales per vendor are found in digital sales category. Five vendors manage to exceed the 5000 sales. When interpreting digital sales, it should be taken into account that these may be underestimated, since the sales are partly based on given feedbacks. Users are more reluctant in giving feedback on digital products.



(a) Distribution of amount of sales per type of good transacted per vendor.

(b) Distribution of generated revenue in USD per vendor.

Figure 5.3: Distribution of sales and revenue per vendor

5.2.3. Estimated revenue

The revenue in USD has been estimated following Grapperhaus (2019), in which the bitcoin exchange rate at the time of purchase is used to convert bitcoins to dollars. To estimate the revenue in the event of missing data, information from the feedbacks is used to infer time of purchase and the transaction value. The administration of Hansa Market does include a cumulative count of revenue generated in bitcoins per vendor. This amount is used to estimate how accurate the revenue estimation presented here is (Appendix A). There is a trend in which the turnover

is being overestimated. Unfortunately, it is also unknown how accurate the cumulative bitcoin earnings per vendor are. It falls beyond the scope of this thesis scrutinise the revenue estimation even further. As a consequence, it is decided to refrain from using revenue as a parameter in the clustering algorithm. Instead, revenue is only used as a robustness check of the clustering algorithm and is interpreted with care when related to the security behaviours.

From the distribution of revenue (Figure 5.3b) can be learnt that about half of the vendors earns more than $1000. About a quarter of the vendor population is estimated to generate more than $10,000, while the other quarter of vendors earns approximately less than $100 in their Hansa Market careers.

5.3. Clustering Vendors

In order to capture multiple vendor characteristics in one variable, the vendors are clustered into vendor profiles through Latent Profile Analysis. In this thesis, the goal is to create homogeneous groups of vendors in as few as possible clusters. It is left to the LPA algorithm to classify the 162 vendors that have both physical and digital sales (see section 5.2.2) in the most cost-effective way.

No model with less than than 10 clusters showed a perfect *global* fit (see under 'BIC' in Table 5.1). Even the Bayesian Information Criterium (BIC), that tends to favour parsimonious and underfitting models (Dziak et al., 2020), did not provide a definitive answer which model with ≤ 10 clusters to select. This is deducted from the fact that each n-cluster model is outperformed by a $n+1$-cluster model, as indicated by their lower BIC scores.

However, achieving full heterogeneity between all clusters is not the most important goal of clustering vendors into vendor profiles. The resulting clusters should be easy to interpret in the context of this research and the sample sizes should be sufficiently large (Masyn, 2013; Meeus et al., 2011). These considerations are important, especially when fitting criteria do not minimise (Collins & Lanza, 2009).

As vendors that specialise in digital items are inherently different from vendors that transact physical items, it is important that these vendor types are separated well. In Table 5.1, under Sig. BVR, is shown that for 5+ cluster models, the local independence assumption holds for the Physical sales - Digital sales pair. This means that within each cluster, the amount of physical and digital sales are statistically unrelated. This is desirable, otherwise it is difficult to attribute potential security differences to having physical or digital sales.

Because a parsimonious model that does differentiate between physical and digital sales is preferred, the 5-cluster model is chosen. All vendor characteristics contribute significantly to the clustering model and thus should be retained (Appendix A.2). Moreover, while 5 bivariate residuals remain significant, 4 out of 6 variable pairs do show a reduction of approximately 90% when the 5-cluster model is compared to the 1-cluster model. Striving for such of BVRs instead of non-significance is valid when sample sizes are large (Notelaers et al., 2006). The BVRs and their reductions are further elaborated upon in Appendix A.2. In this appendix it is discussed how the poor global fit is probably caused by difficulties in separating the effects of the *active on other markets* variable from the other vendor characteristics. The 6-clusters model is not preferred. It results in two small clusters and additional granularity is added where it is not most relevant (Appendix A.3).

These five clusters are named `Novices`, `Drug Dealers`, `Drug Lords`, `Digital Fraudsters` and `Cybercrime Elites`. These labels will be clarified by first by discussing the distributions of the

vendor characteristics *within* the clusters. Then, the categories in which sales have been made are further elaborated upon. This gives insights in the type of vendors within the clusters to and helps with describing and naming the clusters.

Table 5.1: Clustering fit

Model	$BIC(L^2)$	Sig. BVR*	Total BVR	Min. n	Max n	2nd lowest n	2nd highest n
1 cluster	1761741	6	384060	1733	1733	1733	1733
2 clusters	951620	6	52944	214	1519	1519	214
3 clusters	570490	6	13138	51	1403	279	279
4 clusters	385285	6	6806	38	986	131	578
5 clusters	294230	$5^{\dagger}$	3861	23	983	102	510
6 clusters	228205	$4^{\dagger\times}$	2314	21	870	30	537
7 clusters	189753	$5^{\dagger}$	2657	5	868	23	528
8 clusters	151411	$5^{\dagger}$	1283	5	625	23	535
9 clusters	128918	$5^{\dagger}$	1618	5	622	17	520
10 clusters	114286	$5^{\dagger}$	1403	5	505	17	420

* BVR > 3.84

† Non-significant bivariate residuals between Physical and Digital

$^{\times}$Non-significant bivariate residuals between Experience and Digital

5.3.1. Distribution of vendor characteristics within clusters

The biggest cluster, `Novices` ($n = 988$), distinguishes itself from the other clusters by a relatively low amount of physical and digital sales (Figure 5.4), the lowest number of days of experience (5.5a) and a majority of has a revenue below approximately \$150 (Figure 5.5b). Only 40.2% of the users imported their reputation from other markets. This is the lowest compared to all other clusters and below the market average of 52.3%. It should be noted that this cluster also contains vendors which are estimated to have earned more than \$10.000. On the other hand, revenue estimations might be overestimated as shown in Appendix A. Most vendors sell physical goods, but a few vendors do have digital sales. No vendors with more than 100 physical or digital sales are included in the `Novices` profile.

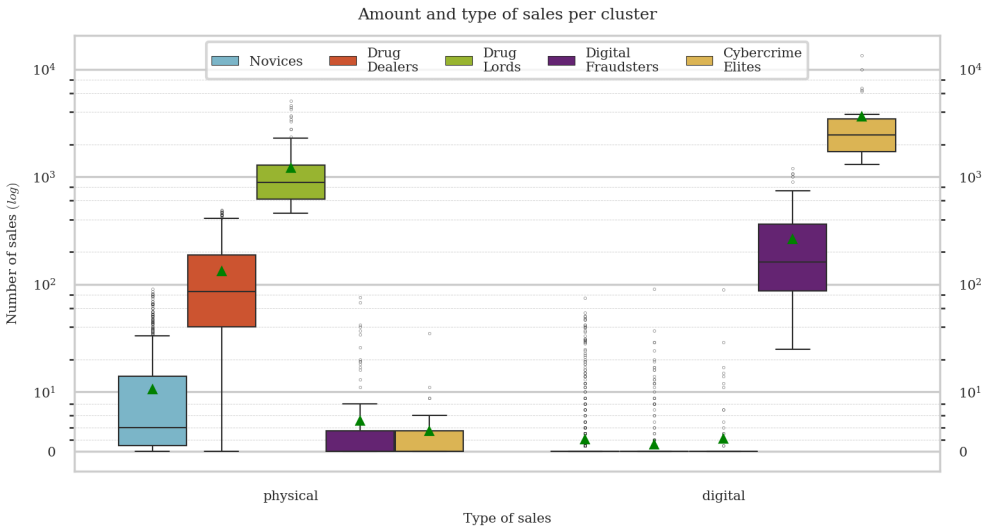
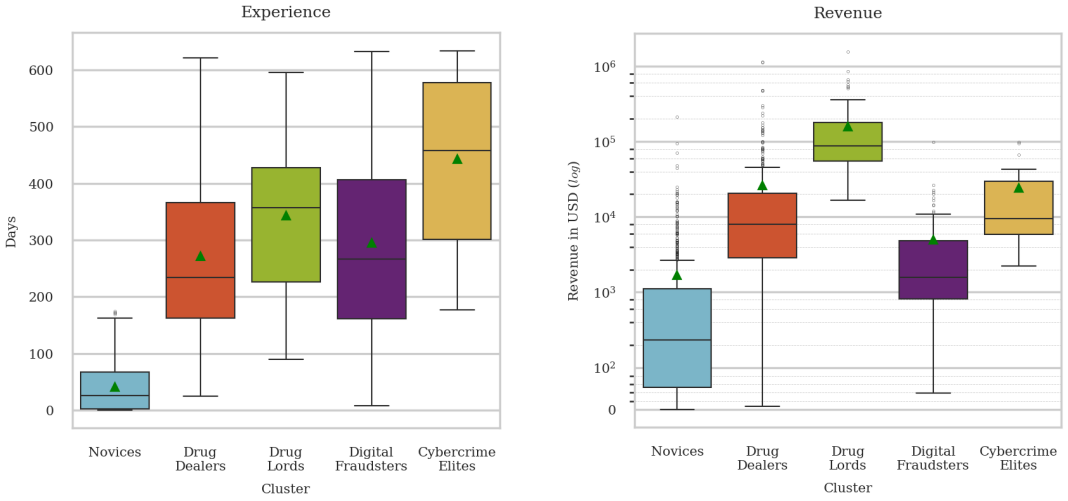


Figure 5.4: Amount of sales per sales type and cluster



(a) Experience

(b) Revenue

Figure 5.5: Distributions per cluster

`Drug Dealers` ($n = 509$) have more physical sales, experience and generate more revenue compared to the `Novices`. In terms of being active on multiple markets, 69.0% of the `Drug Dealers` have their reputation imported, which is higher than the `Novice` cluster. More than half of the vendors clustered as `Drug Dealers` has been active for 230 days, similarly half of the cluster has more than 80 physical sales. There are some vendors with digital sales included in this profile as well.

The last cluster with mainly physical sales are `Drug Lords` ($n = 110$), who do not really differentiate in terms of experience, but do have extremely high amounts of physical sales and generated revenue. Digital sales are mostly absent. Their business success might be caused by their presence on other markets. The average activity on other markets is high: 78.2% indicated to be active on other markets.

The next two clusters thrive in digital sales rather than physical sales. First, the `Digital Fraudsters` ($n = 103$) might have very few or very much days of experience. Theses vendors do have at least 15 sales in te digital domain. About 75% has more than 100 digital sales. Not much can be said on their preference to be active on other markets, since 58.3% showed to import their reputation. Some vendors with mainly digital sales, also made a number of physical sales.

Lastly, `Cybercrime Elites` is a very small group ($n = 23$). It is chosen to accept this cluster with few vendors because indeed, very successful vendors of digital items are scarce and they clearly trump the `Digital Fraudsters` in terms of sales. A large amount of `Cybercrime Elites` is active on other markets (73.9%).

5.3.2. Predominant sales categories within clusters

The clusters are partly based on the amount of sales per vendor of products that need to be physically shipped and the amount of sales per vendor that can be digitally transacted. Since vendors categorise their listing according to a predefined list of categories, it is possible to visualise the dominant sales categories per cluster. This is used to provide more context to the clusters and to check whether indeed the clustering resulted in sensible clusters of fairly homogeneous vendors. The categorisation made by vendors cannot be trusted blindly: vendors tend to post their ads in as many as possible categories, whether they apply to the product offered or not.

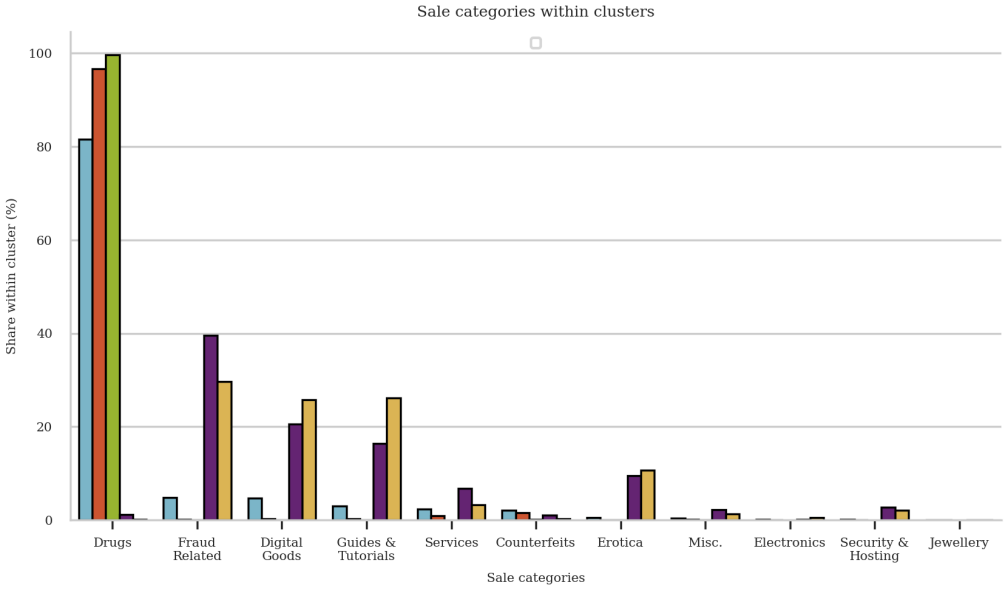


Figure 5.6: Dominant sales categories within clusters

The *Drugs* category is mainly populated by the three clusters in which physical sales are dominant. This is not a surprise, since selling drugs of course has a physical component. `Drug Lords` are selling 100% of their products in *Drugs*, and `Drugs Dealers` have 98% of their sales in this category. However, about 20% of the sales by `Novices` are not drug related. The homogeneity within this cluster is caused by the absence of experience and to a lesser extent, reputation import.

To be clear on what the next three categories entail, examples are given of offerings found in these categories. *Fraud related* items are for example banking credentials, credit card details, identification documents or hacked PayPal accounts. *Digital goods* are for example software keys, email and password combinations acquired in data breaches, accounts for streaming services such as Netflix and Spotify and login details for adult websites. *Tutorials Guides* are carding tutorials, cash-out tutorials, hacking guides, guides on how to use RATs and how to increase operational security while transacting illegal goods. These sales categories are dominant in the digitally focussed clusters `Digital Fraudsters` and `Cybercrime Elites`.

While these three categories are dominant in the digitally focussed clusters, sales generated in these categories are also originating from vendors clustered as the less experienced `Novices`. Because this cluster is fairly large, it involves a significant amount of vendors (with low amount of sales in these categories). The nature of the listings in non-drug related categories created by `Novices` are explored manually. From this, it is concluded that indeed digital cybercrime sales are made by some `Novices`. It therefore is important not to interpret *all* vendors in this cluster as 'amateur drug dealers'.

Lastly, *Services* are cashout services, hacking services but also include the sale of identification documents such as passports and listings in which malware and botnet services are offered for sale. In *Erotica* mostly credentials for adult websites are traded.

6

Security behaviour analysis

> "Sending coin from coinbase to AB was a mistake. Using the market to encrypt your address was a mistake [...] you are the low hanging fruit out of all of us here [...]"
>
> *A pseudonymous cybercriminal on Reddit*
>
> In response to the AlphaBay seizure, as documented by Bradley (2019, p.176)

The security behaviour of darknet market users is assessed by measuring their authentication behaviour (section 6.1), encryption of communication (section 6.2), the linkability of their pseudonyms (section 6.3) and their choice of online financial service providers (section 6.4). Each section starts with a descriptive analysis, continues with a statistical analysis of how the security behaviour compares between vendor types and concludes by presenting an interpretation and discussion of the results.

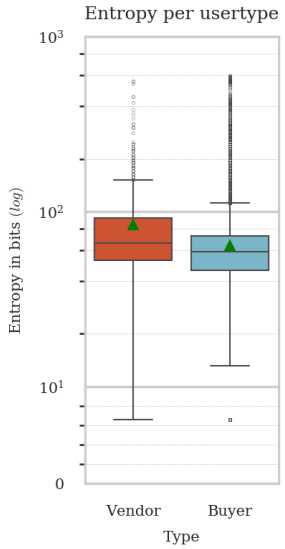
6.1. Authentication Security

Due to the Dutch Police gaining admin rights on the 20th of June 2017, Law Enforcement was in the position to alter the configuration of the market in such a way that passwords were saved as plain text in the markets' connection logs. In total, the passwords of 1081 vendors ($\approx$ 62.4% of all vendors) and 85620 regular members ($\approx$ 20.5% of all members). The strength of these passwords are analysed in section 6.1.1 and section 6.1.2 examines password uniqueness.

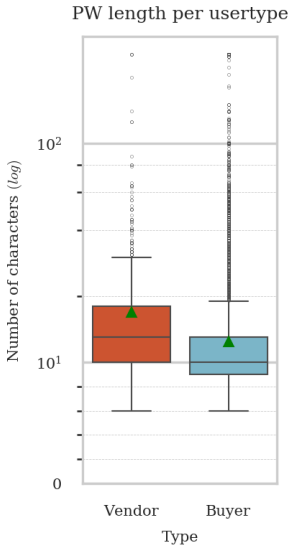
6.1.1. Password Strength

First, the passwords of vendors are compared with those of buyers (Figure 6.1). The distributions of password entropy and password length are displayed per user type in 6.1a and 6.1b respectively. For both password strength measures goes that vendors have higher means (μ) and medians (md) than buyers. Still, there are both vendors and buyers with very strong passwords (e.g. 200+ entropy or 80+ length). In the lowest 25% of each user group, the vendors' passwords have less than approximately 53 entropy bits of password complexity, compared to 46 bits for buyers. These are within the practical limits of brute forcing. At 350 billion guesses per second (Goodin, 2012), brute forcing such passwords is a matter of days and hours respectively[1]. Regarding

[1] This assumes offline cracking. Most current recommendations of the National Institute of Standards and Technology (NIST) (Grassi et al., 2017) do not specify minimum entropy requirements any more. An earlier version (Burr, Dodson, & Polk, 2004) recommends a minimum of 80-bits for the most secure applications.



(a) Entropy distribution per user type (vendor/buyer). $\mu = 83.2/63.6$, $md = 65.7/59.1$



(b) Password length distribution per user type (vendor/buyer). $\mu = 16.8/12.3$, $md = 13/10$

Figure 6.1: Authentication behaviour: password entropy and length per user type. To show the full distribution, when $y > 10$, the axis is log-scaled.

password length, the lowest 25% of the vendors have a password shorter than 10 characters, while 50% of the buyers has a password of less than 10 characters long.

The entropies of the passwords of the Hansa market users are further explored in Figure 6.2. In this figure, users that have a password with more than 200 entropy bits are grouped together for visualisation purposes. The left vertical axis shows the probability density per 10 bit entropy increase. The probability density is a measure of the relative likelihood that a certain 10-bit interval occurs. For each user type, the sum of the coloured area equals to one. Consequently, the bars visualise the differences in distributions between vendors an buyers. It is clearly visible that the lower ranges are more populated by buyers. When entropy increases, vendors become more prominently placed in the figure.

The dotted lines show the *cumulative* distribution on the right vertical axis. From this distribution, can be easily inferred that about 60% of the vendors and 80% of the buyers have a password complexity below 80 entropy bits. Similar trends can be observed when analysing a likewise figure for password length, which is attached in appendix B.1. Roughly 20% of the vendors uses a password longer than 22 characters, whereas about 5% of the buyers have passwords of this length.

To gain further insights in the authentication related security behaviour, descriptives are generated on how entropy correlates with the individual characteristics of vendors. Since entropy is based on password length, the latter is not further correlated with the vendor characteristics. Also, the passwords of buyers are not further explored.

As indicated in section 6.1, only from 1081 out of 1733 vendors the plaintext passwords could be retrieved. About 38% of the vendor population did not log into his or her account in the last month the market was active, thus their passwords remain unknown. This introduces a bias in the available password data (see appendix B.2). In the appendix is shown that vendors with low *experience* or *sales* are more likely to have no password data present compared to their more experienced or successful counterparts. Possibly, these vendors have not logged in to their accounts because they stopped trading or lost interest in the market. Vendors with less

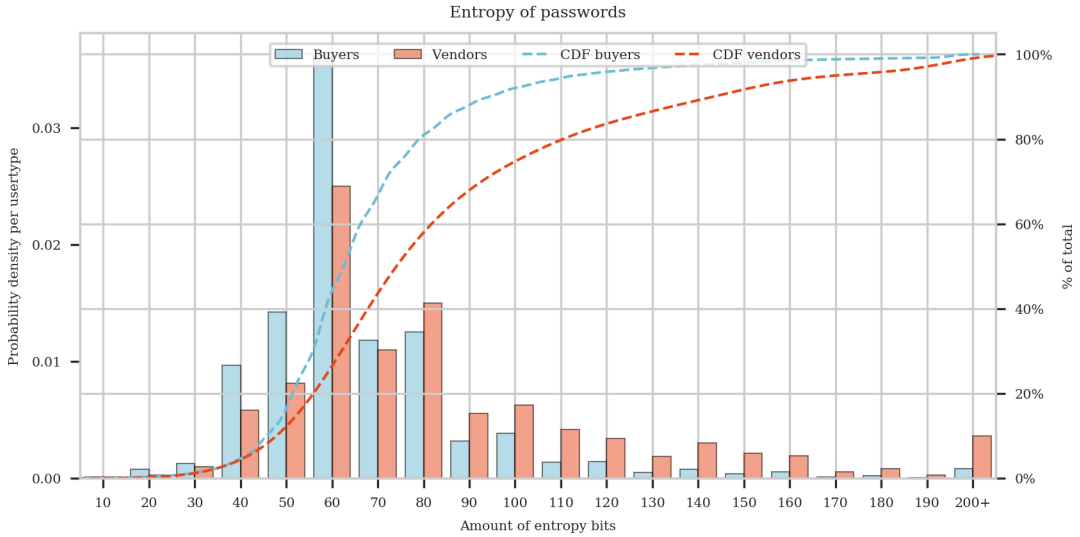
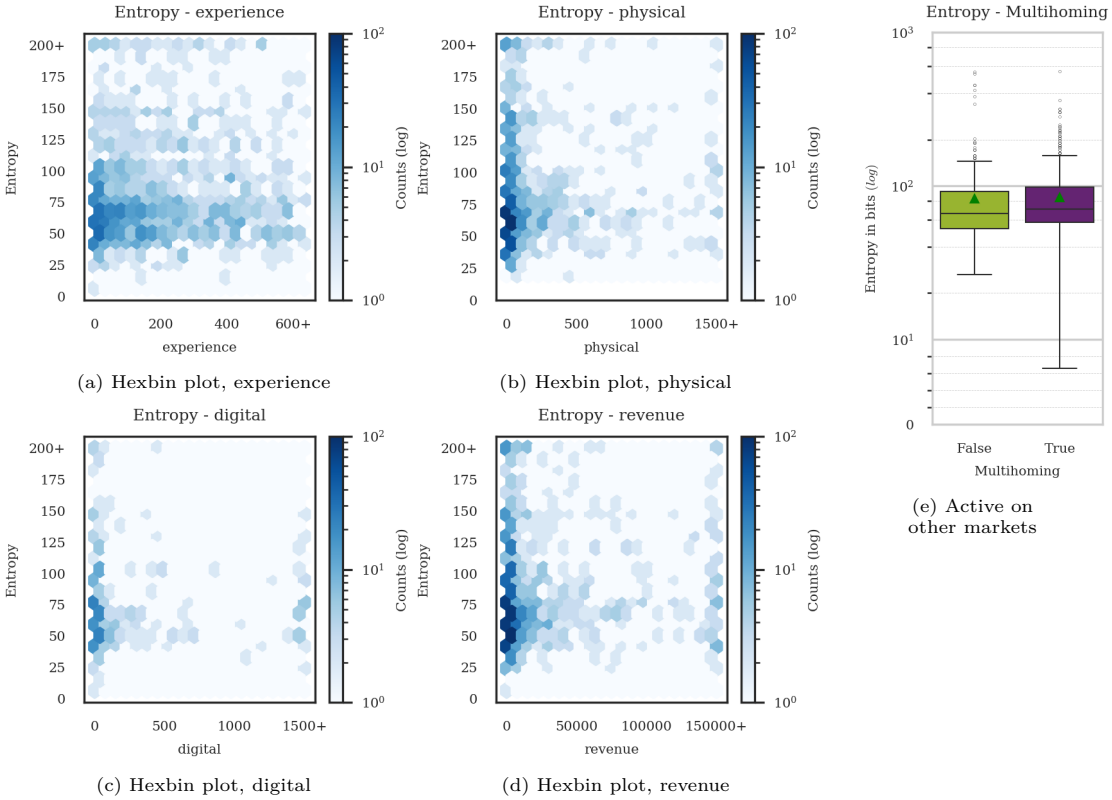


Figure 6.2: Probability density and cumulative distribution of password complexity per user type.

business success (sales) or experience are numerous. Thus, despite the missing data, all clusters of vendors remain well populated. In this part of the analysis, there are 493 `Novices` (-50.1%), 394 `Drug Dealers` (-22.6%), 93 `Drug Lords` (-15.5%), 78 `Digital Fraudsters` (-24.3%) and as much as `Cybercrime Elites` as before ($n = 23$).

Figure 6.3 shows the correlation between each vendor characteristic (horizontal axis) and entropy (vertical axis). Outliers are grouped together at the extreme values of the axes, as indicated by the '+'. For each plot in Figure 6.3, two conclusions regarding the relation between a vendors' characteristics and password complexity are valid. The first conclusion is drawn by examining the entropy ranges below the theoretical limit of brute-forcing (≈ 80 bits). Relatively poor passwords are used by both the low selling and high selling as the less experienced and experienced vendors. The second conclusion is based on the upper limits of the entropy scale, say $H > 125$. Many of the less experienced and low selling vendors have very complex passwords. On the other hand, also experienced vendors or vendors with many sales with complex passwords exist. These conclusions indicate that there are no *strong* correlations between individual vendor characteristics and password entropy.

One should keep in mind that Figure 6.3 shows absolute counts. Thus, vendors with many sales or experience - which are relatively few in number - are accentuated less in the figure. The same goes for the amount of digital sales (Figure 6.3c). This plot is heavily influenced by the fact that digital items are transacted by far less vendors.



(a) Hexbin plot, experience

(b) Hexbin plot, physical

(e) Active on other markets

(c) Hexbin plot, digital

(d) Hexbin plot, revenue

Figure 6.3: Entropy correlated with vendor characteristics

6.1.1.1. Statistical Analysis

Analysing the password complexities of vendors per vendor profile allows for considering multiple vendor characteristics jointly. First, the distribution of entropies within vendor types is visually inspected. Then, an ANOVA-test is performed to statistically determine whether the means of password complexity differ significantly between vendor profiles. The ANOVA-test is followed up with a Tukey-Kramer HSD test. This test clarifies *which* profiles have significantly higher or lower means compared to the other profiles.

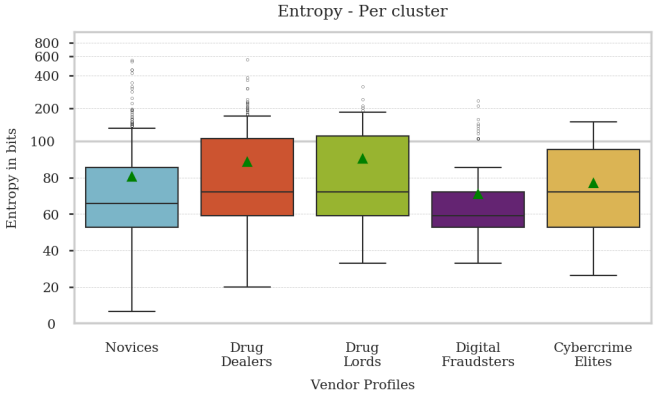


Figure 6.4: Distributions of password complexity per vendor type, log-scaled after 100 entropy bits.

Visually inspecting the distribution of entropies within the clusters of vendors (Figure 6.4) leads to four conclusions. Firstly, the minimum and average password complexity seems to scale with the amount of sales and experience among the physically oriented `Novices`, `Drug Dealers` and

`Drug Lords`. Secondly, the entropy of passwords in `Novices` is relatively dispersed, as indicated by stretched distribution. Thirdly, about 75% of the `Digital Fraudsters` have passwords below the theoretical limit of brute-forcing (80-bits), which is notably worse compared to the other vendor profiles. And fourth, considering the low sample size within `Cybercrime Elites`, the spread is very large. This makes it more difficult to find statistically significant differences between the entropy of `Cybercrime Elites` and the other clusters of vendors.

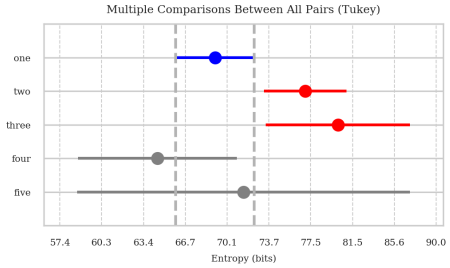
To statistically determine whether there is any difference between the clusters' means, a one-way Analysis of Variance (ANOVA) is performed. This is an *omnibus test*. Thus, it only shows whether at least two vendor profiles are significantly different in terms of password complexity. For an ANOVA-test to render valid results, the data must meet three assumptions. These are extensively discussed in appendix B. The assumption testing is summarised as follows. First, outliers that have a standardised value of $z > 5$ are removed. 488 `Novices` (-5), 391 `Drug Dealers` (-3), 93 `Drug Lords` (-0), 78 `Digital Fraudsters` (-0) and 23 `Cybercrime Elites` (-0) remain in the data. Second, the data is log-transformed to fix the slightly right-skewed distributions. As assessed by normalised histograms and QQ-plots, the data is approximately normally distributed within each vendor profile. Third, the homogeneity of variances assumption is met as assessed through Bartlett's Test ($\chi^2 = 5.63, p = 0.2285$).

The ANOVA-test is significant ($F(4, 1068) = 5.89, p = 0.0001$). This shows that the means of password complexity differ significantly among the vendor profiles. Sec, this result does not provide meaningful insights. An additional post-hoc test is performed to learn how the vendor profiles significantly differ from each other. A Tukey-Kramer HSD is recommended as post-hoc for unbalanced designs, in which the variances are assumed equal and the assumptions for the ANOVA-test have been met (Westfall et al., 2011).

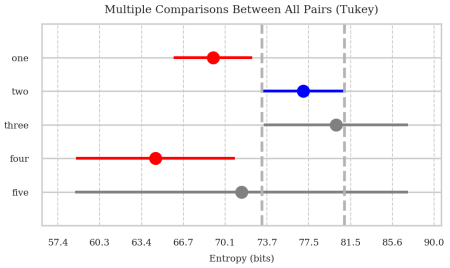
Table 6.1: Tukey-HSD post hoc results.

Profile 1	Profile 2	μ diff.	Adj. p	Lower CI	Upper CI
`Novices`	`Drug Dealers`	0.1079	0.0027*	0.0268	0.1890
`Novices`	`Drug Lords`	0.1470	0.0252*	0.0118	0.2823
`Drug Lords`	`Drug Dealers`	-0.0391	0.9000	-0.1770	0.0988
`Digital Fraudsters`	`Novices`	0.0687	0.6741	-0.0770	0.2145
`Digital Fraudsters`	`Drug Dealers`	0.1767	0.0102*	0.0285	0.3249
`Digital Fraudsters`	`Drug Lords`	0.2158	0.0118*	0.0323	0.3993
`Cybercrime Elites`	`Novices`	-0.0341	0.9000	-0.2891	0.2210
`Cybercrime Elites`	`Drug Dealers`	0.0739	0.9000	-0.1826	0.3303
`Cybercrime Elites`	`Drug Lords`	0.1130	0.7759	-0.1654	0.3913
`Cybercrime Elites`	`Digital Fraudsters`	-0.1028	0.8432	-0.3864	0.1808

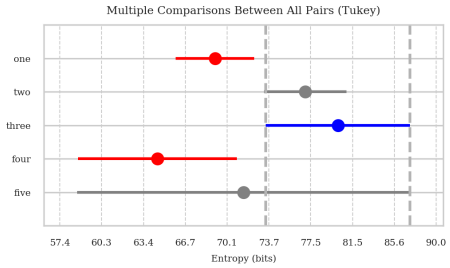
As Table 6.1 shows, only the differences between 4 pairs of vendor profiles are significantly different. No statistically significant differences were found for `Cybercrime Elites`. This may be due to a lack of power as a consequence of the small sample size. Because the pairwise comparisons are slightly cumbersome to analyse, the confidence intervals are be recalculated to account for all comparisons. By doing so, a single plot can be created in which the significance of mean differences of all vendor profiles is compared (Figure 6.5). In this figure the means and confidence intervals are plotted. Whenever the confidence intervals between vendor profiles do *not* overlap, the mean differences are statistically significant. Separate plots are displayed with each a different perspectives to aid interpretation. The vendor type in focus is marked



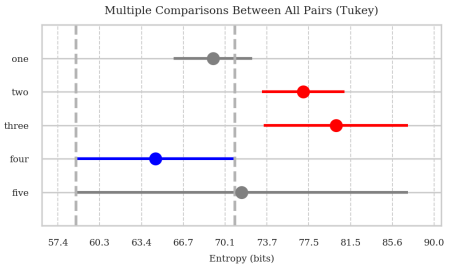
(a) Simultaneous Multiple Comparison between all pairs, `Novices`' perspective



(b) Simultaneous Multiple Comparison between all pairs, `Drug Dealers`' perspective



(c) Simultaneous Multiple Comparison between all pairs, `Drug Lords`' perspective



(d) Simultaneous Multiple Comparison between all pairs, `Digital Fraudsters`' perspective

Figure 6.5: Entropy, multiple comparisons tests

blue, which differs significantly from the vendor profiles marked with red. Additionally, the log transformation of the entropy values is reverted.

In conclusion, it is shown how `Drug Lords` > `Novices` & `Digital Fraudsters`. In which '>' denotes a statistically significant higher average of password complexity. Likewise, `Drug Dealers` > `Novices` & `Digital Fraudsters`.

Lastly, it is taken into consideration that simpler passwords might be chosen when vendors make use of two-factor authentication. Generally, this does not seem to be the case. 2FA-usage correlates *positively* with password complexity, as assessed by a Spearman rank-order correlation ($r_s = 0.219, p < 0.0000$). This shows that vendors do not tend to compensate relatively poor passwords with the additional layer of security that 2FA adds. Therefore, it would not make sense to include 2FA as a covariate in the password complexity analysis presented above. Rather, the positive correlation is an indication that the priority given to security truly differs between vendors. It is decided to treat 2FA-usage as a separate feature of observable security behaviour in section 6.1.3. In section 6.6, multiple security decisions of vendors are considered jointly.

6.1.2. Password uniqueness

Data breaches in which passwords are obtained are plentiful (Hunt, 2020). Databases with leaked passwords are used by password cracking software (e.g. Openwall (2020), creator of the Jack the Ripper software) to quickly guess passwords. As such, password reuse is very non-secure behaviour. A theoretically complex password can be easily guessed, when it is reused on a website which has poor security practices (Ives et al., 2004). For cybercriminals, breaches of account are not the only security risks introduced by password reuse. In databases of leaked credentials, Law Enforcement can find email addresses and usernames of passwords that have been matched. If the matched password is fairly unique, a criminal is de-anonymised quickly.

The 'Have I Been Pwned' database of Troy Hunt includes more than 10 billion leaked passwords, of which 573 million are unique. The passwords of 26540 Hansa Market users could be matched

with the passwords in the PWND database. This amounts to 30.6% of the users whose plaintext password was available. Regarding vendor accounts, 185 passwords are matched (17.1%).

A password match does not necessarily implies that the vendor is reusing his or her password on other websites. Since humans' password generation behaviour follows certain trends (see section 4.2.2.1), it could be very well the case that another person uses the same password. Using such a 'common password' is also poor security behaviour and thus it is also of interest. To assess how common a password is, the amount of occurrences of the password in the PWND data is regarded. Doing so, it becomes clear that 60 vendors have passwords that matched between 1 and 9 times, 54 vendors have passwords that occurred 10-99 times in the PWND data and 71 vendors used passwords that matched 100+ times. The remaining vendors ($n = 896$) could not be matched.

Appendix B.4 shows that most passwords that are not matched tend to be more complex than reused passwords. Exceptions - that showcase the predictability of humans perfectly - exist. A vendor used the password `1q2w3e4r5t6y7u8i9o0p`. By Equation 4.2, this 20 characters long password results in 131 entropy bits. This is fairly high. However the true *true* entropy ('uncertainty') of the password is very low. According to the PWND data, 29,426 other accounts use this password. In practice, this account would be breached within a day. Thus, despite the high estimated theoretical entropy, the entropy in practice is very low due to the predictability of the password.

The amount and type of password matches are visualised in Figure 6.6. Again, `Digital Fraudsters` seem to have the worst password practices. Of this group, 26.9% of the passwords are matched. 9% of these reused passwords are fairly unique and thus might introduce the risk of de-anonymisation. Three out of 23 `Cybercrime Elites` use a very common password that has been matched more than a hundred times.

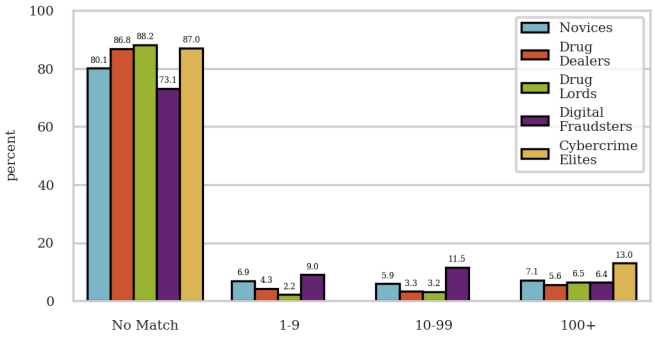


Figure 6.6: Distributions of password uniqueness per vendor type

Considering that (a) any type of password match is a display of non-secure behaviour and (b) the distinguished number of password matches (e.g. '1-9 matches' and '100+ matches') are categorised rather arbitrarily, it is decided that the distinguished types of password matches are not fit for making firm conclusions based on statistical testing. Therefore, a χ^2-test of the proportion password matches/no matches is performed. This proportion differs significantly between vendor profiles ($p = 0.0064$). The χ^2-test is valid, one expected cell count is less than five (3.94). This means that 90% of the cells have expected counts greater than 5.

From a pairwise post-hoc z-test of proportions with FDR-BH correction (see Table 6.2) the following is inferred. First, for most pairwise comparisons no significant difference in the proportion of non-matches/matches is found. This resembles Figure 6.6, in which the percentages of non-matches are alike. Secondly, `Drug Dealers` > `Novices` and `Drug Dealers` >

`Digital Fraudsters`. In which '>' indicates a statistically significant higher proportion of non-matches (i.e. more secure behaviour). Likewise, `Drug Lords` > `Digital Fraudsters`.

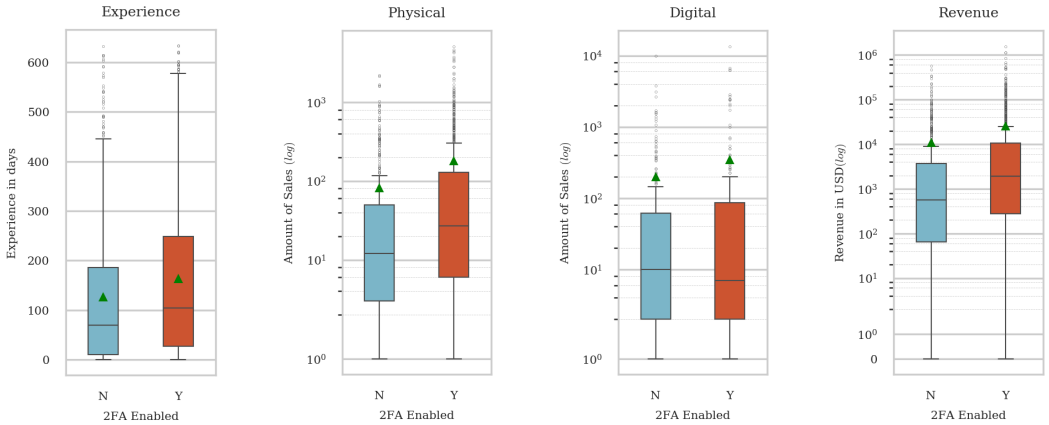
Table 6.2: Results of z-test of proportions, with FDR-BH adjusted p-values

	`Novices`		`Drug Dealers`		`Drug Lords`		`Digital Fraudsters`	
	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.
`Novices`	.	.						
`Drug Dealers`	-2.637	0.0390*	.	.				
`Drug Lords`	-1.830	0.1681	-0.354	0.9041	.	.		
`Digital Frd.`	1.423	0.2822	3.063	0.0219*	2.521	0.0390*	.	.
`Cybercrime Elt.`	-0.808	0.5991	-0.021	0.9830	0.160	0.9697	-1.374	0.2822

* significant with $\alpha = 0.05$

6.1.3. Two-Factor Authentication

For every vendor in the Hansa Market data, it is observed whether Two-Factor Authentication (2FA) is enabled or disabled. From the vendor population 60.5% ($n = 1049$) enabled 2FA. Figure 6.7 shows that more experienced and successful vendors are inclined to have two factor authentication enabled. Regarding vendors that specialise in digital items, this observation does not hold (6.7c). It is also clear that numerous experienced or high selling vendors did *not* enable 2FA. Regarding the vendors that are active on other markets, out of 911 vendors 655 (71.9%) enabled the extra authentication security measure. Likewise, for the remaining 822 vendors that are not active on other markets 394 (47.9%) enabled 2FA.



(a) Experience (b) Physical sales (c) Digital sales (d) Revenue in USD

Figure 6.7: Distributions of vendor characteristics, per group of vendors with 2FA disabled (N) or enabled (Y).

When considering the vendor characteristics jointly by means of the vendor profiles, statistically significant differences in proportions are found using a χ^2 Test of homogeneity ($p < 0.0000$). The proportions of vendors that use 2FA are presented in Table 6.3. The χ^2-test is followed up by a z-test of proportions post-hoc test. The significance levels and the direction of the z-statistic in Table 6.4 indicate whether a *column*-wise vendor profile has significantly higher or lower proportion 2FA-usage compared to a *row*-wise vendor profile. In summary, it is safe to conclude that `Drug Lords` > `Drug Dealers` > `Novices` > `Digital Fraudsters` and `Drug Lords` > `Cybercrime Elites`. In which '>' indicates a statistically significant higher proportion of 2FA usage.

Table 6.3: Amount of vendors with 2FA enabled, per vendor type

2FA	Novices	Drug Dealers	Drug Lords	Digital Fraudsters	Cybercrime Elites
N	446	150	20	58	10
Y	542	359	90	45	13

Table 6.4: Results of z-test of proportions, with FDR-BH adjusted p-values

	Novices		Drug Dealers		Drug Lords		Digital Fraudsters	
	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.
Novices	.	.						
Drug Dealers	-5.868	0.0000*	.	.				
Drug Lords	-5.427	0.0000*	-2.405	0.0269*	.	.		
Digital Frd.	2.164	0.0436*	5.245	0.0000*	5.772	0.0000*	.	.
Cybercrime Elt.	-0.158	0.8741	1.433	0.1898	2.640	0.0166*	-1.116	0.2936

* significant with $\alpha = 0.05$

6.1.4. Interim conclusion & discussion: authentication security

From the perspective of a cybercriminal, authentication related security mechanisms reduce the risk of unauthorised access to or control over cybercriminals' online assets, such as online communication, information (e.g. business partners), transactions or even funds. In case the cybercriminal is active on multiple markets and does not use different passwords for these accounts, information assets distributed over multiple accounts may be accessed. A threat to these assets are LEA that are authorised to perform automated password guessing (Figure 6.8). Next to securing assets from LEA, cybercriminals also protect themselves against from intrusions by other cybercriminals. Skilled cybercriminals try to cheat or scam the less skilled (Herley & Florêncio, 2010) or aim to harm the business of a competitor (Van de Sandt, 2019, p.116). This falls beyond the scope of this thesis (*cf.* the definition of 'security' in section 3.1).

The security mechanism of having a complex and unique password drastically decreases the likelihood of successful password guessing. Passwords can be guessed remotely, albeit heavily limited by connection speeds and server response times. Additionally, security mechanisms on platform level, e.g. CAPTCHA's or the locking of accounts after x failed login attempts during a certain time period, may prevent prolonged password guessing. According to Wang, Zhang, Wang, Yan, and Huang (2016), online password guessing is feasible by feeding password guessing algorithms additional information, such as the website, databases of leaked passwords and personal information (interests, birthday, phone number, hobbies, age, *etc.*). The authors show that *within 100 guesses*, they were able to guess 32% of the passwords of 'security-savvy' users. Of course, not much personal information is known in early stages of investigations and cracking randomly generated passwords does not benefit from targeted guessing. Offline password guessing is much faster, yet can only be performed when web servers or devices are seized.

Note how the deployment of a security mechanism (a password) gives rise to other security risks (Van de Sandt, 2019, pp.90-91). In section 3.1.2.2 it is explained why people tend to reuse complex passwords specifically. People may have a tendency to reuse memorised complex passwords on different websites (Wash et al., 2016). This leads to a false sense of security, since the newly created security risk (i.e. password matching) is more severe than the risk of successful automated guessing when using a simple but unique password (Figure 6.8).

Using two-factor authentication renders online password guessing useless for LEA, when they are not in the possession of the private PGP-key. Enabling 2FA also reduces another security risk. LEA are allowed to engage in social engineering attacks such as (spear)phishing (Ministry of Justice & Security, 2015). In these phishing efforts, fake look-a-like websites are created through which the target sends his or her login credentials to LEA. Bradley (2019, p.176) shows that after the closure of AlphaBay, cybercriminals tried to scam each other through such phishing attacks.

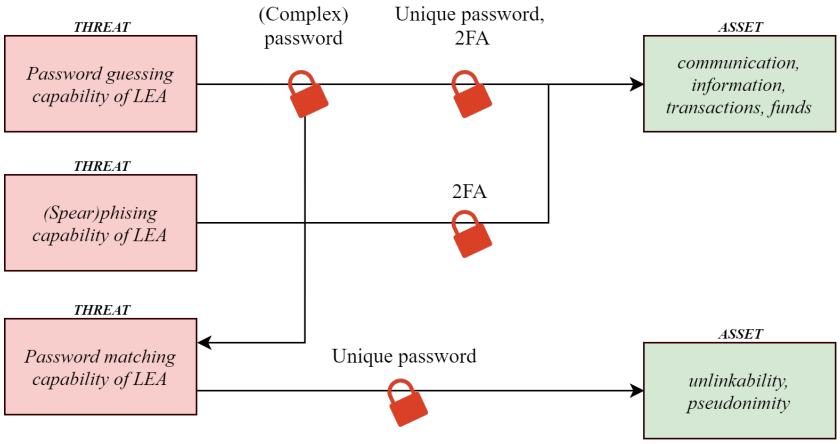


Figure 6.8: Authentication security mechanisms reduce risks to assets created by threats

The security mechanisms analysed in this section do not require large investments in terms of time, knowledge or money. Expectations were that suboptimal security behaviour would barely be present. The analysis of this behaviour shows that this is not the case. Stronger still, not only 'random' mistakes by individuals are found, it is demonstrated that significant differences between vendor types exist. `Drug Lords` and `Drug Dealers` have the most secure authentication security practices. `Digital Fraudsters` are generally the worst performing and `Cybercrime Elites` performance is quite diffused. It is also observed that some high selling or experienced vendors actually do have very poor authentication security practices.

This signals that security mistakes in other domains may be present as well. For example, poor or absent security mechanisms that ought to protect physical or non-IT objects. These are often overlooked by academics discussing cybercrime but highly relevant to Law Enforcemt (Van de Sandt, 2019, p.91). The rationale would be that a vendor that is not bothered to get these simple authentication practices right, does also not prioritise a highest level of security when encrypting data carriers (e.g. hard disks or usb sticks), meeting with business partners, spending large amounts of (tainted) money *etc.*. In this sense, authentication practices are a *proxy* for prioritisation or attention given to security.

6.2. Encryption of Communication

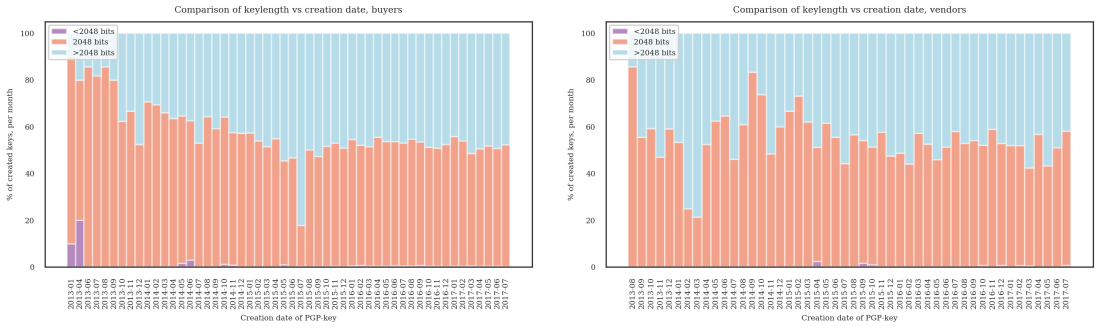
The PGP-adoption among vendors is high. Only 5 vendors do not have a PGP-key listed: two `Novices` and three `Drug Dealers`. It could be, that they removed their PGP keys from their accounts after they stopped trading. The adoption of PGP-keys among buyers is noticeably lower. Only 50,657 out of 415,703 buyers (12.19%) registered a PGP-key.

The public PGP-keys of vendors contain information about the encryption algorithm used, the date the key was generated, the email address the key is registered to and the chosen length of the key. This information is extracted using a Python implementation of GnuPG.

Among the Hansa users, the most-used encryption algorithm within the PGP-protocol is RSA ($\approx$ 98%). Appendix C.1 explains that no security-related conclusions can be drawn from the chosen algorithm. Furthermore, in the appendix is discussed that clearweb email addresses included in te PGP-keys are likely to be fake. Some vendors include dark web addresses or secure email services in their PGP-key. It is less likely that vendors put effort in faking an email address which is hosted at a secure service. The most listed secure email service that is still in service at the time of writing is *safe-mail* (Appendix C.3).

Lastly, the key sizes of the primary[2] RSA PGP-keys are analysed. The key size indicates how 'secure' the key is. Weak keys ($\leq$ 1024 bits) are observed for 9 vendors and 88 buyers. Even by 2015's standards, these key lengths are considered not to be sufficient (Lenstra, 2004; Lenstra & Verheul, 2001). Currently, NIST recommends key sizes of asymmetric cryptosystems based on factorisation problems, such as RSA, to be at least 2048-bits (Barker et al., 2020). Larger keys (2560+ bits (Lenstra & Verheul, 2001) or 3072+ bits (Barker et al., 2020)) are recommended when information needs to remain hidden after the year 2030.

To analyse whether the differences in key strength should be attributed to the creation data of the key, instead of vendors' preferences, the key strengths are plotted over creation dates are per month (Figure 6.9). For buyers, there seems to be a trend in which the most secure keys are created more often when time progresses. Regarding vendors, such a trend is absent.



(a) Buyers, months with <10 keys created omitted.

(b) Vendors, months with <10 keys created omitted.

Figure 6.9: Proportion of key sizes of PGP-keys per date of creation

Considering that the security benefit of any key stronger than 2048-bits is negligible, it is expected that key sizes are chosen 'randomly' or according to whatever is recommended in one of the many PGP-tutorials found on dark web discussion fora. However, extremely secure keys are more often found among vendor profiles that proved to be more 'security aware' in the other analyses as well 6.10. No weak keys of `Drug Lords` are found and among the digitally focused clusters, very strong PGP-keys are observed notably less. In conclusion, the figure suggests that the key strengths are not completely randomly selected by the Hansa vendors.

To statistically determine which vendor types show higher proportions of extremely secure PGP-keys, a χ^2-test with FDR-BH adjusted post hoc z-test is performed on the proportion of >2048-bit keys within the clusters. Thus, in this test all key sizes $\leq$2048-bits are grouped together. The clusters differ significantly, $p < 0.0000$ with all cells having an expected count of more than 5.

From the post-hoc test (Table 6.5) the following conclusions are valid: `Drug Lords` > `Digital Fraudsters` & `Cybercrime Elites` and `Drug Dealers` > `Novices` > `Digital Fraudsters` & `Cybercrime Elites`. In which '>' indicates a statistically significant higher proportion of extremely secure PGP-keys.

[2]Only 0.035% of the users has more than two keys (i.e. within the same PGP-key, a primary key and more than one subkey). Of all [primary key, first subkey]-pairs, only 16 (0.00035%) are not created at the same time. It is concluded that only analysing the primary keys suffices.

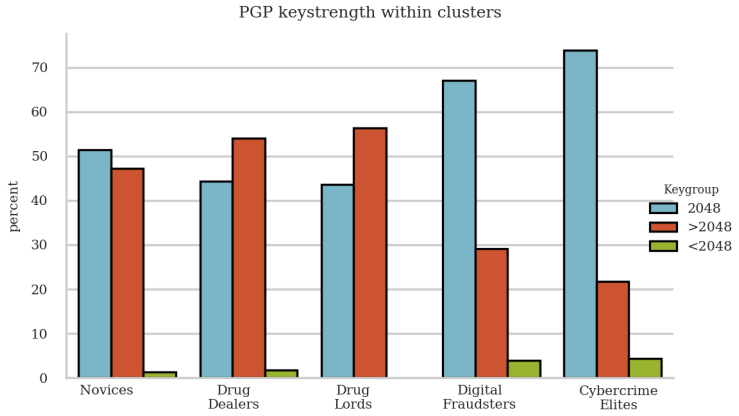


Figure 6.10: Vendors linkable via PGP-matching

Table 6.5: Results of z-test of proportions, PGP-key strength >2048-bits, with FDR-BH adjusted p-values

	Novices		Drug Dealers		Drug Lords		Digital Fraudsters	
	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.
Novices	.	.						
Drug Dealers	-2.447	0.0218*	.	.				
Drug Lords	-1.812	0.0875	-0.460	0.6454	.	.		
Digital Frd.	3.517	0.0015*	4.593	0.0000*	4.010	0.0003*	.	.
Cybercrime Elt.	2.425	0.0218*	3.026	0.0050*	3.020	0.0050*	0.715	0.5273

6.2.1. Interim conclusion & discussion: encryption of communication

If a darknet market gets raided by LEA, unencrypted communication becomes important evidence to the investigators. Before Hansa Market was taken down, other markets have been raided by LEA as well. For example, in *Operation Onymous* (November 2014) 9 darknet markets and a number of other hidden services have been taken down. Supposedly, law enforcement exploited weaknesses in the TOR-routing protocol (Hern, 2014). PGP-encryption of communication compromises the usability of evidence obtained from such raids (Figure 6.11). Soska and Christin (2015) show that in 2012, about 60% of the vendors on darknet markets have PGP-keys listed. After Operation Onymous, the adoption of PGP increased from 80% to 90%. The research presented here shows that the PGP-adoption on Hansa Market is nearly 100% in July 2017. The complete adoption of PGP-encryption indicates that cybercriminals become aware (or, are made aware) of the fact that security mechanisms on platform level not always suffice. As such, cybercriminals continuously react to law enforcements' capabilities and vice-versa (Van de Sandt, 2019, p.224).

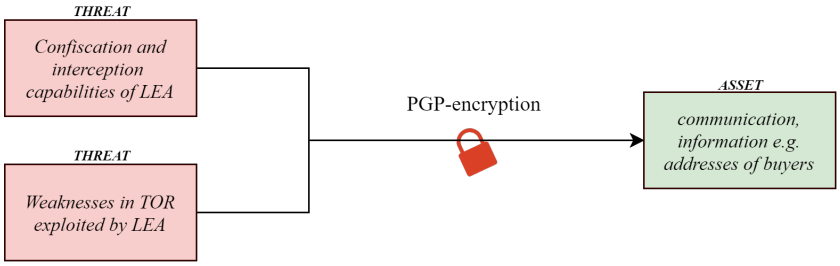


Figure 6.11: Encryption of communication mechanism reduces risks to assets created by threats

It was expected not to recognise any 'logic' in the PGP key sizes used by vendors. In 2015-2017, the security benefits of 4096-bits keys over 2048-bits keys will be negligible for at least 30 years. Assumed was that a random pattern would appear because vendors either a) follow one of the many PGP-tutorials available online, b) select the default value (2048-bits, for most up-to-date key generators) or c) select the highest number possible (often 4096-bits). However, the results presented in this section show that vendors selling larger amounts of drugs (i.e. `Drug Lords` and `Drug Dealers`), use extremely secure keys significantly more often than other types of vendors. Especially compared to those focussing on transacting digital items only. A possible explanation for this distinction is found in Figure 6.11. Addresses of buyers are valuable assets to vendors that specialise in shipping drugs. This is because of two reasons. Firstly, a buyer that receives a visit from law enforcement may slander a vendor for being untrustworthy. Secondly, shipping addresses enable LEA to track packages in the systems of postal services. This reveals the address from which a package is sent. This gives LEA clues about the vendors' identity or whereabouts. Vendors that only transact digital items, do not have physically ship items to addresses. Thus, they do not have names and addresses of their business partners. The information assets in their possession create less risks to de-anonymisation and thus, these vendors might feel that their assets are not in need of 'extreme protection' through 2048+ bit keys.

Van de Sandt (2019) repeatedly mentions that the security behaviour of cybercriminals involves objective and subjective risk assessments. Potentially, this subjectivity becomes apparent through the observation that - despite negligible security benefits - a significantly larger share of `Drug Lords` use extremely secure keys, compared to other vendor types. The 2048-bit keys are estimated to become less secure after the year 2030. This means that Hansa vendors with 2048-bit keys (unknowingly) take the risk that their assets might be accessed after approximately 15 years. It is hard to objectively assess to what extend this risk is considerable.

6.3. Linkability of Pseudonyms

Reusing a PGP-key when active on multiple markets can be beneficial to business success (Van Wegberg & Verburgh, 2018). However, it does introduce potential security risks. LEA are able to link darknet market pseudonyms through this PGP-key. This section shows which vendors chose to reuse their PGP-key over different markets. In the analysis presented here, focusses on vendors that are known to be active on other markets, i.e. those that imported their reputation from at least one other market. Their PGP-keys are matched against the database of the Grams search engine. When the PGP-key of a vendor active on another market is not found in the Grams data, it is likely that this vendor uses different PGP-keys for his or her accounts. This increases security. Figure 6.12 shows how the following groups overlap: vendors with PGP key ($n = 1728$), vendors known to be active on other markets ($n = 908$) and vendors who are linked with Hansa and any other market ($n = 902$). From this figure is concluded there is a group ($n = 265$) who is known to be active on another market but whose PGP-keys could not be matched in the Grams data. Surprisingly, there is also a group ($n = 259$) who did not use the import functionality but whose PGP-keys *were* matched in the Grams data.

When the number of PGP-matches is visualised per cluster of vendors, in which only the vendors are included that are known to be active on multiple markets, it is clear that the differences between clusters are small (Figure 6.13). In the figure, the blue (left) areas are of interest. These represent the vendors that could not be matched. Thus, these vendors might use different PGP-keys on the markets they are active on. A χ^2-test confirms that of the proportion of vendors per cluster that could not be matched (the 'blue' or 'left' areas in Figure 6.13) does not significantly differ between vendor profiles ($p = 0.8425$). The test is valid, all but one (4.96) cells have an expected count of greater than 5.

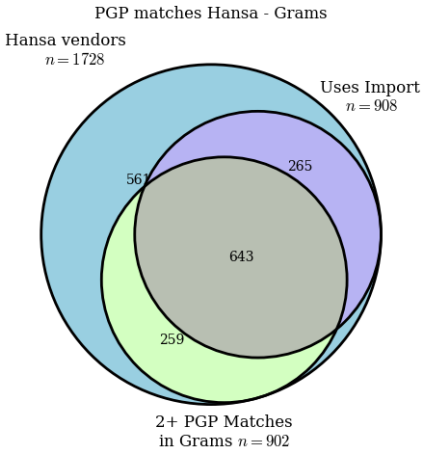
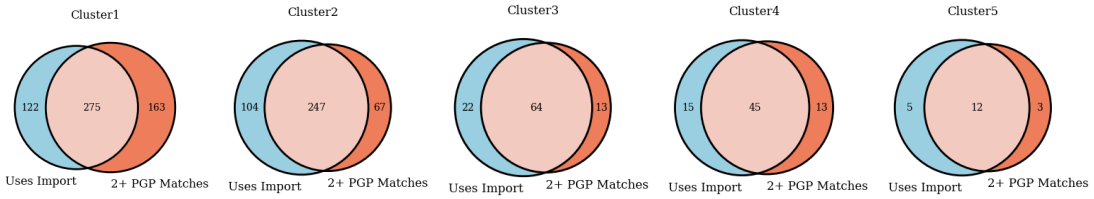


Figure 6.12: Percentage of grouped key sizes per cluster



(a) `Novices` (b) `Drug Dealers` (c) `Drug Lords` (d) `Digital Frd.` (e) `Cybercrime Elt.`

Figure 6.13: The number of PGP-matches of vendors known to be active on other markets, per vendor type.

6.3.1. Interim conclusion & discussion: linkability

PGP-keys have an important function of signalling trust to buyers. While pseudonyms are easily imitated, PGP-keys allow buyers to verify that a vendor account on market x is run by the same entity as the vendor account on market y. PGP-keys therefore are tied to a reputation and allow vendors to monetise this reputation over different markets. If market x unexpectedly shuts down, loyal buyers can find their favourite vendors on market y through PGP-key verification. The search engine *Grams*, that was available to every darknet user, made searching vendors by PGP-key an easy job. However, while reusing PGP-keys over different markets may enhance business success, it potentially reduces security by allowing LEA to link different accounts. A loyal user base, trust and reputation built during a cybercriminals' career may vanish when an evasive measure such as using different PGP-keys is used (Figure 6.14).

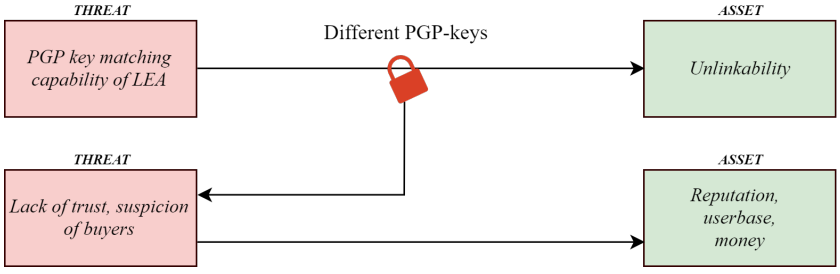


Figure 6.14: Using different PGP-keys reduces risks to assets created by threats and creates a risk to another asset.

The results indicate that most vendors that are active on multiple markets use the same PGP-key. There is also a group that supposedly use different PGP-keys. However, grams data might not be complete or up to date. Of all Hansa PGP-keys belonging to vendors with 5+ sales, 60%

are found in the Grams data *and* are attributed to Hansa market vendors. The Grams data has a small bias towards picking up larger vendors.

6.4. Choice of Online Financial Service Providers

The use of intermediaries to facilitate the conversion of cryptocurrency earnings to spendable fiat currencies is inevitable. At the same time it introduces potential security risks. In the Hansa Market dataset 19.238 unique bitcoin payout addresses could be recovered. Unfortunately, this analysis deals with a large amount of missing data.As elaborated upon in section 4.2.2.4, Hansa generated single use multi-signature addresses for each order. These are not the addresses analysed in this section. It was verified that the payout addresses considered in this section are those on which the earnings are deposited after the order had finalised. These addresses are vendors' own.

Half of the vendors used 3 or more unique bitcoin addresses and 75% of the vendors created more than 7 bitcoin addresses for their payouts. However, due to the missing data, statistics on the amount of bitcoin addresses used by each vendor must be interpreted with care. These numbers are likely to be higher in practice. In appendix D the following observations are made. First, some vendors use multiple payout addresses for a single listing. Second, no payout addresses are found that correspond to more than one listing (with at least one sale). From this follows that vendors with n listings have at least n payout addresses. This observation indicates that Hansa required its vendors to create a new bitcoin address for each listing. It also renders the amount of bitcoin addresses not a valid security behaviour. It would be very much related with the number of listings a vendor has.

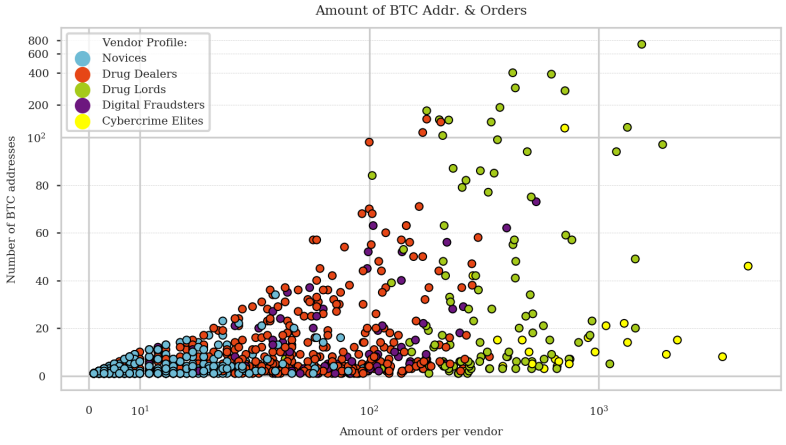


Figure 6.15: Number of bitcoin addresses per vendor compared with the number of orders of a vendor

In Figure 6.15 the amount of orders is plotted over the number of bitcoin addresses per vendor. Here, the amount of orders are corrected for the missing data: only orders in the time frames where bitcoin addresses are retrieved are considered.

6.5. Types of payout addresses

This section discusses the contextual information that is queried via the API of Chainalysis. This section is purely descriptive and shows what type of services are being used by darknet market vendors. In section 5.1 was explained that the last month of data is discarded because the huge influx of vendors made the vendor characteristics unreliable. Because this section does not include the analysis of relations between vendor characteristics and the choice of OFSPs,

this last month *is* included. By doing so, much more data becomes available. This allows for a better insight in the dynamics of preferred OFSPs. It will be clearly indicated when vendor characteristics *are* considered and the more limited data set is used.

From 19.238 bitcoin payout addresses, 2680 ($\approx 14\%$) could be directly attributed to clusters that are identified with known service wallets, such as centrally organised exchanges, peer-to-peer exchanges and bitcoin mixers. In this thesis, payouts directly transacted to such known services are referred to as *direct links*. As expected, the majority of the bitcoin addresses cannot be directly linked with service wallets. This may be due to the fact that a) some service wallets are not identified by Chainalysis and b) vendors do *not* send their criminal proceeds directly to a service wallet. Instead, they first accumulate their earnings on a privately owned (hardware) wallet.

The challenge is to make sense of the remaining 86% that is not directly linked to known services. According to cybercrime investigators, it is very likely that a high amount of vendors use private hardware wallets for storing their bitcoins (scenario b). Therefore, the heuristic presented in section 4.2.2.4 is applied to separate the vendors with hardware wallets (scenario b) from the Chainalysis shortcomings (scenario a). The heuristic results in the 16,564 payout addresses that are not directly linked to a service being categorised as follows. Assumed private wallets ($n = 4165$), private wallets with no exposure ($n = 4037$) and wallets that are either private wallets or service wallets ($n = 8344$). The analysis of 12 addresses threw an error, partly because of parsing issues or due to the address not being recognised by Chainalysis. These numbers are summarised in Figure 6.16.

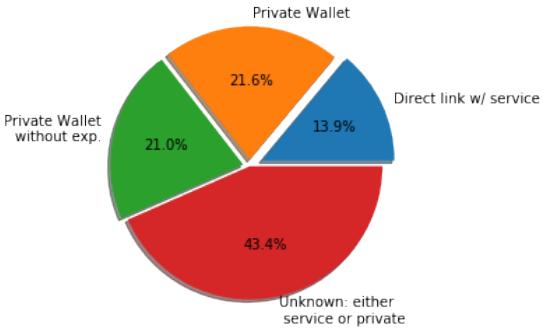


Figure 6.16: Types of wallets identified

6.5.1. Descriptive Analysis

6.5.1.1. Services transacted to

The analysis of the bitcoin addresses reveals what services vendors transact to, either directly or via their private wallets (Figure 6.17). First the output to darknet markets is discussed, followed up by the exchange-like entities and lastly mixing and gambling is elaborated upon.

Darknet Markets Figure 6.17 shows a large peak at vendors transacting to darknet markets. Three reasons explain this large number. First, it is possible that vendors buy goods from darknet markets. Secondly, because vendors are required to pay vendor bonds upon registering on darknet markets, outgoing exposure to markets is created. Thirdly, because a single private wallet can have exposure to multiple markets, the count may rise quickly (every unique market is counted separately).

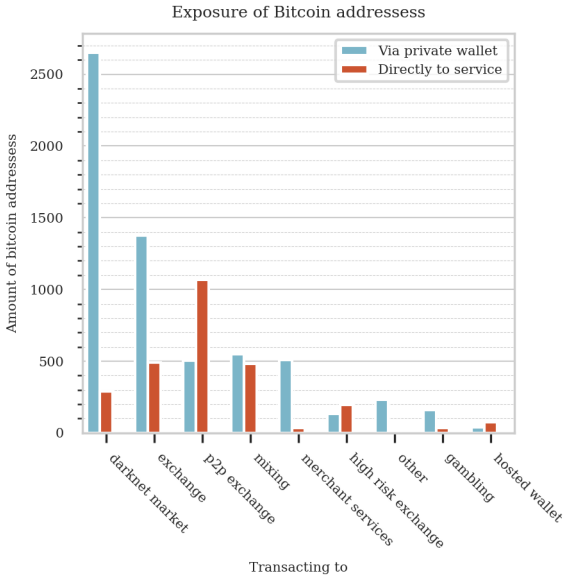


Figure 6.17: Exposure of addresses found in Hansa Market

The red bar indicates a direct link to darknet markets. This is an unexpected result. Further examination learns that 63 vendors use their AlphaBay wallets[3] for their Hansa Market earnings. Of these vendors, 23 did not make use of the Hansa Market reputation import functionality. Possibly, they did so to prevent them from becoming linkable over different markets. Of these 23 vendors, 16 could not be matched in the Grams search engine via PGP-matching. Likewise, with the same criteria (no reputation import, no PGP-match in Grams), 2 Dream Market vendors and 6 vendors active on some lesser known markets were found.

While it is very little effort to create and manage a new wallet, vendors decide to reuse their existing wallets. Even vendors that presumably took action by changing their PGP-key, do create a potential security risk by having a link between their darknet market accounts via reusing wallets.

Exchanges and Hosted Wallets When profits are easily tracked to an exchange or hosted wallet, a vendor exhibits potentially non-secure behaviour. This is due to the fact that these wallets are managed by central entities that can be subpoenaed for information on its users. Hosted wallets carry similar security risks as exchanges. As Figure 6.17 shows, these are not frequently transacted to by vendors and thus not further elaborated upon.

To what extend an exchange responds adequately to a subpoena, depends on the type of exchange and in which jurisdiction it is located. Exchanges that are reluctant in gathering data on its users, or those that do not perform identity checks, are not likely to respond with relevant information such as names, email addresses, bank accounts and other cryptocurrency addresses. These exchanges have poor KYC and AML controls and are labelled as 'high risk exchanges'. Next to these high risk exchanges, peer-to-peer exchanges are also regarded as safe-havens due to minimal identity verification[4]. The 'regular exchanges' have better track records with regard to KYC and AML controls. It is however not guaranteed that subpoenaing a regular exchange

[3]When a user registers on AlphaBay, he or she is assigned an AlphaBay wallet. This is not the case for Hansa Market.

[4]The data shows that LocalBitcoins.com is by far the most used p2p exchange. In the 2015-2017 time frame, LocalBitcoin did not verify identities. At the time of writing, steps have been taken to adhere to AML regulations (https://localbitcoins.com/blog/aml-features-update/)

always results in successfully identifying the vendor. An example is when merely a money mule is identified, rather than the vendor.

Nonetheless, having an observable link between the bitcoin address used in criminal transactions and the exchange at which the earnings are converted to fiat currency, may give usable leads to Law Enforcement. Thus, it is regarded as an exhibit of non-secure behaviour. Especially when the profits are directly transacted to an exchange, investigators are presented a solid reason to subpoena the exchange. Later on, if the bitcoin address can be tied to an identity and if the person is located in a country that is willing to cooperate with the investigating party, such a direct link can be used as undisputable evidence in court.

As indicated by the relatively large red bar, P2P-exchanges are often directly transferred to. This indicates that vendors are confident in using P2P exchanges. Firstly, because they allow a direct link between the illegal earnings and the exchange. And secondly, because they entrust their funds to be stored at the exchange, and not for example on a hardware wallet. About 1350 private wallet addresses show an output to a regular exchange and approximately 500 addresses are directly linked with such exchange.

Mixing and Gambling Transacting to mixers is an exhibit of secure behaviour. About a thousand addresses are directly or indirectly linked with a mixing service. It is likely that the number of addresses that have an output to mixing services is underestimated. The algorithms of mixers are designed to obfuscate the origins of the output. As such, mixing services have a strong incentive to make sure their wallets are not clustered properly.

Among scholars there is debate whether gambling services are used in a structured manner to launder bitcoins. One the one hand, tariffs are often lower than mixing services. When placing many bets, one can calculate the guaranteed returns using the provably fair and public betting algorithms. Fanusie and Robinson (2018) and Paquet-Clouston et al. (2019) regard bitcoin gambling as a form of money laundering. On the other hand, the mixing would not be effective since input and output remain linked (Meiklejohn et al., 2013). As per Figure 6.17, just a few payout addresses are directly linked to gambling services and about 150 assumed private wallets had exposure to these platforms. This indicates that gambling *is not* structurally used to obfuscate money trails.

6.5.1.2. Payout analysis over time

To regard whether the type of wallet used on Hansa is consistent over time, the number of payouts to each wallet type per month were counted (Figure 6.18). The figure shows normalised counts and months with mostly missing data are not displayed. The figure shows that the proportions of wallet types used for the payouts remain somewhat consistent. Even transacting to service wallets directly - which could possibly lead to breaches of pseudonymity - is observed from the beginning of Hansa Market till the end.

Transacting directly to a known service does not equal a security risk per se. As explained in section 6.5, transacting to mixing services and peer-to-peer exchanges are considered relatively more secure behaviour compared to transacting to a regular exchange. Therefore, the types of services directly transacted to are shown in Figure 6.19. Here, non-normalised counts are visualised. The relatively low amount of observations per category would otherwise overestimate the precision of the graph. The figure only displays services *directly* transacted to. Services indirectly transacted to cannot be plotted over time. When payouts are first accumulated on private wallets, the amount of transactions to services and their transaction dates are very much

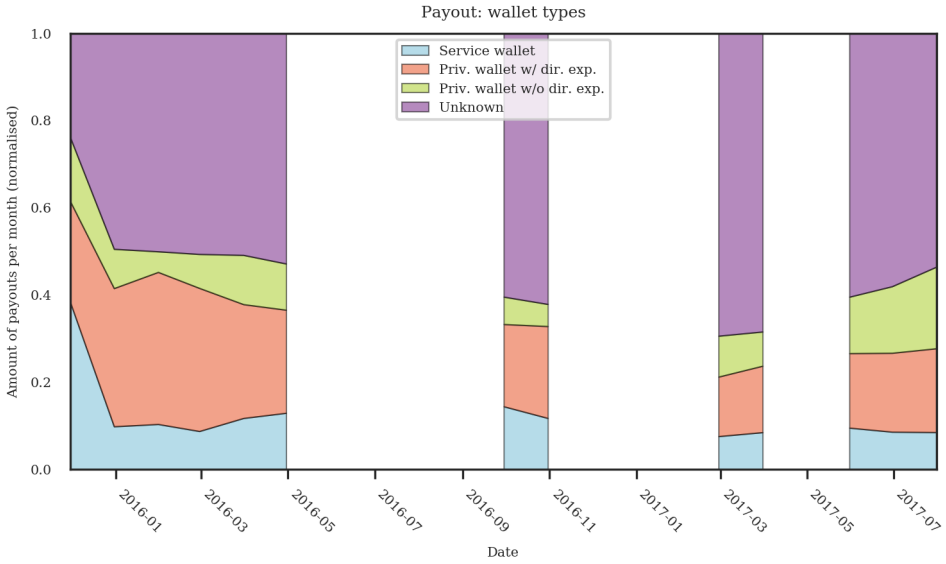


Figure 6.18: Wallet types payouts transacted to

dependent on the preferences of the vendor. How frequently they move their funds from the cold storage to the services for further processing, is up to them.

Direct payouts to exchanges are an ongoing event (Figure 6.19). From the perspective of Law Enforcement, it is good to see that over this extended period of time, this security risk keeps being created by vendors. The occurrences of this non-secure behaviour are further visualised per vendor in appendix D.1. From this analysis follows that 104 vendors have transacted directly to an exchange. For 29 vendors, this security risk was created during the first week of sales. Vendors that transact directly to exchanges, are very likely to do this repeatedly. A clear 'learning effect' could not be observed. Out of 29 vendors that make the mistake in the first two weeks, only 5 did not make the mistake again.

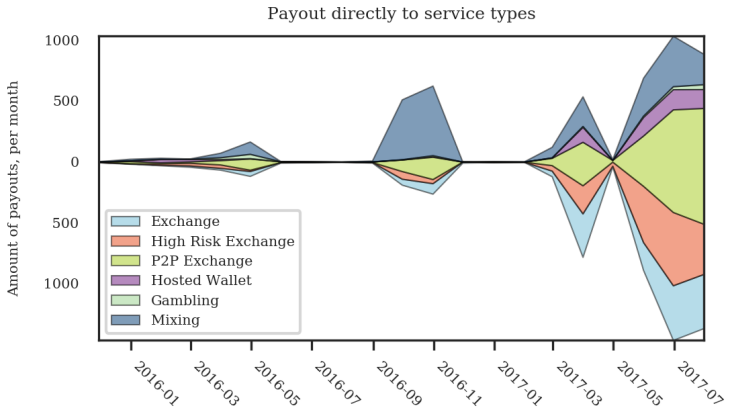


Figure 6.19: Services directly transacted to

Regarding direct payouts to mixers, Figure 6.19 indicates that there is an increasing amount of transactions until early 2017. This increase is both in absolute counts as relatively to the other services transacted to. From 2017 and on, the share of payouts directly going to mixing services seems to decrease.

Whereas the presence of mixing related transactions seem to decrease, the amount of transactions to P2P-exchanges and high risk exchanges increases. This indicates that P2P-exchanges gains

popularity over time. Of the vendors that are (in)directly linked with mixers, 26.8% also use P2P-exchanges. Given those vendors that are (in)directly linked with P2P-exchanges, only 18.5% are also linked with mixers. These numbers should be interpreted carefully, since it is not known how well Chainalysis clusters these mixers. It could be that over time, mixers become increasingly more difficult to cluster. Additionally, since unknown clusters with more than 5 addresses are not considered private wallets, larger vendors might not be included in the analysis.

In appendix D.2 the vendor characteristics are correlated against the choices of OFSPs. The conclusions are presented here. It is observed that many vendors with high numbers of sales can be directly or indirectly linked with central exchanges. This means that they have introduced this potential security risk at least once during their career. Of course, this does not imply immediate de-anonymisation of the vendor. It is not sure whether every exchange collected sufficient user data on its users during the 2015-2017 time frame, some exchanges could be located in jurisdictions that at that time were not likely to cooperate with the investigators. Also, the links may point to intermediaries such as money mules instead of the actual vendors. Despite these considerations, the links with central entities still might provide Law Enforcement with leads to investigate.

It was also observed that P2P-exchanges are popular among vendors that transact many physical or digital items. Vendors that are part of supply chains of drugs might use the cash money obtained through P2P-exchanges for buying new drugs of their suppliers. It is therefore surprising that also high selling digital vendors make use of P2P-exchanges.

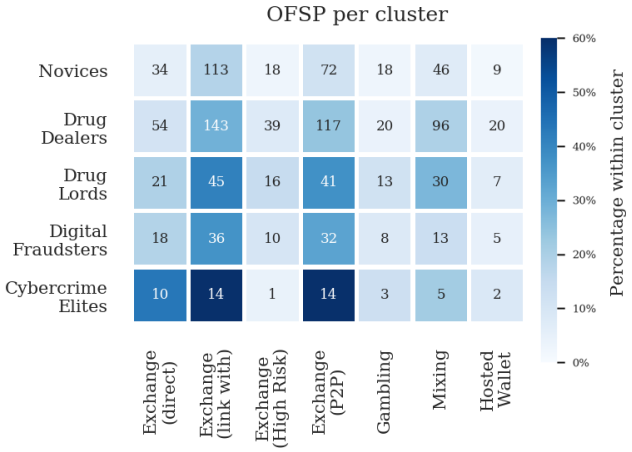


Figure 6.20: OFSPs linked with clusters

Figure 6.20 shows the links that could be established per cluster in absolute counts (numbers) and share within the cluster (color). In general, the differences between vendor profiles are subtle. `Novices` are not often linked to OFSPs, which may be because they transact less and have fewer listings and payout addresses. Relatively many `Cybercrime Elites` (14 out of 23) have transacted directly to exchanges. P2P-exchanges are most popular among `Cybercrime Elites` and `Drug Lords`.

6.5.2. Statistical analysis

Because of the heuristic and its assumptions (section 4.2.2.4) biasses and uncertainties are introduced in analysis of which OFSPs are used by vendors. The link that is irrefutable and that translates to a severe security risk is having the payouts directly transacted to a central exchange. The proportion of vendors of whom this behaviour is *not* observed is compared between

clusters. Differences are significant, as assessed by a χ^2-test ($p < 0.0000$, all cells but one (2.36) have an expected count >5). The results of the post-hoc test (Table 6.6) actually contradict the observations made during analysis of the other security behaviours.

Table 6.6: Results of z-test of proportions of vendors *not* transacting directly to exchanges, with FDR-BH adjusted p-values

	Novices		Drug Dealers		Drug Lords		Digital Fraudsters	
	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.	z-stat.	p-val.
Novices	.	.						
Drug Dealers	3.342	0.0017*	.	.				
Drug Lords	4.982	0.0000*	2.353	0.0233*	.	.		
Digital Fraudsters	4.577	0.0000*	2.067	0.0430*	-0.130	0.8968	.	.
Cybercrime Elites	7.045	0.0000*	4.604	0.0000*	2.489	0.0183*	2.541	0.0183*

Since positive z-values mean that the vendor profile indicated by the columns have higher security than the profiles named in the rows, it is shown that Novices show the most secure behaviour regarding transacting directly to exchanges. In this case, Drug Lords have significantly lower proportions of secure behaviour compared to Novices and Drug Dealers. The proportion of secure behaviour among Cybercrime Elites is the lowest compared to all other vendor profiles, as confirmed by the post-hoc test.

6.5.3. Interim conclusion & discussion: choice of OFSP

More vendors than initially expected transact their earnings directly to exchanges. When an address has a clear link with illegal activities, LEA have a reason to subpoena the exchange for information. This way, LEA may obtain transaction overviews, names, email addresses, residential addresses or IP addresses. Additionally, in extreme cases LEA are able to freeze funds stored at exchanges (Couvee, 2020). This research shows that throughout the the period that Hansa Market was active (2015-2017), each month about 10% of the payouts are directly transacted to exchanges.

Links that are not direct, i.e. via assumed private wallets, introduce more uncertainty in the analysis. A very conservative 5 addresses per unknown cluster is used to reduce the likelihood of service wallets being assumed to be private wallets. This is expected to underestimate the exposure of vendors who transact more often. Despite this possible underestimation, vendor types representing vendors that have a lot of transactions, are linked with regular and P2P-exchanges relatively often.

Regarding Drug Lords and Drug Dealers it is remarkable that these groups of vendors are linked relatively often. This shows that while these groups are considered to prioritise security, poor security practices are observed with regards to their cash-out behaviour. Thus, opportunities for effective cybercrime attribution may lie in this domain.

In earlier analyses it is explained that the suboptimal security practices of Cybercrime Elites and Digital Fraudsters might be related with the fact that they are not part of physical supply chains and as a result, have less evidence to hide. It was therefore expected that these vendors would excel in other parts of security behaviour: cashing out their criminal earnings. Especially because among this group, lots of vendors have financial products for sale (credit cards, PayPal accounts) or offer guides how to cash out.

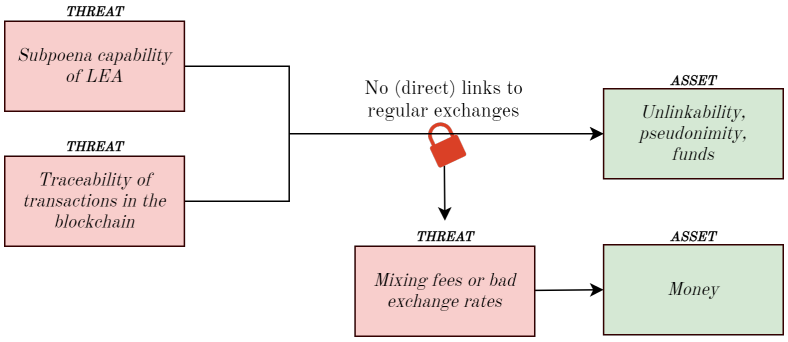


Figure 6.21: Preventing (direct) links to centrally organised exchanges reduces risks to assets created by threats.

6.6. Joint analysis of security behaviours

In this section vendors the overall security behaviour of vendors is assessed by a simple scoring method. Per vendor one point is awarded for for each of the following security practices: the vendor has a higher password complexity than median password complexity of vendors (67-bits), there is no password match in the PWND databse, 2FA is enabled, an 'extremely secure' PGP-key of more than 2048-bits is chosen and no direct transactions to exchanges have been observed. This results in the distribution presented in Figure 6.22.

Only a few vendors managed to score 0 out of 5 security points. About half of the `Digital Fraudsters` (49%) and half of the `Cybercrime Elites` (52%) have a score below 3. This is in congruence with the analyses of individual security behaviours, which showed that these clusters consistently have low proportions of secure behaviour. Important to note is that 15% of `Drug Lords` have a security score below 3. This is in line with the observations made when examining the relation between physical sales and the individual security behaviours. While the tendency is that security behaviour scales with the amount of physical sales, for every security behaviour exceptions were observed.

Since the scoring method is a simple sum of dichotomous scores, it is of no surprise that the score of 3 is common. The number of `Drug Lords` with a score of 3 is more than expected. However, not every security malpractice has an equal amount of 'impact' on the security of vendors. For example, the use of 2FA and extremely high PGP-key sizes should be interpreted as an indicator of security awareness or priority given to security rather than security mechanisms that prevent immediate de-anonymisation.

About a third of the `Cybercrime Elites` (35%) is scored with a 4 or 5. This shows that within this cluster of vendors there is actually a group of vendors that behaves securely (>3) and a group that does not (<3). However, the sample size within this cluster is not large ($n = 23$), this translates in relatively large differences increases or between security scores when these differences are expressed as percentages. The products sold by these two groups within the `Cybercrime Elites` are manually compared for differences. No obvious differences are noticed, most vendors in both groups have listings in numerous categories and do not tend to specialise in a single digital item sold. From this is inferred that these vendors are not skilled hackers themselves, but resellers of digital items purchased elsewhere.

No poor security behaviours are observed for 27% of the `Drug Lords`. Given this observation, it would be concluded that about a quarter of the `Drug Lords` are vendors that strive for 'maximum security'. When including vendors with slightly suboptimal security that score a 4, the `Drug Lords` (51%) do not differ much from the `Drug Dealers` (51%).

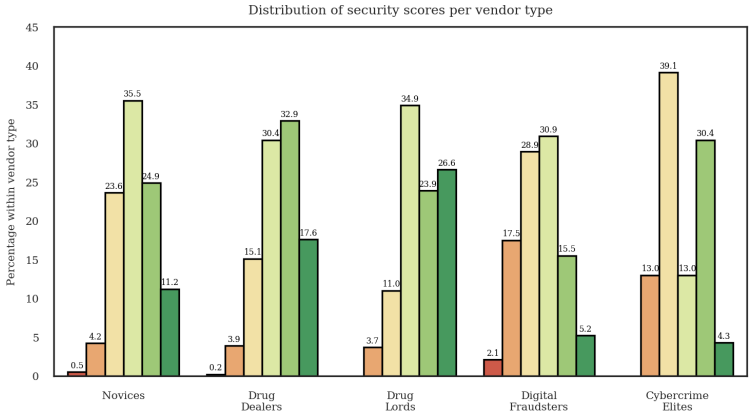


Figure 6.22: Distributions of security scores (0-5) per vendor type

6.7. Summary of findings and insights

Security Mechanism	*Result*	*Insights*
Complex password	Buyers use less complex passwords than vendors. `Drug Lords` and `Drug Dealers` have the most complex passwords. `Digital Fraudsters` have the least complex passwords. Some passwords of `Cybercrime Elites` are relatively simple while others are complex.	Authentication related security mechanisms do not require large investments in terms of time, knowledge or money. These security practices may be used as a proxy for how a cybercriminal prioritises security. Poor authentication related practices, might signal that security mistakes in other domains may be present as well. Paradoxically, the use of security mechanisms may introduce new risks to security, as demonstrated by how passwords can be used to link pseudonyms with other usernames or email addresses. When vendors have valuable informational assets, the protection of these assets tends to be more secure. However, exceptions have been observed: vendors with high sales and estimated revenue, that show relatively poor security behaviour.
Unique password	Non-unique passwords are more often found among buyers than vendors. Vendors belonging to `Digital Fraudsters`, and to a lesser extent `Novices`, relatively often make use of non-unique passwords. A few `Cybercrime Elites` have non-unique passwords.	
2FA	Buyers use less 2FA than vendors. A clear majority of `Drug Lords` and `Drug Dealers` use 2FA. Less than half of the `Digital Fraudsters` and slightly more than half of `Cybercrime Elites` and `Novices` make use of this additional layer of security.	

Security Mechanism	*Result*	*Insights*
PGP keys	Compared to older markets, the adoption rate of PGP is high (almost 100%). On Hansa, the adoption rate of PGP among buyers is lower compared to vendors. Only few vendors registered weak keys. Such weak keys are absent among `Drug Lords`. Together with `Drug Dealers`, this group of vendors registered extremely secure keys the most often. `Digital Fraudsters` and `Cybercrime Elites` registered extremely secure keys the least often. The security benefit of an extremely secure key over other keys is negligible till the year 2030.	As judged by the high adoption rate, cybercriminals tend to react to LEAs capabilities. Obfuscating capabilities might benefit their effectiveness. Additionally, vendors with valuable informational assets tend to use stronger encryption to secure these assets for a very long period of time. From this is inferred that some cybercriminals perform a kind of (subjective) risk assessment.
Linkability PGP	Within each vendor type, of roughly 20-25% of the vendors it is assumed that they use different PGP-keys between markets. Regarding this security measure, there are no significant differences between the vendor types.	The capability to match PGP-keys between darknet markets may disrupt cybercriminals' business. Choosing to use different PGP-keys seems to happen 'at random'. Since security aware vendors do not behave differently from vendors that behave generally less securely, it is inferred that a) not every one is aware of this capability b) there is no 'consensus' about whether this is a risk to security (hints towards subjectivity in risk assessments) or c) this risk is accepted for the gains of not using different PGP-keys.
OFSPs	During the time that Hansa Market was active, about 10% of the monthly payouts are directly transacted to exchanges. Vendors that transact more, are more often linked with OFSPs. `Novices` and `Drug Dealers` transact the least often directly to OFSPs, `Drug Lords` and `Cybercrime Elites` the most.	Relatively often security aware vendors make the mistake of being linkable to OFSPs. Such a link is a potential risk to pseudonymity. Since all vendors have to cash out eventually and 'mistakes' are somewhat prevalent, LEA should focus on behaviour related to OFSPs.

Security Mechanism	*Result*	*Insights*
Joint Analysis	Overall, `Drug Lords` and `Drug Dealers` show the most secure behaviour. However, only a quarter of `Drug Lords` have 'maximum security'. The security behaviour of `Cybercrime Elites` seems to be divided in two subgroups, no differences in items sold between these groups are found.	Maximum security is observed less than expected. Some vendors with high business success and experience, score low on multiple security behaviours.

7

Discussion: a deeper understanding

"The dark web is not as dark as you think"

Catherine De Bolle, executive director of Europol

In response to darknet market related arrests (Europol, 2019).

This thesis explores which factors influence security behaviour. This is done through analysing how 'vendor types', that represent vendors comparable with respect to their experience, activity on other markets and number of physical sales and digital sales, are related with security behaviour. Since most of the security behaviours differ significantly between vendor types, an approximately causal link between these concepts may be assumed. The research goal (section 1.2.2) stipulates that it is aimed to develop a deeper understanding of cybercriminal security behaviour. To this end, this section discusses the findings presented in this thesis in terms of a more abstract behaviour.

The results indicate that cybercriminals may perform a kind of (subjective) risk assessment. As elaborated upon throughout the interim-discussions, the behaviour of vendors may be explained in terms of risks to assets created by threats. Risk is determined by a) the probability of an event occurring and b) the resulting costs of such event. The costs are high when assets of high value are compromised. Examples of a cybercriminals' assets are: informational assets containing incriminating evidence, reputation, loyal customers, funds, pseudonimity, unlinkability and of course, freedom (as in, not being in jail). When there is less of [asset] to loose, the estimated cost of a 'security breach' is lower. The probability of a breach of security in cybercriminal context is determined by a) the likelihood that Law Enforcement starts an investigation and b) the likelihood that this investigation leads to the prosecution of the cybercriminal.

The risk paradigm described above can be used to explain why `Drug Lords` tend to behave more securely than `Novices`, who on their turn are exhibiting secure behaviour more often compared to `Digital Fraudsters`. Although jurisdictions vary greatly with respect to this, drug trade is mostly combatted through a law-and-order war on drugs (Revier, 2019) and intervening in drug trade is seen as an important objective of LEA (Alexandris Polomarkakis, 2017). Additionally, sentences for drug trade are high in many jurisdictions. Death penalties for drug offences apply in at least 35 countries (Girelli, 2019). Thus, `Drug Lords` play a risky game indeed. Combined with the observation that vendors transacting drugs have more valuable informational assets to secure from LEA, the risk paradigm might explain why the authentication related security mechanisms of vendors dealing drugs are more secure than vendors trading digital items.

This reasoning also explains why buyers tend to exhibit less secure behaviour compared to vendors. For the average buyer, the likelihood of being targetted by LEA is lower compared to the average vendor. Arguably, in case of prosecution, the punishment will be lower as well. Increased security as a result from more severe potential punishment, is also observed among a different kind of cybercriminals. Those engaging in the consumption or production of online child abuse material also seem to prioritise their security based on the severity of potential punishment and the likelihood of LE prosecution (National Rapporteur on Trafficking in Human Beings, 2011). 'Simple downloaders' often lack technological knowledge and are easily identified, whereas producers of online child abuse material have very high security standards Moran (2010).

The assessment of risk would be subjective, due to an inability to objectively assess the value of assets and the probability of a threat occurring, nor is it possible to estimate 'by how much' security mechanisms reduce risks to assets. Subjectiveness may also explain the differences *within* vendor types. For example, even `Drug Lords` do not behave consistently with regards to authentication security practices or preferred PGP-key lengths. Indeed, how much risk is reduced through using extremely secure passwords and PGP-keys or 2FA is based on a gut feeling at best. Subjective risk assessment does not imply full rationality, which could not be achieved due a lack of information. It may coexist with behavioural biasses, such as the status-quo biases mentioned earlier in this thesis. This may be the case for the `Digital Fraudsters` and `Cybercrime Elites` that are reselling digital items that they have purchased on other markets. Such vendors might not see themselves as true criminals being targetted by LEA (a subjective risk assessment). This perception does not create an urge to think about security and because this behaviour did not result in any issues, they might stick to their poor security practices (a behavioural bias). Similarly, Van de Sandt (2019, p.85) reasons that cybercriminals who are not fully aware of the criminality of their acts, tend to have no or little security mechanisms in place.

The findings in this thesis show that, compared to other vendor types, `Drug Lords` and `Drug Dealers` tend to have the most secure authentication practices and use extremely secure PGP-keys more often. However, regarding the linkability of pseudonyms through PGP-matching, no differences between vendor types are observed. With respect to (direct) links to exchanges, these vendor types show even less secure behaviour. Still, this does not invalidate the hypothesis of subjective risk assessments being performed.

Authentication related security mechanisms and the choice of PGP-key length have in common that they require little investments: to be secure, only some basic knowledge is needed. When using different PGP-keys between markets or when cashing out via mixers and P2P-exchanges, costs are incurred. For using different PGP-keys, these entail a potential loss of user base and reputation. Mixing services require mixing fees to be paid (Möser et al., 2013) and the exchange rates on the largest P2P-exchange is unfavourable compared to regular exchanges, as analysis of Pieters and Vivanco (2017) shows. Based on the observation that the generally securely behaving vendor types, show less or equally secure behaviour when security mechanisms are costly, the following may be inferred. When vendors are forced to trade-off between business and security, vendors may favour 'business' through accepting some risks.

The complexity of security mechanisms may also be used to explain suboptimal security behaviour. Throughout this thesis, it is shown that deploying security mechanisms may introduce new risks to security. This is the case for a) passwords that can be used create links with usernames in password databases, b) PGP-keys that may be exploited to link different darknet market pseudonyms with and c) even bitcoin transactions, that have been mistaken for being completely anonymous in the past (Reid & Harrigan, 2013). Currently, it is commonly known that transactions in the blockchain may be traced to central entities that can be subpoenaed

for information by LEA. Only when the workings of security mechanisms are fully understood by everyone using them, they provide adequate levels of security (Dourish & Anderson, 2006). When security mechanisms are not fully understood, they only address the subjective feeling 'that something is done' (Van de Sandt, 2019, p.86), which creates a false sense of security. An overestimation of the effectiveness of security mechanisms and an underestimation of their complexity leads wrong estimates of the risks exposed to assets.

From this discussion regarding subjective security risk assessments in a cybercriminal context a twofold conclusion is drawn. On the one hand, this research shows that suboptimal security behaviour is observed in numerous occasions. This behaviour differs significantly between vendor types and is thus not completely random. The risk paradigm provides explanations for less secure behaviour and the significant differences in security behaviour observed between vendor types. Generally put, cybercriminals deploy better security mechanisms when the perceived risk increases. This does not hold when these security mechanisms are costly (risks are accepted), or when security mechanisms have side-effects that may not be understood well (risks are underestimated). On the other hand, security behaviour varies highly within vendor types. The joint analysis of security behaviours and the distribution of security behaviour within vendor types made this clear. While it is attempted to explain this through the subjectiveness in risk assessment (and to a lesser extend, behavioural biasses) it should be taken into account that the findings may also be interpreted otherwise. Additionally, the chosen approach, measurement methodology and data all have their limitations. These are elaborated upon in section 8.3.

8

Conclusion

> "The digital part is only one aspect of the entire process [...] you have the entire digital process, the physical process and the financial process. That all has to be flawless, every day again. That's a challenge"
>
> *An anonymous cybercrime expert*
>
> Interviewed by Van Hardeveld (2018, p.154)

8.1. Answer to the research questions

> *Which factors influence the security behaviour of darknet market vendors active on Hansa Market?*

Approximately causal relationships are inferred between on the one hand vendor types, that represent a combination of business success in terms of physical and digital sales, experience and activity on other markets and on the other hand security behaviour. Among vendors selling mainly physical items, an increasing amount of business success, experience and activity on other markets, results in an increase of more secure behaviour regarding simple security measures such as authentication related practices. In general, vendors that transact mainly digital items perform relatively poor with respect to these behaviours, although the spread within the group of highly successful digitally focussed vendors is large. When security mechanisms are costly and vendors have to trade-off between increased profits and security, relatively more non-secure behaviour is observed within the groups of vendors that exhibited very secure behaviour otherwise. When security mechanisms are more complex, similar observations are made. In this thesis, the (differences in) suboptimal security behaviours observed are successfully explained by stipulating that vendors on Hansa Market conduct *subjective risk assessments*. This implies that the probability of being targeted by LEA and the value of the vendors' assets that are at stake (informational assets containing incriminating evidence or 'years of freedom') are of influence on security behaviour.

8.1.1. Answer to the sub-questions

- **SQ1**: *What security behaviours can potentially be observed on darknet marketplaces?*

Knowledge from different academic disciplines is accumulated to draft a theoretical framework of security behaviours that apply to darknet market vendors. First, security is defined as acts

of "compromising the availability or usefulness of evidence to the investigative process". Four course categories of security behaviour have been defined, being data hiding, trail obfuscation, data destruction and data minimisation. *Data hiding practices* that apply to darknet market vendors are those entailing: a) the encryption of communication (PGP-usage, PGP-key length, auto-encryption usage, using other communication services), b) authentication security (2FA usage, password strength and uniqueness), c) crossing jurisdictions (shipping preferences), actively increasing linkability of darknet market pseudonyms (similar usernames, reusing PGP-keys over different markets) and d) passively allowing such linkage (being susceptible to advanced linking techniques such as authorship analysis). *Data minimisation and destruction* practices may be mentioned in vendors' profile descriptions, terms and conditions or listings. *Trail obfuscation* mechanisms are a) not having clear links to central bitcoin exchanges, b) using mixing services and c) using gambling services, although the effectiveness of the latter is debated. Some security practices are combinations of data hiding, minimisation and trail obfuscation. P2P-exchanges and central bitcoin exchanges with weak AML/KYC controls do not perform strict identity checks. This means that no personal information (data minimisation) or falsified identities (trail obfuscation) can be registered at these exchanges. P2P-exchanges hide links between payment systems (e.g. bitcoins and physical fiat currencies) from the investigator because transactions via peers are facilitated. Central bitcoin exchanges with weak KYC/AML controls are located in 'crypto-friendly' jurisdictions and thus are able to hide data because these jurisdictions do not take part in international agreements.

- **SQ2**: *What vendor characteristics are relevant to include when assigning vendor types to vendors?*

  Indications are found that business success, experience, activity on other markets and whether a vendor transacts physical or digital items relate to security behaviour. Physical items offered for sale on darknet markets are mostly drugs, e.g. cocaine, cannabis and heroin. A broad selection of digital items can be found on darknet markets, e.g. banking credentials, credit card details, identification documents, email and password combinations acquired in data breaches, accounts for streaming services such as Netflix and Spotify, login details for adult websites and cash-out solutions that facilitate in money laundering. It is decided to measure business success by regarding the amount of physical items sold and the amount of digital items sold.

- **SQ3**: *What types of vendors can be distinguished on Hansa Market?*

  Based on the defined characteristics, LPA clustering is performed to group similar vendors together. A 5-cluster model is selected. The resulting vendor types sufficiently differ in terms of the vendor characteristics, as assessed through statistical testing. The clusters are labelled and explained as follows. `Novices` ($n = 988$) score low on all vendor characteristics and only 40.2% of the these vendors is active on other markets. 80% of the sales made by `Novices` are drugs related. This means that not *all* vendors in this cluster are low selling drug dealers. The `Drug Dealers` ($n = 509$) have more physical sales, experience and generate more revenue compared to `Novices`. 69.0% of the `Drug Dealers` are active on other markets. Approximately 98% of the sales are drugs related. `Drug Lords` ($n = 110$) have extremely high amounts of physical sales and are mostly (78.2%) active on other markets. They sell only drugs (100%). Two clusters of vendors that thrive in digital sales are identified. `Digital Fraudsters` ($n = 103$, 58.3% active on other markets). 75% sold 100+ digital items. And `Cybercrime Elites` ($n = 23$, of which 73.9% active on other markets), all vendors sold 1000+ digital items. All kinds of digital items are offered by the vendors within both vendor types.

- **SQ4**: *How do the vendor types compare relatively to each other in terms of the security behaviours analysed?*

From the identified security behaviours, three authentication security related practices (password strength, password uniqueness and 2FA usage), the encryption of communication (PGP-key lengths used), the linkability of pseudonyms (PGP-matches in the *Grams* search engine database) and the choice of Online Financial Service providers (OFSPs) are analysed.

Authentication security Vendors tend to have a higher **password complexity** compared to buyers. Both the `Drug Lords` and `Drug Dealers` have statistically significant higher password complexities than the `Novices` and `Digital Fraudsters`. No statistically significant differences are found between any cluster and the `Cybercrime Elites`. This is due to the large spread within this cluster combined with the small sample size. It is concluded that password complexity seems to scale with the amount of sales and experience for the vendors that sell physical items. Vendors that sell digital items have relatively simple passwords in general. Additionally, observations of highly successful vendors with very weak passwords have bee made.

A **unique password** is important. For a cybercriminal, the reuse of a password introduces two security risks. Firstly, libraries of leaked passwords are used by password cracking software to guess passwords very quickly. Secondly, these databases often consist of passwords paired with usernames or email addresses. Law Enforcement may be able to match the passwords to these usernames or email addresses. This introduces the risk of de-anonymisation.

This research matched passwords of Hansa users in the 'Have I Been PWND' database, which consists of more than 10 billion (573M unique) passwords. 30% of buyers' passwords and 17% of vendors' passwords were matched in the PWND data. Within all clusters, fairly unique (1-9 matches) and common (100+ matches) passwords were found. This means that again, some successful vendors do make severe security mistakes. With about 27% of the `Digital Fraudsters`' passwords being matched in the Hansa data, this group shows the most non-secure behaviour, which is confirmed through statistical testing.

Only 60.5% of the vendor population had **Two-Factor Authentication** enabled. There is a negative correlation between password complexity and 2FA usage. This indicates that 2FA is *not* used to compensate poor passwords with. The analysis of the vendor characteristics correlated with password reuse shows that many experienced and high selling vendors did not enable 2FA. When regarded per vendor profile, it is revealed that most `Drug Lords` (82%) and `Drug Dealers` (71%) have 2FA enabled. Statistical tests confirm that this is significantly more than the proportions of 2FA usage within `Novices` (55%) and `Digital Fraudsters` (44%). Of the `Cybercrime Elites` only 57% had enabled 2FA, which is a low amount, given their business success. Because of the low sample size, only the difference with `Drug Lords` resulted in a significant test.

Encryption of Communication All but 5 vendors have PGP-keys listed on their profiles. The PGP-adoption among buyers is noticeably lower (12%). The metadata enclosed in the PGP keys are extracted and analysed. The following observations are made on the cryptographic strength of the PGP-keys. Only 9 vendors used weak ($\leq$1024-bits) keys. The other keys of vendors are 2048-bits (53%) or stronger (47%). An analysis of the creation dates and of keys and the key sizes shows that PGP-keys do not get stronger or weaker over time. Till at least 2030, there are no security benefits of key sizes beyond 2048-bits. Thus, it was expected that key strengths are selected randomly or, more likely, based on one of many online tutorials. However, differences in key sizes are significant between vendor types. The proportions of extremely secure 2048+ bit PGP-keys are highest among `Drug Lords` and `Drug Dealers`.

Actively increasing the linkability of pseudonyms This analysis only applies to the vendors that are active on other markets. Their PGP-keys are matched against the database of the *Grams* darknet market search engine. Out of the 908 vendors known to be active on other markets, 643 could be linked via their PGP-key. This implies that 265 vendors use different keys on other markets. No significant differences were found between the vendor types.

Choice of OFSPs The bitcoin addresses to which the payouts of vendors are transacted are analysed to gain understanding which OFSPs are used. Out of 19,238 payout addresses, 14% could be directly attributed to a known service. The other payout addresses belong to either assumed private wallets (42%) or to wallets of which it remains unknown whether these belong to services (43%).

Findings include that each month, roughly 10% of the transactions could be directly linked to central exchanges. These direct links have the potential to be severe risks to security and are observed among `Drug Lords` and `Cybercrime Elites` the most. P2P-exchanges are popular among `Drug Lords` and `Cybercrime Elites`.

- **SQ5**: *How do the vendor types compare relatively to each other when all security behaviours are considered jointly?*

  Overall, `Drug Lords` and `Drug Dealers` show the most secure behaviour. About a quarter of the `Drug Lords` have 'maximum security'. `Digital Fraudsters` and `Cybercrime Elites` show poor security behaviour most often. However, still 15% of the `Drug Lords` exhibit relatively much non-secure behaviour. Likewise, a third of the `Cybercrime Elites` do show mostly secure behaviour.

- **SQ6**: *How can these differences in security behaviour be explained?*

  In this thesis it is posited that differences in security behaviour can be explained via the notion of subjective risk assessment. Because most security behaviours between vendor types significantly differs, the behaviour is not completely random. The risk paradigm provides explanations for non-secure behaviour and the significant differences in security behaviour observed between vendor types. Generally put, cybercriminals deploy better security mechanisms when the perceived risk increases. The level of risk is related to the estimated probability of being investigated by LEA and the estimated costs of compromised assets in case this investigation is successful (informational assets containing evidence, pseudonimity, years of freedom). Stronger security mechanisms are needed when risks increases. The estimations of probabilities, costs and amount of risk reduced by security mechanisms are subjective. These cannot be objectively assessed. When security mechanisms are costly, risks may be accepted. When the complexity of security practices is not understood well, e.g. when they have side-effects, a false sense of security may be created. An overestimation of the effectiveness of security mechanisms and an underestimation of their complexity may lead to non-secure behaviour.

8.2. Scientific relevance

In section 1.1 the following gaps in academic knowledge are identified. These are addressed as follows.

- *A limited conceptual insight in the security behaviour of darknet market users*

  Two contributions are made that provide a comprehensive conceptual insight in the security behaviour of darknet market users. Firstly, this research presents a theoretical framework (Figure 3.1) that draws from security behaviours described in multiple academic disciplines. A definition of 'security' from the perspective of darknet market users is given and high level behaviours (e.g. 'data hiding') are related to very specific behaviours that may be observed on darknet markets (e.g. the use of auto-encryption). Additionally, it shows that cybercriminals may increase their security by choosing OFSPs wisely. These provide opportunities for data minimisation, data hiding and trail obfuscation. A few relations between behaviours and their influence on security are unclear. The work presented here shows that it is unlikely that gambling platforms are used to obfuscate money trails at scale. Other ambiguous relationships between behaviours and security are not investigated further in this thesis. However, the findings presented here can be used to clarify these relations (see section 8.3.2). Compared to earlier research, e.g. Soska and Christin (2015) who regard mentioning a PGP-key as a simple proxy for security behaviour, Van Wegberg and Verburgh (2018) in which evasive measures are analysed in terms of PGP-key and username changes and Décary-Hétu et al. (2016), who operationalise 'risk' as the willingness to ship internationally, this framework is a significant addition to the understanding of the concept of security behaviour on darknet markets. Secondly, empirical evidence is presented on the relationship between concepts related to characteristics of vendors and security behaviour. For vendors that trade in physical items, business success and experience are approximately causally related with security behaviour.

- *It is unknown to what extend suboptimal security behaviour can be observed on darknet markets, among what types of darknet market users suboptimal security behaviour is most prevalent and what might cause this suboptimal security behaviour.*

  This exploratory research addresses these gaps in current knowledge by showing that 'maximum security' is often not achieved on a darknet market. Vendors that trade in digital items, especially those with a relatively low amount of sales, exhibit non-secure behaviour most often. Through quantitative analysis of differences in security behaviours an a qualitative interpretation of these results, it was found that cybercriminal security behaviour on darknet markets may be explained using a risk paradigm. Thus, the work presented in this thesis extends existing research on the microeconomic approaches to cybersecurity, by providing evidence these principles also apply in a cybercriminal context. This supports the predominantly qualitative findings of van de Sandt, who argues that many 'deviant security practices' can be explained by risk assessments conducted by cybercriminals (Van de Sandt, 2019, pp.64-65).

8.3. Recommendations and future work

8.3.1. Recommendations to LEA

- Consider targeting vendors that sell illicit digital items on darknet markets. Vendors that sell less than a thousand digital items, are most likely not to prioritise security. On Hansa Market, about half of the vendors that transacted 1000+ digital items has below average security practices. Although simply reselling digital items purchased on other markets may not be regarded as very serious crime, LE may impact the cybercriminal ecosystem by intervening early. Regarding vendors selling digital items, LE may prevent cybercriminals

growing into being involved with more severe digital crimes. This can be combined with programs to divert youngsters from cybercrime, such as Hack_Right (Dutch Police, n.d.).

- Additionally, while vendors transacting extreme amounts drugs tend to behave more securely, exceptions to this rule are observed. Despite the business success and experience of a vendor, there is always a probability that security mistakes are made.

- Exploit the subjectiveness in risk assessments. For example, cybercriminals who regard themselves as not being of interest to LEA, might have less secure mechanisms in place. Additionally, the effects of security mechanisms may be misunderstood, creating a false sense of security. When new security technologies are introduced, it may take a while for cybercriminals to apply these correctly. This provides a window of opportunity for LEA.

- When there is a trade-off between financial gains and security, opportunities arise for LEA. As shown in the analysis of OFSPs and the analysis of linkability of pseudonyms, relatively more non-secure behaviour among otherwise very secure vendors is observed when security mechanisms are costly in terms of money or reputation. When focussing on the behaviours that involve a trade-off between increased financial gains and security, in the best case vulnerabilities in cybercriminals' security behaviour are found and in the worst case LEA disrupt profitable businesses of darknet market vendors by enforcing extra security costs on them.

8.3.2. Recommendations for further research

This section discusses the limitations and provides recommendations for further academic research. Firstly, limitations and subsequent further research regarding the approach is elaborated upon. Then, suggestions are made based on the limitations of the measurement methodology. Lastly, shortcomings and recommendations regarding the interpretation of results are made.

8.3.2.1. Research approach

- Regarding the research approach, it must be noted that only a single case is studied. This negatively impacts the generalisability, or external validity, of the findings (Seawright & Gerring, 2008; Yin, 1993). The Hansa Market was active from 2015 till 2017. Because of large internationally coordinated LE interventions, such as Operation Bayonet in which Hansa was involved, it is likely that security behaviour has changed since then. When follow-up research on another darknet market is conducted, it can be observed how security behaviour changes over time and whether significant differences between security practices of vendor types still exist. Within the context of cybercriminals trading illicit goods on the dark web, also other types of platforms may be studied.

- Secondly, because of the exploratory nature of this research, approximate causal relations between vendors and their security are inferred by comparing the security behaviour of vendor types. This gives way to aggregation biasses, in which the heterogeneity within the group-level measure is not taken well into account. It is therefore suggested to perform similar research that focuses on a single type of vendors active on a large darknet market (e.g. drug dealers on AlphaBay). Doing so, may allow researchers to describe the influence of vendor characteristics on security behaviours via proportions of explained variance (R^2), which is a more common approach in behavioural research.

- Lastly, this research takes vendor types as units of analysis. Distinguishing between types of buyers could confirm or reject the hypothesis that subjective risk assessments are conducted by darknet market users.

8.3.2.2. Measurement methodology & data

- Because not all data from the Hansa Market could be recovered, this research has to deal with missing data. The number of sales could be estimated using data on the number of feedbacks per listing. This introduces a certain bias, since buyers provide feedbacks on digital items less frequently compared to physical items. Whereas the number of sales could be estimated, some data used for measuring security behaviour could not be inferred from other data. As a result, the analysis of password strength, password uniqueness and choice of OFSPs is performed on a more limited dataset compared to the other analyses. In future research, it is recommended to only include vendors of whom full information is available, if the sample size permits. Additionally, the *Grams* database is not fully accurate. Further research into the accuracy of such databases, will aid in interpretation of the analysis of the linkability of pseudonyms.

- This thesis only distinguishes physical items from digital items. Since the findings indicate that security behaviour is not at random, additional insights may be gained by considering a more refined classification of items transacted. Unfortunately, digital items are listed in many (irrelevant) categories on darknet markets. Therefore it is suggested to train a classifier to cluster similar listings (Van Wegberg, Tajalizadehkhoob, et al., 2018).

- Compared to the theoretical framework (Figure 3.1) not all security behaviours identified are included in the conceptual model (Figure 4.1). Vendors that prioritise security have been identified in this research. This can be leveraged when investigating the causal link between e.g. mentioning data destruction practices on a profile and actual security behaviour.

- Because of privacy considerations (section 8.4), only hash values of passwords and no usernames are made available to the researcher. Therefore, no differentiation could be made between randomly generated and human-generated passwords. As a result, Equation 4.2 was used to estimate the complexity of both types of passwords. Further research should consider separate methods for estimating password entropy of human-generated and randomly generated passwords. Linkability of pseudonyms can be better assessed when not only PGP-keys, but also usernames are considered.

- *Chainalysis* is queried for information on blockchain transactions. When unknown clusters of addresses are returned, all clusters with more than 5 addresses are not considered further in this thesis. It is reasoned that these might belong to services, instead of individuals. This number is very conservative number. Further research should experiment with this threshold. Additional checks, such as analysing the types of addresses clustered or the amount of transactions (per unit of time) could be considered to estimate whether an unknown cluster belongs to a service entity or not with greater precision.

- The scoring method used to analyse the security behaviours jointly is limited. Different types of security behaviours are regarded as equals, which is contra-intuitive. In addition, taking the median password complexity as cut-off point is arbitrary. Factor analysis may be used for a more insightful scoring system.

8.3.2.3. Interpretation of findings

- This thesis adopts a microeconomic perspective to explain the security behaviours observed. While full rationality is not implied, some rationality is assumed when (un)knowingly estimating risks. To what extend this rationality may be assumed, is up to debate. Other researchers should reject the notion of subjective risk assessments through explaining the results using other theories or should confirm the explanation by investigating some phenomena more closely. For example, it is suggested to research whether vendors that transact drugs to, or from, jurisdictions with heavy punishment for drug offences have higher security standards compared to those who do not ship to or from these jurisdictions.
- Some data might be misinterpreted due to security mechanisms that could not be observed at scale. This would be the case for high selling vendors that use money mules for cashing out. From the data non-secure behaviour might be inferred (a direct link to an exchange), the behaviour is in practice very secure. Such behaviour can be better understood by focusing on a small group op vendors through analysis of case law or interviews.

8.4. Ethical considerations

It is important to consider research ethics when studying online communities (Buchanan & Ess, 2009). Thomas, Pastrana, Hutchings, Clayton, and Beresford (2017) provide a framework for evaluating the ethical considerations of researching illicitly obtained datasets. While the Hansa Market data has been legally obtained by a mandated LEA, has been cleared for research and the researcher has been authorised to conduct research on this data, this framework helps in discussing these matters in a structured manner. The authors use this framework to evaluate research ethics regarding scraped darknet market data. Such data is resemblant to the data used in this thesis. Next to the Hansa data, the PWND password database and Grams darknet market search engine data are used in this research. These are considered as well when deemed necessary. Following Thomas et al. (2017), potential legal issues and ethical considerations are reflected upon in table 8.1 and table 8.2 respectively.

Table 8.1: Assessment of legal issues relevant to this research

Legal Issue	*Explanation*	✓/✗*
Computer misuse	The Hansa data is legitimately obtained by LEA. The password database uses data from illicit sources. This is of no legal concern because a) only password hashes are made publicly available, b) the owner of the password database behaves ethically and does not charge for the use the anonymised data. For an elaboration on the legitimisation of using this password database, see Hunt (2018). The copy of the Grams data has been acquired via institutional resources.	✗
Copyright	The Hansa data used for this research will not be shared among researchers. While this reduces the reproducibility of findings and thus impacts research ethics, it ensures that data are not unlawfully distributed.	✗

Data privacy	The market data has been anonymised. Only unique identifiers or hashes of potentially personally identifiable data are made available. The data have been cleared for research purposes and are approved by the appointed privacy officer of the FIOD. The password database does not contain any personally identifiable data. For the Grams database analysis, only unique identifiers and PGP-keys are used. These are not made public in this thesis.	✗
Terrorism	Acts of terrorism are not discussed on darknet markets. Thus, failing to report terrorist activity is not likely.	✗
Indecent images	Photos hosted by the darknet market have been not been made available to the researcher.	✗
National security	Data have been cleared for research purposes and the researcher is authorised to make use of these data. No legal risks apply to the researcher.	✗

* ✓ this issue of significant concern, ✗ this is not of significant concern

To discuss potential ethical issues, it is important to identify the major stakeholders Thomas et al. (2017). The authors define primary stakeholders, intermediaries and key players. Hansa market users and LEA are the primary stakeholders. The analysis of ethical issues analyses potential ethical concerns of this group (Table 8.2). In this research, the intermediaries used for data gathering and analysis are HaveIBeenPwnd, Chainalysis and Grams. This research did not put a significant strain on their networks: the PWND databse is downloaded once and shared within the FIOD, Chainalysis is a large service provider contracted by LEA and the Grams database consists of an offline copy. Lastly, key players are the researcher and academic committee, who are aware of their names being published in this work.

Table 8.2: Assessment of ethical issues relevant to this research

Consideration	*Explanation*	✓/✗/•*
Harms	*Illicit measurement*: unlike darknet market scrapes, there is no debate on the legality of the means used to acquire the Hansa Market dataset. The password database however, is consists of hacked or leaked data (Hunt, 2018). The use of these data are justified below, using the 'not the first' principle.	•
	Potential abuse: insights presented in this thesis might be used by cybercriminals to enhance their security. This is justified below using the 'no significant additional harm' and 'not the first' principles.	•
	De-anonymisation: instead of personally identifiable data, hashes and unique identifiers are used. The data have been approved by the privacy officer of the FIOD.	✓

	Sensitive information: the findings may negatively impact cybercriminals, due to new insights to LEA. Using a utilitarian perspective, this is easily justified because the positive impact on society is larger than the negative impact on the individual. The other way round also holds. Some information disclosed here may be perceived as sensitive information on LEA benefiting cybercriminals. This is justified through the 'no significant additional harm' and 'not the first' principles below.	•
	Researchers harm: according to Barratt and Maddox (2016), harm being done to the researcher is of low ethical concern when: the researcher did not interact with darknet market users, findings do not harm individuals directly and no legal risks apply to the researcher (Table 8.1).	✓
	Behavioural change: darknet market users may change their behaviour in such a way that future data may become unreliable, e.g. through not leaving behind feedbacks or faking sales. This is justified below through the 'public data' principle. Another potential harm is that users may change their suboptimal security practices to more secure behaviour. This would make it more difficult for LEA to gather evidence on future darknet market users. The 'no significant additional harm' and 'not the first' principles justify the analyses presented in this thesis.	•
Safeguards	*Secure storage*: data are stored within a secured virtual environment only accessible by authorised users. Raw data have not been taken out of this environment.	✓
	Privacy: The market data has been anonymised, only unique identifiers or hashes are made available.	✓
	Controlled Sharing: data will not be shared with other researchers.	✓
Justice	Data used in this research does not put certain demographic groups (in terms of race, gender, age *etc.*) more in danger than other groups	✓
Public interest	This research serves the public interest, as described extensively in section 1.2.4.	✓
Benefits	*Uniqueness*: the Hansa data is unique in a way that some data cannot be obtained through scrapes. There are no alternatives for the password database that are not based on hacks or data leaks. These are valid reasons to use the data, if the research is also deemed useful (Thomas et al., 2017).	✓
	Defence mechanisms: the data are used to develop insights for more effective enforcement of the law, protecting democracy against disruption by cybercriminals	✓
	Anthropology and transparency: the research goal (section 1.2.2) can only be attained through using empirical real-world data. Lab experiments or surveys will not produce valid insights, because these cannot simulate the cybercriminal context. Likewise, there is also no valid simulated alternative to the PWND data for assessing password uniqueness.	✓

Justifications	*Not the first*: existing research uses the same darknet market data used in this research (Grapperhaus, 2019; Stinenbosch, 2019). Additionally, cybercriminal behaviour has been extensively researched via darknet market scrapes. Moreover, academic work in which accounts are linked via PGP-matching (Van Wegberg & Verburgh, 2018) exists and there seems to be a consensus on the justification of using databases of hacked or leaked passwords for research purposes (Martin & Christin, 2016)	✓
	Public data. Information displayed on darknet markets may not be intended to be publicised (Flick & Sandvik, 2013). Still, compared to for example conventional drug trade, darknet markets are transparent and open. Users deploy pseudomising technologies, which allow information to be accessible without immediate consequences. Because registration on darknet markets is open to everyone, Christin (2013) considers darknet market data to be public. Next to the absence of significant entry barriers, Eysenbach and Till (2001) gives other reasons why such data should be considered public: there is a large membership and it can be assumed that most darknet market users are aware of the fact that information being posted is monitored, because these markets have been under scrutiny LEA and academics for years. However, this research also makes use of non-public information that can only be found in the back-end of the market.	✗/✓
	No significant additional harm: through anonymising the data, individuals will not be directly impacted by this research. Also, the data have been acquired in 2017 and are not seized for the sole purpose of doing research. Regarding harm being done to LEA by disclosing investigative capabilities, the following is important to note. None of the displayed techniques are based on inside information given by LEA. They are the researchers' own and are based on previous academic work. Additionally, this thesis is not written as a 'cybercriminals handbook'. All information on secure illicit trading found in this thesis is widely available on both clearnet and darknet. Guides and tutorials can be freely obtained on discussion fora and darknet markets. These will provide a more holistic and up-to-date perspective on 'opsec' than this thesis does.	✓
	Necessary data: this justification applies to every harm and is valid when there is sufficient public interest (Thomas et al., 2017). Without the data sources used in this thesis, the research could not be conducted.	✓
Ethics section	Ethical considerations are elaborated upon extensively	✓

* ✓ applicable to this research, ✗ not applicable to this research, • applicable, but justified

Bibliography

Afroz, S., Islam, A. C., Stolerman, A., Greenstadt, R., & McCoy, D. (2014). Doppelgänger finder: Taking stylometry to the underground. In *2014 ieee symposium on security and privacy* (pp. 212–226). IEEE.

Aldridge, J., & Askew, R. (2017). Delivery dilemmas: How drug cryptomarket users identify and seek to reduce their risk of detection by law enforcement. *International Journal of Drug Policy*, *41*, 101–109.

Aldridge, J., & Décary-Hétu, D. (2014). Not an'ebay for drugs': The cryptomarket'silk road'as a paradigm shifting criminal innovation.

Alexander, J. C. (1987). *The micro-macro link*. Univ of California Press.

Alexandris Polomarkakis, K. (2017). Drug law enforcement revisited: The "war" against the war on drugs. *Journal of Drug Issues*, *47*(3), 396–404.

Anderson, R. E., & Srinivasan, S. S. (2003). E-satisfaction and e-loyalty: A contingency framework. *Psychology & marketing*, *20*(2), 123–138.

Athey, S., Parashkevov, I., Sarukkai, V., & Xia, J. (2016). Bitcoin pricing, adoption, and usage: Theory and evidence.

Baldwin, D. A. (1997). The concept of security. *Review of international studies*, *23*(1), 5–26.

Bancroft, A., & Scott Reid, P. (2017). Challenging the techno-politics of anonymity: The case of cryptomarket users. *Information, Communication & Society*, *20*(4), 497–512.

Barker, E., Barker, W., Burr, W., Polk, W., Smid, M., et al. (2020). *Recommendation for key management: Part 1: General, 5th rev.* National Institute of Standards and Technology, Technology Administration.

Barker, E., & Dang, Q. (2015). *Recommendation for key management, part 3: Application-specific key management guidance*. National Institute of Standards and Technology.

Barker, E., & Roginsky, A. (2010). Recommendation for the transitioning of cryptographic algorithms and key sizes. *NIST SP 800-131, 2010, Technical Report*.

Barr, W. (2020). Statement from attorney general william p. barr on introduction of lawful access bill in senate. Retrieved August 5, 2020, from https://www.justice.gov/opa/pr/statement-attorney-general-william-p-barr-introduction-lawful-access-bill-senate

Barratt, M. J., & Maddox, A. (2016). Active engagement with stigmatised communities through digital ethnography. *Qualitative research*, *16*(6), 701–719.

Bauer, J. M., & van Eeten, M. J. (2009). Cybersecurity: Stakeholder incentives, externalities, and policy options. *Telecommunications Policy*, *33*(10-11), 706–719.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, *57*(1), 289–300.

Bokhorst, R., Steeg, M., & de Poot, C. (2011). *Rechercheprocessen bij de bestrijding van georganiseerde criminaliteit*. WODC.

Bonneau, J. (2012). Statistical metrics for individual password strength. In *International workshop on security protocols* (pp. 76–86). Springer.

Boxerman, S. J., & Schwerin, M. F. (2016). Its bark is worse than its bit (e): Regulatory and criminal law implications of virtual currency. *Crim. Just.*, *31*, 10.

Bradley, C. (2019). *On the resilience of the dark net market ecosystem to law enforcement intervention* (Doctoral dissertation, UCL (University College London)).

Bradley, C., & Stringhini, G. (2019). A qualitative evaluation of two different law enforcement approaches on dark net markets. In *2019 ieee european symposium on security and privacy workshops (euros&pw)* (pp. 453–463). IEEE.

Branwen, G. (2020). Darknet market archives. Retrieved August 5, 2020, from https://www.gwern.net/DNM-archives

Broadhurst, R., Ball, M., & Trivedi, H. (2020). Fentanyl availability on darknet markets. *Broadhurst R, Ball M & Trivedi H.*

Broséus, J., Morelato, M., Tahtouh, M., & Roux, C. (2017). Forensic drug intelligence and the rise of cryptomarkets. part i: Studying the australian virtual market. *Forensic science international*, *279*, 288–301.

Broséus, J., Rhumorbarbe, D., Mireault, C., Ouellette, V., Crispino, F., & Décary-Hétu, D. (2016). Studying illicit drug trafficking on darknet markets: Structure and organisation from a canadian perspective. *Forensic science international*, *264*, 7–14.

Buchanan, E. A., & Ess, C. M. (2009). Internet research ethics and the institutional review board: Current practices and issues. *ACM SIGCAS Computers and Society*, *39*(3), 43–49.

Burnett, M. (2006). *Perfect password: Selection, protection, authentication*. Elsevier.

Burr, W., Dodson, D., & Polk, W. (2006). *Information security: Electronic authentication guideline nist*. Technical report, Tech. Rep. Special Rep. 800-63.

Burr, W., Dodson, D., & Polk, W. (2004). *Electronic authentication guideline*. National Institute of Standards and Technology.

Cabaj, K., Caviglione, L., Mazurczyk, W., Wendzel, S., Woodward, A., & Zander, S. (2018). The new threats of information hiding: The road ahead. *IT Professional*, *20*(3), 31–39.

Calis, T. (2018). Multi-homing sellers and loyal buyers on darknet markets.

Campbell-Verduyn, M. (2018). Bitcoin, crypto-coins, and global anti-money laundering governance. *Crime, Law and Social Change*, *69*(2), 283–305.

Carnavalet, X. D. C. D., & Mannan, M. (2015). A large-scale evaluation of high-impact password strength meters. *ACM Transactions on Information and System Security (TISSEC)*, *18*(1), 1–32.

Carr, T., Zhuang, J., Sablan, D., LaRue, E., Wu, Y., Al Hasan, M., & Mohler, G. (2019). Into the reverie: Exploration of the dream market. In *2019 ieee international conference on big data (big data)* (pp. 1432–1441). IEEE.

Chainalysis. (2020). *Crypto crime report 2020*. Retrieved August 5, 2020, from https://go.chainalysis.com/rs/503-FAP-074/images/2020-Crypto-Crime-Report.pdf

Christen, P. (2012). *Data matching: Concepts and techniques for record linkage, entity resolution, and duplicate detection*. Springer Science & Business Media.

Christin, N. (2013). Traveling the silk road: A measurement analysis of a large anonymous online marketplace. In *Proceedings of the 22nd international conference on world wide web* (pp. 213–224). ACM.

Collier, D., & Mahoney, J. (1996). Insights and pitfalls: Selection bias in qualitative research. *World Politics*, *49*(1), 56–91.

Collins, L. M., & Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. John Wiley & Sons.

Conlan, K., Baggili, I., & Breitinger, F. (2016). Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy. *Digital investigation*, *18*, S66–S75.

Couvee, K. (2020). Cryptocurrency exchange ordered to freeze bitcoins, identify suspected hackers. Retrieved August 5, 2020, from https://www.moneylaundering.com/news/cryptocurrency-exchange-ordered-to-freeze-bitcoins-identify-suspected-hackers/

Cox, J. (2016). Staying in the shadows: The use of bitcoin and encryption in cryptomarkets. *The Internet and drug markets*, 41–48.

Décary-Hétu, D., & Giommoni, L. (2017). Do police crackdowns disrupt drug cryptomarkets? a longitudinal analysis of the effects of operation onymous. *Crime, Law and Social Change*, *67*(1), 55–75.

Décary-Hétu, D., & Leppänen, A. (2016). Criminals and signals: An assessment of criminal performance in the carding underworld. *Security Journal*, *29*(3), 442–460.

Décary-Hétu, D., Paquet-Clouston, M., & Aldridge, J. (2016). Going international - risk taking by cryptomarket drug vendors. *International Journal of Drug Policy*, *35*, 69–76.

Décary-Hétu, D., & Quessy-Doré, O. (2017). Are repeat buyers in cryptomarkets loyal customers? repeat business between dyads of cryptomarket vendors and users. *American Behavioral Scientist*, *61*(11), 1341–1357.

del Castillo, M. (2015). The 'great bitcoin exodus' has totally changed new york's bitcoin ecosystem. *New York Business Journal*.

Dittus, M., Wright, J., & Graham, M. (2018). Platform criminalism: The'last-mile'geography of the darknet market supply chain. In *Proceedings of the 2018 world wide web conference* (pp. 277–286). International World Wide Web Conferences Steering Committee.

Dixit, A. K. (2011). *Lawlessness and economics: Alternative modes of governance*. Princeton University Press.

Dolliver, D. S., Ericson, S. P., & Love, K. L. (2018). A geographic analysis of drug trafficking patterns on the tor network. *Geographical review*, *108*(1), 45–68.

Doong, H.-S., Wang, H.-C., & Shih, H.-C. (2008). Exploring loyalty intention in the electronic marketplace. *Electronic Markets*, *18*(2), 142–149.

Dourish, P., & Anderson, K. (2006). Collective information practice: Exploring privacy and security as social and cultural phenomena. *Human-computer interaction*, *21*(3), 319–342.

Dread Forums. (2019). Psa: Check the strength of your vendors' pgp key! Retrieved August 5, 2020, from https://dreadytofatroptsdj6io7l3xptbet6onoyno2yv7jicoxknyazubrad.onion.pet/post/b397dbdd5cd37a956578/

Dutch Police. (n.d.). Hack right. Retrieved August 5, 2020, from https://www.politie.nl/themas/hack_right.htm

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in bioinformatics*, *21*(2), 553–565.

Eckstein, H. (2000). Case study and theory in political science. *Case study method*, 119–164.

Ermilov, D., Panov, M., & Yanovich, Y. (2017). Automatic bitcoin address clustering. In *2017 16th ieee international conference on machine learning and applications (icmla)* (pp. 461–466). IEEE.

Eurojust and Europol. (2019). Common challenges in combating cybercrime as identified by eurojust and europol. Retrieved August 5, 2020, from %7Bhttp://www.eurojust.europa.eu/doclibrary/Eurojust-framework/Casework/Joint%20report%20of%20Eurojust%20and%20Europol%20on%20Common%20challenges%20in%20combating%20cybercrime%20(June%202019)/2019-06_Joint-Eurojust-Europol-report_Common-challenges-in-combating-cybercrime_EN.PDF%7D

Europol. (2019). Global law enforcement action against vendors and buyers on the dark web. Retrieved August 5, 2020, from https://www.europol.europa.eu/newsroom/news/global-law-enforcement-action-against-vendors-and-buyers-dark-web

Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *Bmj*, *323*(7321), 1103–1105.

Fanusie, Y., & Robinson, T. (2018). Bitcoin laundering: An analysis of illicit flows into digital currency services. *Center on Sanctions and Illicit Finance memorandum, January*.

FATF. (2015). Fatf guidance for a risk based approach. Retrieved August 5, 2020, from https://www.amlc.nl/wp-content/uploads/2018/07/FATF-Guidance-RBA-Virtual-Currencies-2015.pdf

FATF. (2019a). Fatf guidance. Retrieved August 5, 2020, from http://www.fatf-gafi.org/media/fatf/documents/recommendations/RBA-VA-VASPs.pdf

FATF. (2019b). Fatf recommendations. Retrieved August 5, 2020, from http://www.fatf-gafi.org/media/fatf/documents/recommendations/pdfs/FATF%20Recommendations%202012.pdf

FATF. (2019c). Mitigating risks from virtual assets. Retrieved August 5, 2020, from http://www.fatf-gafi.org/publications/fatfrecommendations/documents/regulation-virtual-assets-interpretive-note.html

Flick, C., & Sandvik, R. A. (2013). Tor and the darknet: Researching the world of hidden services. In *Proceedings of the thirteenth international conference, the possibilities of ethical ict* (pp. 150–157).

Foley, S., Karlsen, J. R., & Putniņš, T. J. (2019). Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies? *The Review of Financial Studies*, *32*(5), 1798–1853.

Furnell, S. (2011). Assessing password guidance and enforcement on leading websites. *Computer Fraud & Security*, *2011*(12), 10–18.

Furnell, S., Esmael, R., Yang, W., Li, N., et al. (2018). Enhancing security behaviour by supporting the user. *Computers & Security*, *75*, 1–9.

Galbally, J., Coisel, I., & Sanchez, I. (2016). A new multimodal approach for password strength estimation—part i: Theory and algorithms. *IEEE Transactions on Information Forensics and Security*, *12*(12), 2829–2844.

Gambetta, D. (2009). Codes of the underworld: How criminals communicate.

Garfinkel, S. (2007). Anti-forensics: Techniques, detection and countermeasures. In *2nd international conference on i-warfare and security* (Vol. 20087, pp. 77–84).

Gastwirth, J. L., Gel, Y. R., & Miao, W. (2009). The impact of levene's test of equality of variances on statistical theory and practice. *Statistical Science*, 343–360.

Gefen, D. (2002). Customer loyalty in e-commerce. *Journal of the association for information systems*, *3*(1), 2.

Gerring, J. (2006). *Case study research: Principles and practices*. Cambridge university press.

Girelli, G. (2019). The death penalty for drug offences. Retrieved August 5, 2020, from https://www.hri.global/files/2019/02/22/HRI_DeathPenaltyReport_2019.pdf

Goldschlag, D., Reed, M., & Syverson, P. (1999). *Onion routing for anonymous and private internet connections*.

Golla, M., Wei, M., Hainline, J., Filipe, L., Dürmuth, M., Redmiles, E., & Ur, B. (2018). "What was that site doing with my facebook password?"designing password-reuse notifications. In *Proceedings of the 2018 acm sigsac conference on computer and communications security* (pp. 1549–1566).

Goodin, D. (2012). 25-gpu cluster cracks every standard windows password in< 6 hours. *Ars Technica*.

Grapperhaus, F. (2018). Kamerbrief over integrale aanpak cybercrime. Retrieved August 5, 2020, from https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/kamerstukken/2018/04/20/tk-integrale-aanpak-cybercrime/tk-integrale-aanpak-cybercrime.pdf

Grapperhaus, F. (2020). Antwoorden kamervragen over het bericht over encryptie. Retrieved August 5, 2020, from https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/kamerstukken/2020/01/24/antwoorden-kamervragen-over-het-bericht-minister-grapperhaus-pleit-bij-europese-unie-voor-achterdeur-in-encryptie/antwoorden-kamervragen-over-het-bericht-minister-grapperhaus-pleit-bij-europese-unie-voor-achterdeur-in-encryptie.pdf

Grapperhaus, V. (2019). From zero to hero: Identifying vendor characteristics that impact vendor performance on darknet markets.

Grassi, P. A., Fenton, J. L., Newton, E., Perlner, R., Regenscheid, A., Burr, W., ... Choong, Y.-Y., et al. (2017). Nist special publication 800-63b: Digital identity guidelines. *Enrollment and Identity Proofing Requirements*.

Greenberg, A. (2018). Operation bayonet: Inside the sting that hijacked an entire dark web drug market. Retrieved August 5, 2020, from https://www.wired.com/story/hansa-dutch-police-sting-operation/

Hammond, J. L. (1973). Two sources of error in ecological correlations. *American Sociological Review*, 764–777.

Hardy, R. A., & Norgaard, J. R. (2016). Reputation in the internet black market: An empirical and theoretical analysis of the deep web. *Journal of Institutional Economics*, *12*(3), 515–539.

Harlev, M. A., Sun Yin, H., Langenheldt, K. C., Mukkamala, R., & Vatrapu, R. (2018). Breaking bad: De-anonymising entity types on the bitcoin blockchain using supervised machine learning. In *Proceedings of the 51st hawaii international conference on system sciences*.

Harris, R. (2006). Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem. *digital investigation*, *3*, 44–49.

Hartel, P., & Van Wegberg, R. (2019). Crime and online anonymous markets. *International and Transnational Crime and Justice*, 67.

Haslhofer, B., Karl, R., & Filtz, E. (2016). O bitcoin where art thou? insight into large-scale transaction graphs. In *Semantics (posters, demos, success)*.

Herley, C., & Florêncio, D. (2010). Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. In *Economics of information security and privacy* (pp. 33–53). Springer.

Hern, A. (2014). Operation onymous may have exposed flaws in tor, developers reveal. *The Guardian*, *11*.

Hetzl, S., & Mutzel, P. (2005). A graph–theoretic approach to steganography. In *Ifip international conference on communications and multimedia security* (pp. 119–128). Springer.

Hinteregger, A., & Haslhofer, B. (2018). An empirical analysis of monero cross-chain traceability.

Ho, T. N., & Ng, W. K. (2016). Application of stylometry to darkweb forum user identification. In *International conference on information and communications security* (pp. 173–183). Springer.

Hoffman, P. (2009). Dsa with sha-2 for dnssec. *Network Working Group - Draft*. Retrieved from https://tools.ietf.org/html/draft-hoffman-dnssec-dsa-sha2-00

Holt, T. J., Smirnova, O., Chua, Y. T., & Copes, H. (2015). Examining the risk reduction strategies of actors in online criminal markets. *Global Crime*, *16*(2), 81–103.

Holt, T. J., Smirnova, O., & Hutchings, A. (2016). Examining signals of trust in criminal markets online. *Journal of Cybersecurity*, *2*(2), 137–145.

Huisman, S., Princen, M., Klerks, P., & Kop, N. (2016). Handelen naar waarheid, sterkte- en zwakteanalyse van de opsporing. Retrieved August 5, 2020, from https://www.rijksoverheid.nl/documenten/rapporten/2016/05/19/tk-bijlage-sterkte-en-zwakteanalyse-van-de-opsporing

Hunt, T. (2018). The legitimisation of have i been pwned. Retrieved August 5, 2020, from https://www.troyhunt.com/the-legitimisation-of-have-i-been-pwned/

Hunt, T. (2020). Have i been pwnd. Retrieved August 5, 2020, from https://www.troyhunt.com/10b/

Ives, B., Walsh, K. R., & Schneider, H. (2004). The domino effect of password reuse. *Communications of the ACM*, *47*(4), 75–78.

James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of applied psychology*, *67*(2), 219.

Janze, C. (2017). Are cryptocurrencies criminals best friends? examining the co-evolution of bitcoin and darknet markets.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, *5*(1), 193–206.

Kankane, S., DiRusso, C., & Buckley, C. (2018). Can we nudge users toward better password management? an initial study. In *Extended abstracts of the 2018 chi conference on human factors in computing systems* (pp. 1–6).

Kessler, G. C. (2007). Anti-forensics and the digital investigator.

Kethineni, S., Cao, Y., & Dodge, C. (2018). Use of bitcoin in darknet markets: Examining facilitative factors on bitcoin-related crimes. *American Journal of Criminal Justice*, *43*(2), 141–157.

Komanduri, S., Shay, R., Kelley, P. G., Mazurek, M. L., Bauer, L., Christin, N., ... Egelman, S. (2011). Of passwords and people: Measuring the effect of password-composition policies. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 2595–2604).

Koops, B.-J. (2016). Megatrends and grand challenges of cybercrime and cyberterrorism policy and research. In *Combatting cybercrime and cyberterrorism* (pp. 3–15). Springer.

Kumar, R., Yadav, S., Daniulaityte, R., Lamy, F., Thirunarayan, K., Lokala, U., & Sheth, A. (2020). Edarkfind: Unsupervised multi-view learning for sybil account detection. In *Proceedings of the web conference 2020* (pp. 1955–1965).

Lenstra, A. K. (2004). Key length. contribution to the handbook of information security.

Lenstra, A. K., & Verheul, E. R. (2001). Selecting cryptographic key sizes. *Journal of cryptology*, *14*(4), 255–293.

Lischke, M., & Fabian, B. (2016). Analyzing the bitcoin network: The first four years. *Future Internet*, *8*(1), 7.

Lusthaus, J. (2012). Trust in the world of cybercrime. *Global crime*, *13*(2), 71–94.

Magidson, J., & Vermunt, J. (2002). Latent class models for clustering: A comparison with k-means. *Canadian Journal of Marketing Research*, *20*(1), 36–43.

Magidson, J., & Vermunt, J. K. (2004). Latent class models. *The Sage handbook of quantitative methodology for the social sciences*, 175–198.

Martin, J., & Christin, N. (2016). Ethics in cryptomarket research. *International Journal of Drug Policy*, *35*, 84–91.

Masyn, K. E. (2013). 25 latent class analysis and finite mixture modeling. *The Oxford handbook of quantitative methods*, 551.

Maxwell, S. E., Delaney, H. D., & Kelley, K. (2017). *Designing experiments and analyzing data: A model comparison perspective*. Routledge.

Meeus, W., van de Schoot, R., Klimstra, T., & Branje, S. (2011). Personality types in adolescence: Change and stability and links with adjustment and relationships: A five-wave longitudinal study. *Developmental psychology*, *47*(4), 1181.

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G. M., & Savage, S. (2013). A fistful of bitcoins: Characterizing payments among men with no names. In *Proceedings of the 2013 conference on internet measurement conference* (pp. 127–140).

Ministry of Justice & Security. (2015). Memorie van toelichting computercriminaliteit iii. Retrieved August 5, 2020, from https://www.bitsoffreedom.nl/wp-content/uploads/memorie-van-toelichting-deel-1-2-computercriminaliteit-iii.pdf

Moore, D., & Rid, T. (2016). Cryptopolitik and the darknet. *Survival*, *58*(1), 7–38.

Moore, T., & Christin, N. (2013). Beware the middleman: Empirical analysis of bitcoin-exchange risk. In *International conference on financial cryptography and data security* (pp. 25–33). Springer.

Moran, M. (2010). Online child abuse material offenders. are we assigning law enforcement expertise appropriately. *Unpublished manuscript. Dublin, Ireland: University College Dublin.*

Moriarty, K., Kaliski, B., Jonsson, J., & Rusch, A. (2016). Pkcs# 1: Rsa cryptography specifications version 2.2. *Internet Engineering Task Force, Request for Comments*, *8017*.

Morselli, C., Giguère, C., & Petit, K. (2007). The efficiency/security trade-off in criminal networks. *Social Networks*, *29*(1), 143–153.

Morselli, C., & Tremblay, P. (2004). Criminal achievement, offender networks and the benefits of low self-control. *Criminology*, *42*(3), 773–804.

Möser, M., Böhme, R., & Breuker, D. (2013). An inquiry into money laundering tools in the bitcoin ecosystem. In *2013 apwg ecrime researchers summit* (pp. 1–14). Ieee.

Mossholder, K. W., & Bedeian, A. G. (1983). Cross-level inference and organizational research: Perspectives on interpretation and application. *Academy of Management Review*, *8*(4), 547–558.

Narayanan, A., & Möser, M. (2017). Obfuscation in bitcoin: Techniques and politics.

National Rapporteur on Trafficking in Human Beings. (2011). Child pornography – first report of the dutch national rapporteur. the hague: Bnrm.

Nemec, M., Sys, M., Svenda, P., Klinec, D., & Matyas, V. (2017). The return of coppersmith's attack: Practical factorization of widely used rsa moduli. In *Proceedings of the 2017 acm sigsac conference on computer and communications security* (pp. 1631–1648).

Notelaers, G., Einarsen, S., De Witte, H., & Vermunt, J. K. (2006). Measuring exposure to bullying at work: The validity and advantages of the latent class cluster approach. *Work & Stress*, *20*(4), 289–302.

Odinot, G., Verhoeven, M., Pool, R., & De Poot, C. (2017). Organised cybercrime in the netherlands: Empirical findings and implications for law enforcement. *Cahiers*.

Openwall. (2020). John the ripper password cracker. Retrieved August 5, 2020, from https://www.openwall.com/john/

Paquet-Clouston, M., Décary-Hétu, D., & Morselli, C. (2018). Assessing market competition and vendors' size and scope on alphabay. *International Journal of Drug Policy*, *54*, 87–98.

Paquet-Clouston, M., Haslhofer, B., & Dupont, B. (2019). Ransomware payments in the bitcoin ecosystem. *Journal of Cybersecurity*, *5*(1), tyz003.

Peron, C. S., & Legary, M. (2005). Digital anti-forensics: Emerging trends in data transformation techniques. In *Proceedings of*.

Pfitzmann, A., & Hansen, M. (2010). A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management.

Pieters, G., & Vivanco, S. (2017). Financial regulations and price inconsistencies across bitcoin markets. *Information Economics and Policy*, *39*, 1–14.

Popper, N. (2015). The tax sleuth who took down a drug lord. New York Times. Retrieved March 5, 2020, from https://www.nytimes.com/2015/12/27/business/dealbook/the-unsung-tax-agent-who-put-a-face-on-the-silk-road.html

Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Univ of California Press.

Ragin, C. C., Becker, H. S. et al. (1992). *What is a case?: Exploring the foundations of social inquiry*. Cambridge university press.

Reid, F., & Harrigan, M. (2013). An analysis of anonymity in the bitcoin system. In *Security and privacy in social networks* (pp. 197–223). Springer.

Revier, K. (2019). *Policing pain: Opioids, crisis, and a shifting war on drugs* (Doctoral dissertation, State University of New York at Binghamton).

Reynolds, P., & Irwin, A. S. (2017). Tracking digital footprints: Anonymity within the bitcoin system. *Journal of Money Laundering Control*.

Rhumorbarbe, D., Staehli, L., Broséus, J., Rossy, Q., & Esseiva, P. (2016). Buying drugs on a darknet market: A better deal? studying the online illicit drug market through the analysis of digital, physical and chemical data. *Forensic science international*, *267*, 173–182.

Rogers, M. (2006). Anti-forensics: The coming wave in digital forensics. *Retrieved September*, *7*, 2008.

Ruoti, S., Andersen, J., Zappala, D., & Seamons, K. (2015). Why johnny still, still can't encrypt: Evaluating the usability of a modern pgp client.

Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of risk and uncertainty*, *1*(1), 7–59.

Sawilowsky, S. S., & Blair, R. C. (1992). A more realistic look at the robustness and type ii error properties of the t test to departures from population normality. *Psychological bulletin*, *111*(2), 352.

Seawright, J., & Gerring, J. (2008). Case selection techniques in case study research: A menu of qualitative and quantitative options. *Political research quarterly*, *61*(2), 294–308.

Shay, R., Bauer, L., Christin, N., Cranor, L. F., Forget, A., Komanduri, S., ... Ur, B. (2015). A spoonful of sugar? the impact of guidance and feedback on password-creation behavior. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 2903–2912).

Shay, R., Komanduri, S., Kelley, P. G., Leon, P. G., Mazurek, M. L., Bauer, L., ... Cranor, L. F. (2010). Encountering stronger password requirements: User attitudes and behaviors. In *Proceedings of the sixth symposium on usable privacy and security* (pp. 1–20).

Soska, K., & Christin, N. (2015). Measuring the longitudinal evolution of the online anonymous marketplace ecosystem. In *24th usenix security symposium* (pp. 33–48).

Spitters, M., Klaver, F., Koot, G., & van Staalduinen, M. (2015). Authorship analysis on dark marketplace forums. In *2015 european intelligence and security informatics conference* (pp. 1–8). IEEE.

Sremack, J. C., & Antonov, A. V. (2007). Taxonomy of anti-computer forensics threats. *IMF*, *103*, e12.

Stanton, J. M., Stam, K. R., Mastrangelo, P., & Jolton, J. (2005). Analysis of end user security behaviors. *Computers & security*, *24*(2), 124–133.

Stinenbosch, B. (2019). *Measuring darknet markets* (Master's thesis, TU Delft).

Sun Yin, H., & Vatrapu, R. (2017). A first estimation of the proportion of cybercriminal entities in the bitcoin ecosystem using supervised machine learning. In *2017 ieee international conference on big data (big data)* (pp. 3690–3699).

Sundaresan, S., McCoy, D., Afroz, S., & Paxson, V. (2016). Profiling underground merchants based on network behavior. In *2016 apwg symposium on electronic crime research (ecrime)* (pp. 1–9). IEEE.

Tai, X. H., Soska, K., & Christin, N. (2019). Adversarial matching of dark net market vendor accounts. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining* (pp. 1871–1880).

Thomas, D. R., Pastrana, S., Hutchings, A., Clayton, R., & Beresford, A. R. (2017). Ethical issues in research using datasets of illicit origin. In *Proceedings of the 2017 internet measurement conference* (pp. 445–462).

Tzanetakis, M., Kamphausen, G., Werse, B., & von Laufenberg, R. (2016). The transparency paradox. building trust, resolving disputes and optimising logistics on conventional and online drugs markets. *International Journal of Drug Policy*, *35*, 58–68.

Tziakouris, G. (2018). Cryptocurrencies—a forensic challenge or opportunity for law enforcement? an interpol perspective. *IEEE Security & Privacy*, *16*(4), 92–94.

Van Buskirk, J., Bruno, R., Dobbins, T., Breen, C., Burns, L., Naicker, S., & Roxburgh, A. (2017). The recovery of online drug markets following law enforcement and other disruptions. *Drug and alcohol dependence*, *173*, 159–162.

Van Buskirk, J., Naicker, S., Bruno, R., Breen, C., & Roxburgh, A. (2016). Drugs and the internet. *DrugsDrugs*.

Van Buskirk, J., Naicker, S., Roxburgh, A., Bruno, R., & Burns, L. (2016). Who sells what? country specific differences in substance availability on the agora cryptomarket. *International Journal of Drug Policy*, *35*, 16–23.

Van de Sandt, E. (2019). *Deviant security: The technical computer security practices of cyber criminals* (Doctoral dissertation, University of Bristol).

Van de Steur, A. (2016). Naar een effectieve en toekomstbestendige opsporing, voortangsnota. Retrieved August 5, 2020, from https://www:rijksoverheid:nl/documenten/rapporten/2016/06/20/tk-bijlage-3a-voortgangsnota-versterking-opsporing-juni-2016

Van Hardeveld, G. J. (2018). *Deviating from the cybercriminal script: Exploring the contextual factors and cognitive biases involved in carding* (Doctoral dissertation, University of Southampton).

Van Hardeveld, G. J., Webber, C., & O'Hara, K. (2017). Deviating from the cybercriminal script: Exploring tools of anonymity (mis) used by carders on cryptomarkets. *American Behavioral Scientist*, *61*(11), 1244–1266.

Van Hout, M. C., & Bingham, T. (2014). Responsible vendors, intelligent consumers: Silk road, the online revolution in drug trading. *International Journal of Drug Policy*, *25*(2), 183–189.

Van Valkenburgh, P. (2015). Tracking bitcoin regulation state by state. CoinCentre.

Van Wegberg, R., Miedema, F., Akyazi, U., Noroozian, A., Klievink, B., & van Eeten, M. (2020). Go see a specialist? predicting cybercrime sales on online anonymous markets from vendor and product characteristics, 816–826.

Van Wegberg, R., Tajalizadehkhoob, S., Soska, K., Akyazi, U., Ganan, C. H., Klievink, B., ... van Eeten, M. (2018). Plug and prey? measuring the commoditization of cybercrime via online anonymous markets. In *27th usenixsecurity symposium (usenix security 18)* (pp. 1009–1026).

Van Wegberg, R., & Verburgh, T. (2018). Lost in the dream? measuring the effects of operation bayonet on vendors migrating to dream market. In *Proceedings of the evolution of the darknet workshop* (pp. 1–5).

Van Wegberg, R., Oerlemans, J., Deventer, O., et al. (2018). Bitcoin money laundering: Mixed results? an explorative study on money laundering of cybercrime proceeds using bitcoin. *Journal of Financial Crime*, *25*, 17.

Van Wijk, A., & Scholten, L. (2006). Onbenutte kansen.

Vargha, A., & Delaney, H. D. (1998). The kruskal-wallis test and stochastic homogeneity. *Journal of Educational and behavioral Statistics*, *23*(2), 170–192.

Veitch, W. R., & John T, R. (1974). Homogeneity of variance: An empirical comparison of 4 statistical tests. *The Journal of Experimental Education*, *43*(2), 73–78.

Vermunt, J. K., & Magidson, J. (2013). Technical guide for latent gold 5.0: Basic, advanced, and syntax. *Belmont, MA: Statistical Innovations Inc.*

Wang, D., Zhang, Z., Wang, P., Yan, J., & Huang, X. (2016). Targeted online password guessing: An underestimated threat. In *Proceedings of the 2016 acm sigsac conference on computer and communications security* (pp. 1242–1254).

Wang, X., Peng, P., Wang, C., & Wang, G. (2018). You are your photographs: Detecting multiple identities of vendors in the darknet marketplaces. In *Proceedings of the 2018 on asia conference on computer and communications security* (pp. 431–442).

Wash, R., Rader, E., Berman, R., & Wellmer, Z. (2016). Understanding password choices: How frequently entered passwords are re-used across websites. In *Twelfth symposium on usable privacy and security ({soups} 2016)* (pp. 175–188).

Wehinger, F. (2011). The dark net: Self-regulation dynamics of illegal online markets for identities and related services. In *2011 european intelligence and security informatics conference* (pp. 209–213). IEEE.

Westfall, P. H., Tobias, R. D., & Wolfinger, R. D. (2011). *Multiple comparisons and multiple tests using sas*. SAS Institute.

Wheeler, D. A., & Larsen, G. N. (2003). *Techniques for cyber attack attribution*.

Whitten, A., & Tygar, J. D. (1999). Why johnny can't encrypt: A usability evaluation of pgp 5.0. In *Usenix security symposium* (Vol. 348, pp. 169–184).

Yin, R. K. (1993). *Case study research: Design and methods*.

Zhang, Y., Xiao, Y., Ghaboosi, K., Zhang, J., & Deng, H. (2012). A survey of cyber crimes. *Security and Communication Networks*, *5*(4), 422–437.

Zhou, G., Zhuge, J., Fan, Y., Du, K., & Lu, S. (2020). A market in dream: The rapid development of anonymous cybercrime. *Mobile Networks and Applications*, *25*(1), 259–270.

A

Vendor Results

A.1. Accuracy of revenue

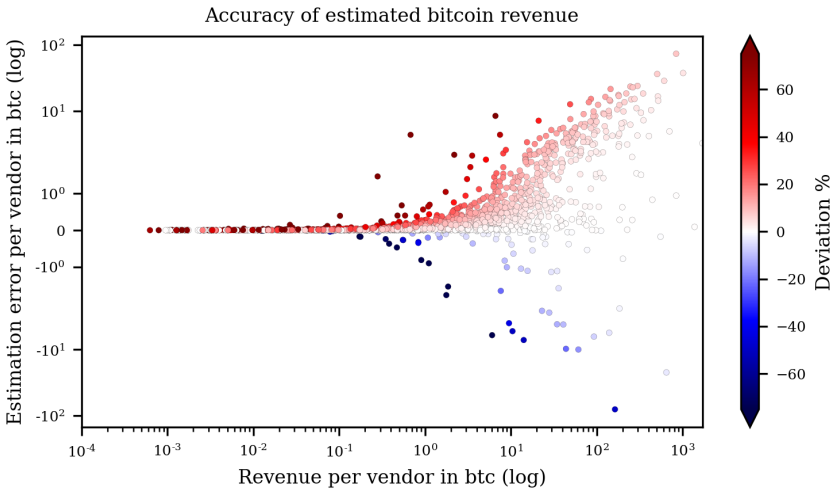


Figure A.1: Estimated accuracy of revenue accuracy

A.2. BVR statistics

BVRs statistics are sensitive to large sample sizes. Scholars therefore work with a reduction of the BVRs relative to the 1-cluster model. A rule of thumb is to require a reduction of 90% instead of striving for non-significance of the BVRs (Notelaers et al., 2006). Representing the BVRs as relative reductions is insightful in understanding the global fit as well. From Table A.1 it becomes clear that difficulties arise in reducing the local independence of all import-related variable pairs. This makes sense, because vendors with ample experience mostly are not active on other markets. Vendors that are active on other markets might have an established customer base resulting in more sales. To leave *import* out of the clustering model does result in better clustering statistics, but this goes beyond the goal of grouping similar vendors. It is therefore decided to keep import and to accept the local dependencies. When interpreting the correlation between clusters and security differences, it is of importance not to attribute all differences to experience, physical or digital sales but consider these jointly with import.

To make sure each vendor characteristic adds significant value to the clustering result, Wald statistics are calculated. The null hypothesis is that all parameters are equal to zero. The

Table A.1: Bivariate residuals of the 5-cluster models, including relative decrease in percentages compared to 1-cluster model.

	5-cluster			5-cluster, decrease in %		
	experience	physical	digital	experience	physical	digital
experience						
physical	556	.		247.4%		
digital	35	1	.	4318.0%	7517.2%	
import	2318	497	454	10.5%	86.7%	37.8%

alternative hypothesis stipulates that at least one parameter is non-zero. For all vendor characteristics $p < 0.05$ (Table A.4). Therefore, it is safe to conclude that there is a relation between each vendor characteristic and all latent profiles. This means that all vendor characteristics contribute significantly to the clustering model and thus should be retained. Positive indicator values indicate a positive influence on the probability that a vendor belongs to a certain cluster, while negative values decrease this possibility. Inspecting these values already gives a first idea on the type of clusters formed.

A.3. Assessing the 6-clusters model

Evidently, a good local fit for especially the digital and physical sale variables is important. The 6-cluster model seems to be a better choice than the 5-cluster model. Here, an additional variable pair has non-significant residuals and the sum of all BVRs is remarkably low. However, further inspection of the 6-cluster model shows that it does not provide a more relevant separation of vendor types. The 6-cluster model seperates vendors with a lot of physical sales into two clusters ('very high selling' and 'extremely high selling' vendors). Remarkably, this results in the Digital-Experience pair becoming locally independent. The extra granularity in the top selling segment of physical sales is not desired, so the less complex model is preferred. As indicated by the sample sizes per cluster, the 5-model cluster includes only one 'small' cluster, whereas the 6-model cluster includes two.

A.4. Significance of differences between clusters

It is important that the clusters differ significantly in terms of the vendor characteristics. Kruskal-Willis H tests are performed to test this. This test is used to asses whether the medians or distributions of each vendor characteristic differ for any of the created clusters. A non-parametric test is performed because none of the vendor characteristics are normally distributed within clusters. This remains to be the case after log-transforming the data. Because the data is rather imbalanced, violations of the normality assumption should not be ignored.

First, it is regarded whether the medians or distributions should be compared. The distributions are plotted in Figure 5.5. There is no generally accepted procedure for determining to what extent distributions are similar (Vargha & Delaney, 1998). It is however clear that in the case of physical and digital, distributions are not similar at all. The spread of the whiskers, the proportion of upper and lower quartile differ clearly among groups. In these cases, via Kruskal Willis will be tested whether there are differences in distributions, scores or mean ranks. For experience, it is concluded that the distributions are somewhat similar. The medians are placed more or less evenly between the upper and first quartiles and few outliers are present for each group. Therefore, for experience the differences in medians can be judged.

Table A.2: Pairwise Kruskal Willis H tests, per cluster and vendor characteristic

	Vendor Characteristic											
	experience				physical				digital			
	p-value per cluster pair				*p*-value per cluster pair				*p*-value per cluster pair			
	c1	c2	c3	c4	c1	c2	c3	c4	c1	c2	c3	c4
c1	.				.				.			
c2	0.0000*	.			0.0000*	.			0.0000*	.		
c3	0.0000*	0.0000*	.		0.0000*	0.0000*	.		0.2490	0.0744	.	
c4	0.0000*	0.2229	0.0113*	.	0.0000*	0.0000*	0.0000*	.	0.0000*	0.0000*	0.0000*	.
c5	0.0000*	0.0000*	0.0028*	0.0002*	0.0000*	0.0000*	0.0000*	0.8199	0.0000*	0.0000*	0.0000*	0.0000*

* significant with $\alpha = 0.05$

The Kruskal-Wallis H test was run to determine if there were differences in experience between the five generated clusters of vendors. Since the distributions of experience were similar throughout the clusters, the test indicates to what extend the *medians* of experience are significantly different between groups. By testing the clusters pairwise, it becomes clear that except for the c2-c4 pair, all pairs of clusters have significantly different medians (see Table A.2. This indicates successful clustering. As would show later, the non-significance of the c2-c4 pair is not problematic, since c2 describes vendors selling mainly physical items and c4 mostly digital ones.

Table A.3: Mean ranks per cluster for physical and digital

	Clusters				
Vendor characteristic	c1	c2	c3	c4	c5
physical	622.4	1296.1	1678.4	347.1	324.3
digital	835.8	744.5	795.9	1656.8	1722.0

For physical and digital, the Kruskal-Wallis H tests are also pairwisely performed. For these vendor characteristics, the distributions do differ between groups. Therefore the interpretation of the test results is slightly different. It is tested whether the *mean ranks* significantly differ instead of the medians. The mean ranks are presented in Table A.3. Regarding the vendor characteristic physical, all but one mean ranks significantly differ (A.3. The digital-oriented clusters c4 and c5 do not statistically differ, which is not surprising and does not harm interpretation of the clusters. The same trend is visible in the results of digital: all clusters significantly differ, with the exception of the physical oriented clusters c1, c2 and c3.

Table A.4: Contribution of vendor characteristics to clusters

	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5	Wald	p-value
experience	-1.69	0.22	0.45	0.31	0.71	117944.86	0.000
physical	-1.07	1.45	3.67	-1.86	-2.21	344824.98	0.000
digital	-2.40	-2.90	-2.34	2.51	5.12	260740.81	0.000
import							
FALSE	0.50	-0.10	-0.33	0.14	-0.21	142.52	0.000
TRUE	-0.50	0.10	0.33	-0.14	0.21		

B

Authentication security

B.1. Password Length

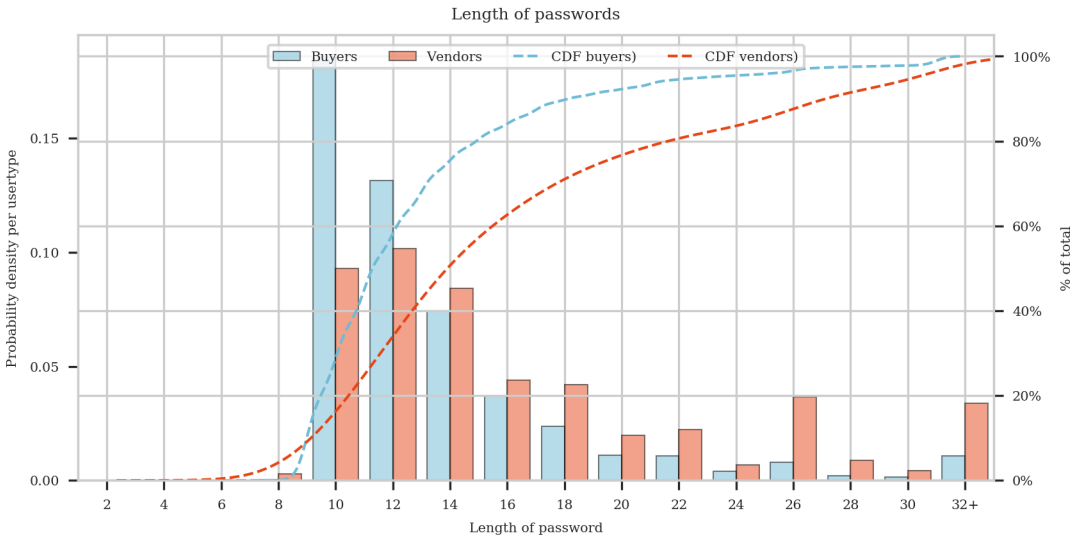


Figure B.1: Probability density and cumulative distribution of password length per user type.

B.2. Assessment of missing password data

About 38% of the vendor population did not log into his or her account in the last month the market was active. This introduces a bias in the available password data. This bias is visualised in Figure B.2. Relatively inexperienced or unsuccessful vendors are more likely to have no password data present compared to their more successful counterparts. Possibly, these vendors haven't logged in because they stopped trading or lost interest in the market. Fortunately, the less successful vendors are numerous, so the clusters of vendors remain well populated. In this part of the analysis, there are 493 `Novices` (-50.1%), 394 `Drug Dealers` (-22.6%), 93 `Drug Lords` (-15.5%), 78 `Digital Fraudsters` (-24.3%) and as much as `Cybercrime Elites` as before ($n = 23$).

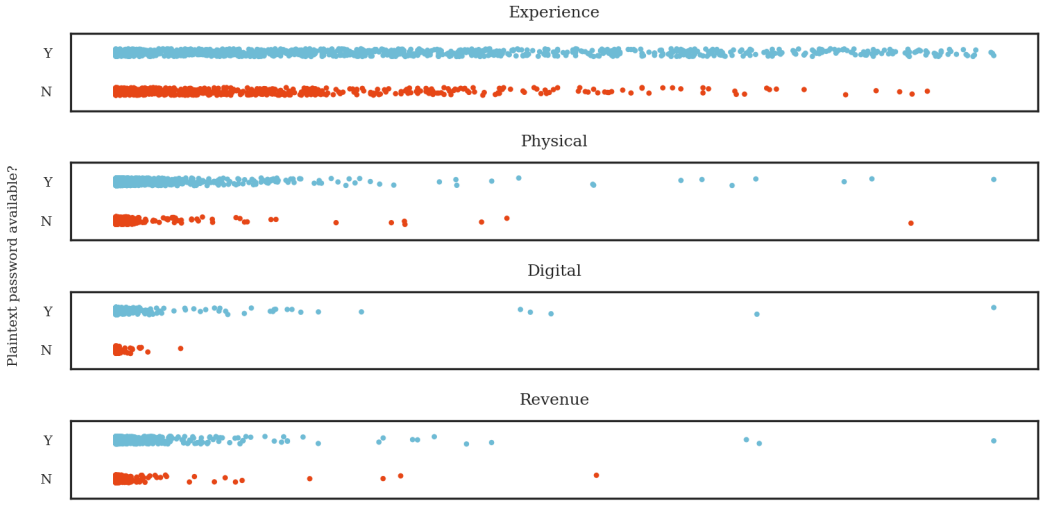


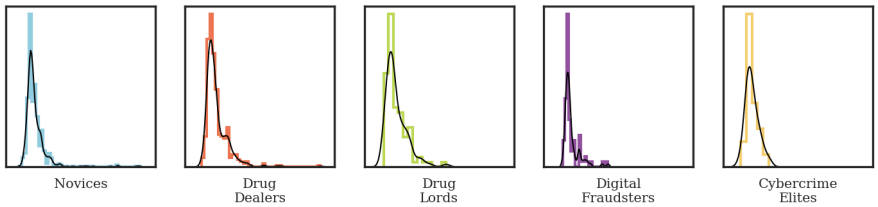
Figure B.2: Missing plaintext passwords of vendors plotted per vendor characteristic

B.3. Assumption testing ANOVA

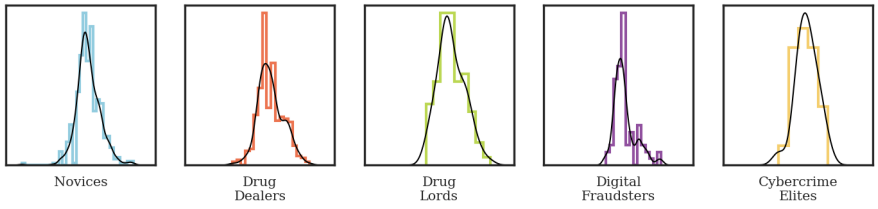
An ANOVA-test should only be performed when certain assumptions are met.

B.3.1. Normality assumption

Firstly, the data should be approximately normally distributed within each vendor profile. The normalised histograms and estimated density functions are plotted in Figure B.3a. At first glance, the distributions follow an approximately normal curve. However, most groups are slightly skewed to the right and some inconsistencies are present at the right tails of the distributions. These issues are fixed by log-transforming the data and removing any outliers that have a standardised value of $z > 5$ (Figure B.3b). This means that only the 'most extreme' outliers are removed. Ultimately, 488 `Novices` (-5), 391 `Drug Dealers` (-3), 93 `Drug Lords` (-0), 78 `Digital Fraudsters` (-0) and 23 `Cybercrime Elites` (-0) remain in the data.



(a) Distribution of password complexity per vendor profile, normalised with fitted estimated density function



(b) Distribution of log-transformed password complexity per vendor profile, normalised with fitted estimated density function. After removing outliers $z > 5$.

Figure B.3: Distribution of password complexity per vendor profile, before and after transforming the data

The normality assumption is further assessed by plotting the theoretical quantities against the sample quantities in a QQ-plot (Figure B.4). The closer the data points are to the dotted line,

the better a normal distribution fits the data. The small deviations from normality (as observed in the right tails) are not problematic since a) they are rather small, b) ANOVA is fairly robust[1] to deviations from normality (Maxwell, Delaney, & Kelley, 2017) and c) the groups tend to be similarly skewed (Sawilowsky & Blair, 1992).

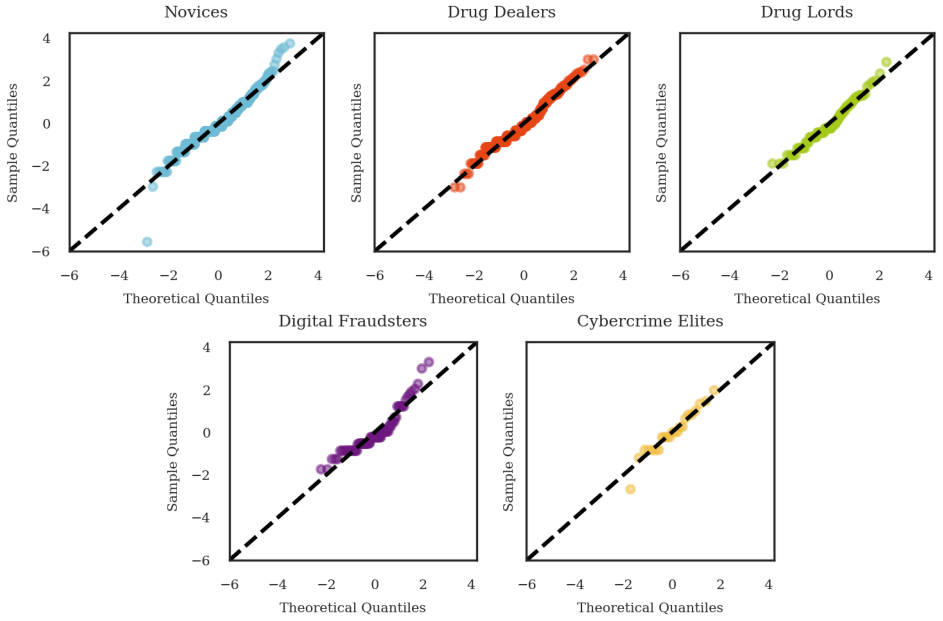


Figure B.4: QQ-plot to assess whether the data is approximately normally distributed

B.3.2. Homogeneity of Variances assumption

Secondly, the homogeneity of variances is regarded through Bartlett's Test. Because the data is normally distributed and the sample size of `Novices` is moderately high, Bartlett's test is preferred over Levene's Test of Homogeneity. Levene's Test depends on sample size and produces lower p-values when n increases (Gastwirth, Gel, & Miao, 2009). Furthermore, Bartlett's Test is suited for unbalanced designs (Veitch & John T, 1974). From Bartlett's Test ($\chi^2 = 5.63, p = 0.2285$) it was concluded that there is homogeneity of variances among all vendor profiles.

The ANOVA test is significant ($F(4, 1068) = 5.89, p = 0.0001$). This means that between the vendor profiles the password complexity differs significantly.

B.4. PWND Matching

Most passwords that are not matched tend to be more complex than reused passwords (Figure B.5). Likewise, rarely used passwords tend to be stronger than often reused passwords.

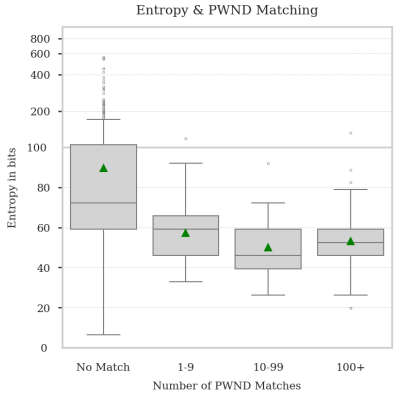


Figure B.5: Password complexity per number of matches

[1] w.r.t. Type I errors

C

Encryption of Communication

C.1. Peculiar key sizes

Occurrences of PGP-keys of length 1023, 2047 or 4095 bits are found in the Hansa database. These aberrant key sizes are the result of how RSA keys of length N are generated. The length N refers to the modulus used by the RSA algorithm. N is created by multiplying two randomly chosen primes $p \cdot q$ of length $N/2$. In most cases, this will result in a key of length N. It can happen by chance, that a key of $N - 1$ is generated. Some implementations of RSA correct for this, but it is not mandatory as per the RSA specification (Moriarty, Kaliski, Jonsson, & Rusch, 2016). Most importantly, the atypical keys do not have significantly lower cryptographic strengths as shown by Nemec, Sys, Svenda, Klinec, and Matyas (2017).

C.2. Algorithms

The most-used encryption algorithm to be used with the PGP protocol is RSA. About 2% of the vendors' and 1.6% of the buyers' keys is generated to be used in conjunction with DSA (for signing) and Elgamal (for encryption) algorithms. Without going into the advantages and disadvantages of each algorithm, it is important to note that DSA, Elgamal and RSA have about the same cryptographic strength at equal key sizes (Barker & Roginsky, 2010; Hoffman, 2009). While it is unusual to generate a PGP-key with a different algorithm then RSA, this does not equal to non-secure behaviour per se.

C.3. Email addresses

The names and email addresses included in the PGP-keys may or may not be fake. While exploring these data, names could be derived from some email addresses. It is however infeasible to check whether these names actually do belong to the entities as represented by the usernames. So no solid conclusions w.r.t. security behaviour can be made. The same goes for email providers listed within the PGP keys. While some vendors' email addresses end with 'gmail.com' or 'fake.com' others register their PGP-key to a dark web email address. For vendors the most popular services are (1) Sigaint, (2) safe-mail, (3) mail2tor, (4) Protonmail and (5) Torbox. Unfortunately it is again infeasible to check whether these are valid email addresses, thus no security related conclusions can be drawn.

D

Choice of OFSP

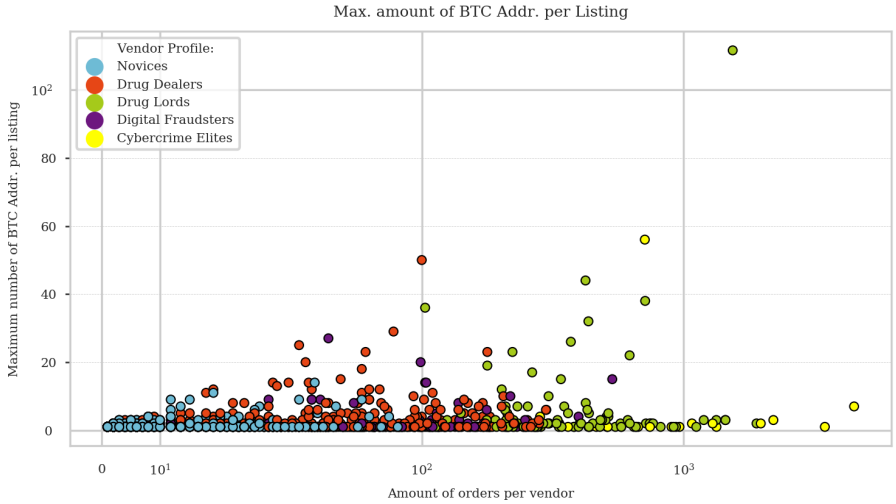


Figure D.1: Maximum number of bitcoin addresses used for a single listing. Per vendor and plotted over the number of orders a vendor has.

Figure D.1 shows how often vendors change the payout addresses of listings. Most vendors do not change the payout addresses of their listings. Some have listings of which the payout addresses are changed numerous times.

D.1. Directly transacting to exchanges

Figure D.2 shows at what points in their Hansa careers vendors transact directly to exchanges. The dates on which these transactions took place are aggregated over weeks. If one or more payouts have been transacted directly to an exchange in a certain week, this week is marked red. Weeks in which a vendor has sales of which the payouts have not been directly transacted to exchanges, are marked blue.

While it was expected that most vendors would make this mistake early in their career, this is not the case. To be precise, 29 out of 104 vendors create this security risk in their first week of sales. Only 5 vendors make this mistake in their first two weeks and later never again. As shown, most vendors repeatedly transact directly to exchanges. It seems that vendors have to specify an unique payout address for each listing, since no payout addresses are found that are used for more than one listing. The pattern that is observed in figure D.2 might be due to a certain

listing being sold repeatedly. On the other hand, the listing entries in the Hansa database do not contain any bitcoin addresses. The bitcoin payout addresses are specified in the individual orders. This suggests that vendors specify their payout preferences upon accepting an order from a buyer.

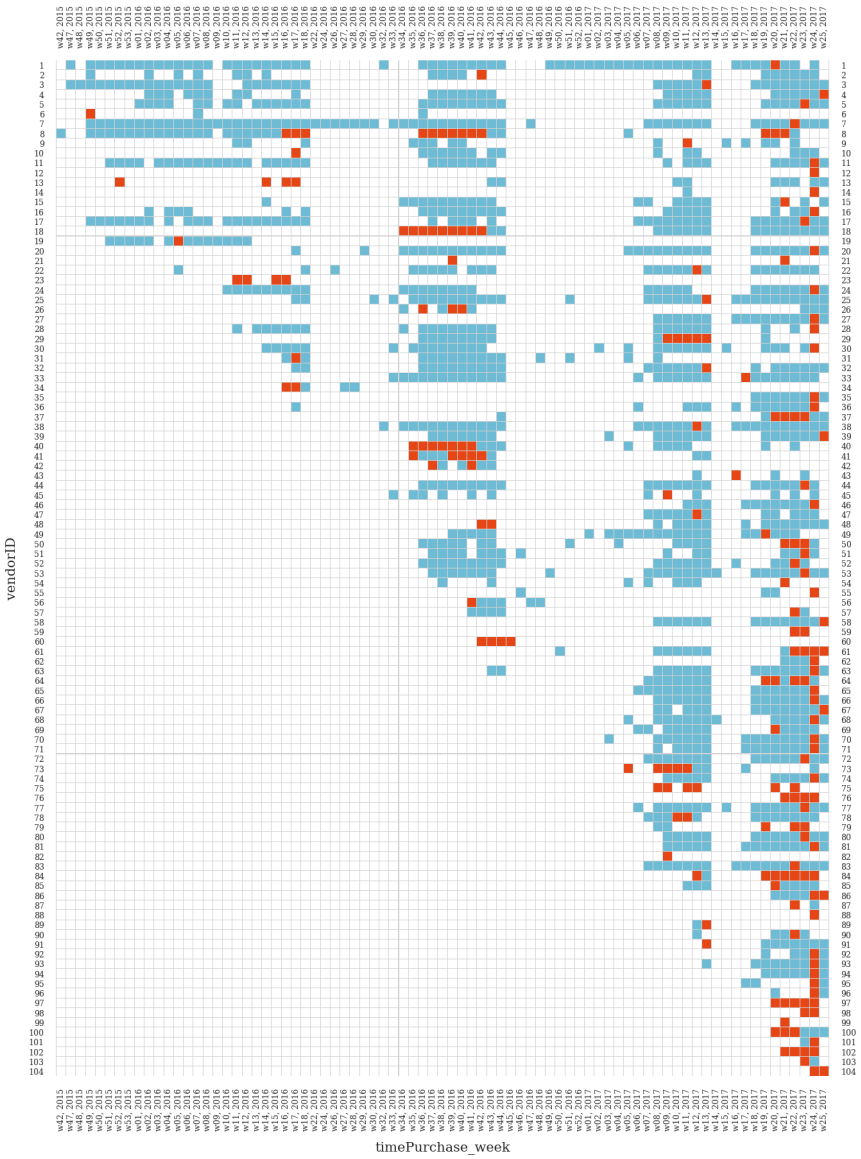


Figure D.2: Direct transactions to exchanges, per vendor per week

D.2. Vendor Characteristics & Choice of OFSP

In this section the vendor characteristics are visually compared with the results of the OFSP-analyses. Not all payout addresses are retrieved. The figures therefore might not sketch a complete picture of the transactional behaviour. For example, if a vendor has been mostly active in the months in which the data is missing, the vendor is less likely to appear in these figures. Vendors with little sales and experience are expected to have less links with OFSPs because these venors transact less. The individual data points are plotted. Therefore, for each OFSP it can be observed whether they are used by (un)successful or (in)experienced vendors. Lastly, as described in section 4.2.2.4, it is not regarded whether a vendor has transacted one, or multiple

times to an OFSP. From a security perspective, it makes sense to only analyse whether a link could be established between the vendor and an OFSP. It does not matter how often or in what phase of the cybercriminals' career.

Figure D.3 shows that both the experienced as non-experienced could be linked with any of the OFSPs. From all vendor characteristics, *experience* is the most difficult characteristic to draw conclusions from. Regarding a link with mixers, the figure shows that given an established link with a mixer, relatively many experienced vendors are observed.

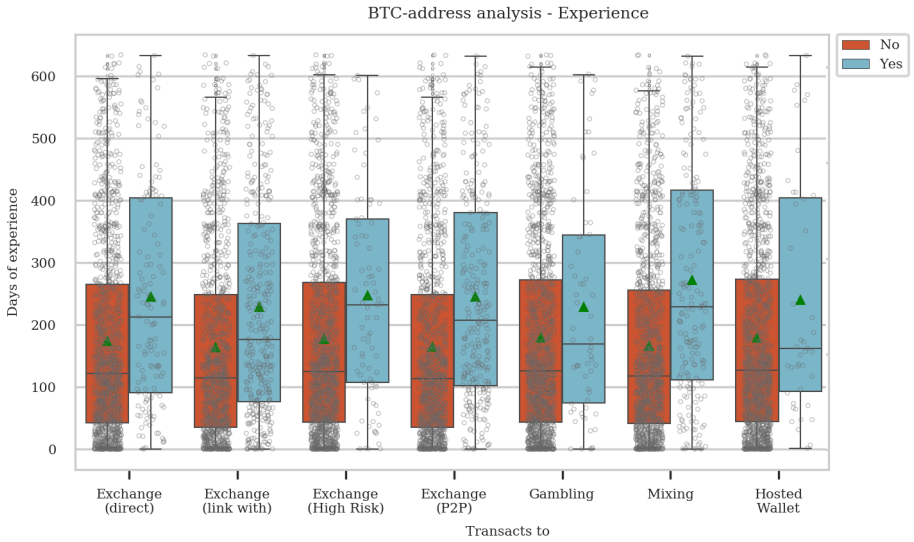


Figure D.3: Experience

Figure D.4 shows that half of the vendors that could be directly linked with exchanges, have more than approximately 80 physical sales. Quite a few vendors with hundreds of sales, transacted directly to an exchange during their Hansa career (although, many more did not). Regarding transacting to a central exchange, either directly or indirectly via a private wallet, larger number of high selling vendors could be linked. Regarding digital sales, similar conclusions are drawn (Figure D.5).

Not that many vendors transact to high risk exchanges, but 50% of those that do, have more than about a hundred physical sales. High risk exchanges are not popular among vendors that sell many digital goods. More vendors are observed to make use of P2P-exchanges. Among both physically and digitally focused vendors, these exchanges seem to be popular with vendors that have many sales. It is easy to understand how vendors that are part of physically oriented supply chains of drugs, can use cash money of the P2P exchanges for buying drugs of their suppliers. It is therefore surprising that also high selling digital vendors make use of P2P exchanges. It was expected that mixing would occur more often among the high selling vendors.

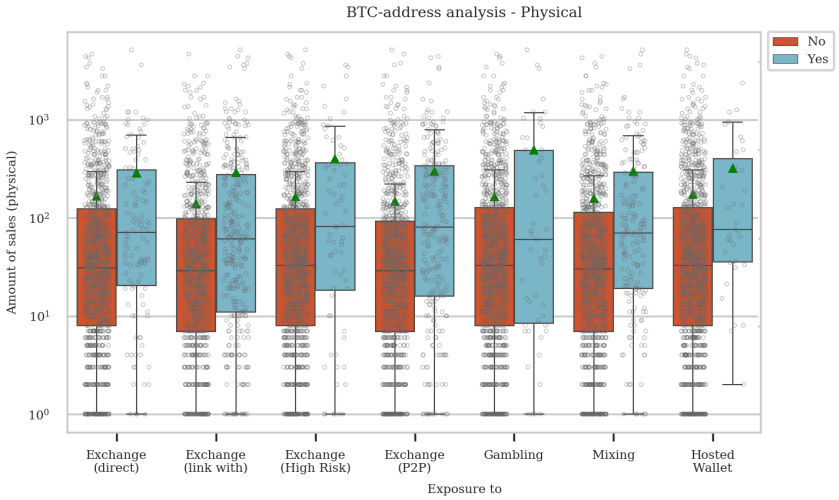


Figure D.4: Physical Sales

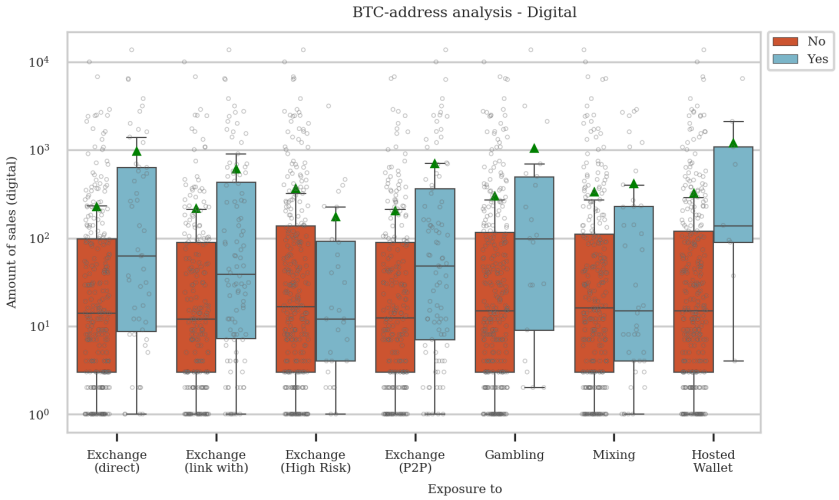


Figure D.5: Digital Sales

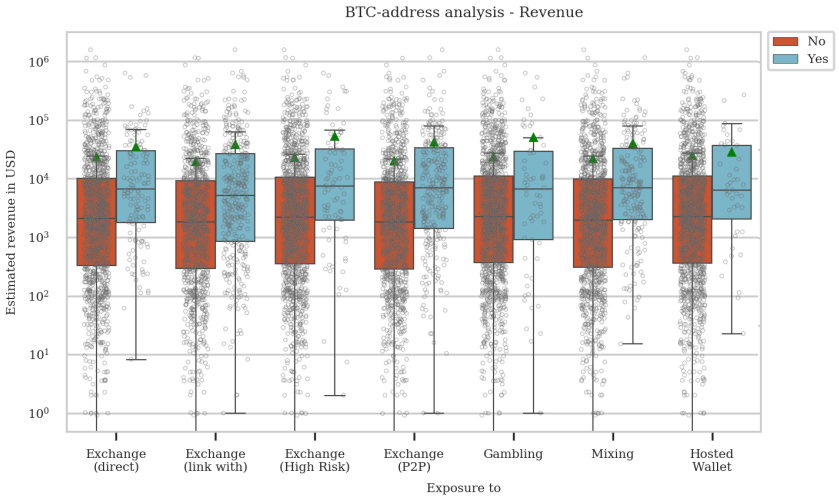


Figure D.6: Revenue

Always fully secure?

Analysing the security behaviour of cybercriminals active on the Hansa darknet market

J.W. van de Laarschot
September 2020
Master Thesis - Scientific Article
Delft University of Technology

ABSTRACT

Darknet market users trouble attribution efforts of Law Enforcement Agencies by investing in additional security mechanisms. These decrease the amount of useful evidence available to Law Enforcement Agencies. In the full administration of the Hansa Market, acquired in Operation Bayonet and originating from the server that hosted the market, observations are made on *a)* what types of vendors can be differentiated, in terms of their experience gained, the amount of sales made and the type of goods sold, *b)* what kind of security behaviour vendors exhibit. For each vendor, we measure what kind of encryption they use, how strong their password is, whether they use different PGP-keys when active on other markets and whether they can be easily linked to a central bitcoin exchange. We show that vendors that are active for a longer period of time and with relatively many drug related sales show the most secure behaviour. Vendors specialised in transacting digital items, have bad security practices most often. However, within every vendor type, there are vendors with poor security practices observed.

1 INTRODUCTION

Cybercrime is one of the biggest challenges of law enforcement agencies [1]. Regarding the prosecution of cybercrime in The Netherlands, Van de Sandt [2] even speaks of an *effectiveness crisis*. Online anonymous marketplaces (or *darknet markets*) are prominently placed in today's cybercrime ecosystem [3, p. 67]. Currently, the yearly estimated revenue of all darknet markets combined is more than $790 million worth of cryptocurrencies [4].

Roughly two types of products are sold on darknet markets: products that have to be shipped physically and items that can be transacted digitally. The physical goods are mostly drugs, e.g. cocaine, cannabis, heroin or other psychoactive substances [5]. Digital items found on darknet markets are e.g. credit card details, hacked accounts (for example, PayPal, adult websites and streaming services such as Netflix and Spotify), databases of names and e-mail addresses and money mule recruitment ads [6].

In the past, Law Enforcement Agencies (LEA) have successfully shut down darknet markets in large scale operations. According to Bradley [7, pp. 228-230], these internationally coordinated operations impact the capability of cybercriminals to trade in the darknet market ecosystem. In contrast, recent analyses of cryptocurrency transactions show that, despite law enforcement scrutiny, the transaction volume and the number of active markets in the ecosystem have an upward trend since 2014 [4]. Besides, the number of convictions remain relatively low [7].

Darknet markets feature security mechanisms that create an attribution problem, i.e. linking a cybercriminal or its machine to an identity or location [8]. Examples of these features are the use of cryptocurrencies for payment, reputation mechanisms, security policies, content moderation, user banning or auto-encryption functionalities [9, 10, 11, 12, 13, 7]. However, this does not suffice, as darknet market users have been arrested [14, 15]. Thus, cybercriminals also have to invest in security themselves. In this research, it is regarded that 'security of cybercriminals' is created through compromising the availability or usefulness of evidence to the forensic process [16].

Cybercriminals do not always achieve maximum security. Van de Sandt [2] argues that 'perfect security' is not economically viable for cybercriminals. Darknet market users have to invest in maintaining their knowledge, skills, equipment and secure work routines. These practices are extensively discussed on darknet discussion fora [7]. From an micro-economical point of view, it could be stipulated that parallelly to the legitimate online world, security in the illegitimate world comes at a cost [*cf.* 17]. Using a behavioural economics perspective, Van Hardeveld [18] argues that cybercriminals are subjected to behavioural biasses that impede their security.

This research aims to explore the security practices of vendors active on a darknet market at scale. This is done by analysing the full administration of the Hansa Market (2015-2017). The Hansa Market administration was acquired in Operation Bayonet and originates from the server that hosted the market. It was made available for research by the Fiscal Information and Investigation Service (FIOD). Based on this data, observations are made on *a)* what types of vendors can be differentiated, in terms of their experience gained, the amount of sales made and

the type of goods sold, *b)* what kind of security behaviour vendors exhibit. These two insights are combined to answer the research question: *"Which factors influence the security behaviour of darknet maket vendors on Hansa Market?"*.

First, through literature review security behaviours that potentially can be observed on darknet markets are identified (section 2). In this review it is also shortly discussed which distinguishing characteristics of vendors are expected to have a relation to security behaviour. Then, five behaviours are selected that are clearly beneficial or detrimental to the cybercriminals' security. The conceptual model summarises how the security behaviours are related to the vendor characteristics by assigning 'vendor types' to the vendors (section 3). The methodology (section 4) elaborates on how vendors are clustered into vendor types, based on their characteristics and on how each security behaviour is measured in the Hansa data. Additionally, a scoring method is presented which scores each vendor by considering the security behaviours jointly. This produces an estimation of vendors' overall security. The results of the analyses are presented in section 5. While research on cybercriminals' security practices is scarce, a number of academic works are related to this research. How these relate to this research is discussed in section 7. Lastly, conclusions are presented in section 8 and are further discussed in section **??**.

2 THEORETICAL BACKGROUND

In this thesis, security is defined from a digital forensics' point of view as *"any attempts to compromise the availability or usefulness of evidence to the forensics process"* [16]. Within the field of digital anti-forensics, high level security behaviours have been identified which are related to security behaviour that can be observed on darknet markets: data hiding, trail obfuscation and data minimisation. It was found that some behaviours in the financial domain are combinations of these behaviours. Therefore the additional subclass 'choice of Online Financial Service Providers' (OFSPs) is added.

Data Hiding

Data hiding is *"the act of removing evidence from the view so that it is less likely to be incorporated into the forensic process"* [16]. Data hiding techniques allow the data to only be accessed and used by those who hid the data [19].

Encryption of communication. On darknet markets, users use PGP-encryption to hide their communication [20]. The PGP setup procedure is infamously known to be difficult to understand for the layman [21, 22]. Tutorials for PGP-encryption are widely available in the cybercriminal community [23]. To help those that do not want to go through this process, Hansa featured an auto-encryption functionality. However, in the case of Hansa Market this functionality proved to be a risk to security, LEA was able to disable the encryption [18]. It was found that cybercriminals also use alternative communication methods such as Skype, ICQ, Jabber, Privnote and Exploit.im [24, 25, 26]

Authentication. Authentication is the process of confirming the identity of a user. On a darknet market, a user is proving that he is who he says he is by entering a secret password that matches his darknet market username. After a successful login, i.e. when the user is authenticated, data that the user is authorized to see is shown ('unhidden'). Despite the fact that password authentication has been around for decades, its use still comes with a significant amount of bad practice [27]. Improving and understanding password hygiene is well researched in the field of information security behaviour [28]. Results show that even forcing, or nudging [29, 30], users to adhere to strict password requirements does not withhold users from picking predictable and easy-to-hack passwords [31, 32]. Research suggests that in the legitimate world, password reuse is very common [33]. Even when people are aware that strong passwords are important, these strong passwords are often reused over different websites [34]. A theoretically complex password can be easily breached, when it is reused on a website which has poor security practices [35]. As such, password reuse is non-secure behaviour. The most recent NIST security recommendations [36] therefore advise that it should be checked whether passwords used are truly unique or that they can be found in leaked password databases. Next to logging in with only a password, some markets allow users to enable two-factor authentication (2FA) [p.146 18]. First, the user is presented a text that is encrypted with their public PGP-key. Then, the user is challenged to obtain and submit the original unencrypted text, which only can be done using their secret private key [37, 38].

Linkability. Van de Sandt [2, p.153] states that cybercriminal activities leave behind fragments of information. These fragments are decentrally stored in a variety of databases. To ensure that these fragments are not included in the forensic process (*cf.* 'data hiding'), it is in the cybercriminals' interest to keep these fragments dispersed. Consequently, unlinkability from a security perspective is an attribute of confidentiality [39]. The unlinkability of darknet market pseudonyms refers to the inability for Law Enforcement to link two or more usernames to the same real world identity. Acts of linking, 'matching' [40], 'record linkage' [41] or 'Sybil account detection' [42] describe finding the pseudonyms that presumably refer to the same real-world entity. When multiple pseudonyms belonging to a single cybercriminal can be connected, a security risk for this criminal is created. LEA may accumulate advanced knowledge on a persons behaviour and identity, which may result in bringing this person to justice [43, 2].

Vendors active on multiple darknet markets may knowingly increase the linkability of their darknet market pseudonyms with the goal to increase their sales. By having a clear link between user accounts, valuable reputations can be transferred to other markets. In the pseudonymous world of darknet marketplaces, reputation distinguishes frauds from high quality vendors making 'reputation' an important asset to the cybercriminal [44]. Because darknet market users trade pseudonimously, usernames (instead of real names) come with a certain reputation [45]. They represent a brand [46] and are signals of trust [47]. Because vendors have the incentive of reusing a username over different markets, a variety research is done on matching usernames across markets. Ranging from obtaining an exact match to more elaborate techniques where similar but not identical usernames are matched [48, 49, 50, 40].

Not only the username signals trust and is tied to a reputation. This also goes for the public PGP-key listed by a vendor [44]. PGP-keys are suitable for signalling trustworthiness, because their legitimacy can be verified by asking the signalling party to decrypt a text [40]. Because PGP-keys are a reliable indicator of trust, vendors have incentives to register the same PGP-key over different markets. This behaviour has been successfully used to link pseudonyms with [44, 48, 51, 40]. Still, it must be taken into account that vendors may choose to use more than one key as an evasive strategy, because keys can expire or due to lost private keys [44, 51, 40]. Online darknet market search engine service Grams used username and PGP-key matching to offer linkability insights to a large audience [52]. Therefore, when a vendor reuses an username or a PGP-key over multiple markets, he actively increases the linkability of his or her pseudonyms.

Crossing jurisdictions. Security behaviour of cybercriminals is heavily shaped by the (absence) of laws and regulations [2, pp.76-81, p.198]. Cybercriminals create information asymmetries between key players (offenders, victims, law enforcement agencies) by distributing evidence across multiple jurisdictions [2, p.120]. In this section, this behaviour is described as 'crossing jurisdictions'. This refers to *"limiting what evidence can be captured due to inability to access data in one or more jurisdictions* [53]. Van de Sandt [2, pp.168-173] regards this type of security behaviour as a 'distribution countermeasure'. The author, referring to all types of cybercrimes, argues that three geographical locations give LEA jurisdiction: the location where the attack originates from, where the victims of the attack are located and the location of infrastructures that support the attack. Behaviour regarding the former two geographical locations of interest can be observed in darknet market context: a vendors' country of residence and the countries the vendor is willing to ship to are often denoted in the listing of vendors [54, 55, 56]. Marketing products in another jurisdiction may increase security [2] or decrease them [54].

Trail obfuscation

Trail obfuscation is defined *"the deliberate activity to disorient and divert a forensic investigation"*. Rogers [57] defines it as *"adding misdirection to the evidence"* which is very similar to the 'evidence counterfeiting' of Sremack and Antonov [53]. In the darknet market ecosystem, obfuscation techniques can be observed in the financial domain. Bitcoins (and other cryptocurrencies) facilitate transactions among cybercriminals [58, 59]. Estimations of the share of bitcoin transactions that are linked with illicit activities range from 1.1% [4], to 10%-30% [60] or even 50% [61].

Because money streams in the Bitcoin blockchain can be tracked, cybercriminals obfuscate these financial trails by 'mixing' or 'tumbling' their bitcoins [p.128 18, 62]. This is a type of 'cooperative obfuscation', which consists of the mixing of funds of various users [63]. In return for a transaction fee, these services generate a stream of transactions that turn investigating the money stream into highly complex procedures [64]. Mixing of criminal proceeds obtained through darknet market transactions is not an irregularity. Janze [58] shows that usage of transaction obfuscation services is related to the amount of sales on darknet markets. Stronger still, the first Silk Road included mixing functionalities on their platform [20]. Next to mixing services, online gambling sites that accept Bitcoin also receive high proportions of bitcoins linked with illicit activities [65, 66]. The highly popular gambling sites receive huge amounts of relatively small transactions [67, 68]. Some scholars e.g. [65, 66, 69] assume that gambling is to obfuscate money trails. However, Meiklejohn et al. [70] stipulates that - at least in 2013 - gambling sites such as SatoshiDice are not mixing bitcoin effectively. According to the authors, the addresses belonging to SatoshiDice are publicly known and users have to specify a payout address. This makes an permanent link between the bitcoins that are placed as bets and the bitcoins paid out.

Choice of OFSPs: obfuscation, data minimisation & crossing jurisdictions

As described in section 2, a few intermediaries in the Bitcoin ecosystem can subjected to regulation [71]. A bitcoin exchange is an intermediary that functions as a digital currency exchange office. At these exchanges, bitcoins can be traded in for fiat currencies (hard currencies such as euros and dollars). When a cybercriminal wants to convert the bitcoins earned with criminal activities to spendable money at scale, using an exchange at some point is unavoidable [70]. Because of the public and transparent nature of bitcoin transactions, it can be observed to what exchanges cybercriminals transact to [65].

Bitcoin exchanges can be subjected to regulation and thus, from the perspective of a cybercriminal, form a security risk. Among exchanges, large inconsistencies in identity verification and monitoring transactions exist,

which give way to fraudulent behaviour [72]. For years, bitcoin exchanges did not have to comply to any *universal* anti-money laundering regulations [73]. Bitcoin exchanges relocate their offices to jurisdictions that have less stringent AML requirements [74, 75]. Thus, bitcoin exchanges located in these jurisdictions remain capable of laundering money originating from criminal activities [76, 77]. While numerous exchanges require identification and apply 'know your customer' (KYC) principles correctly, others do not to serve clients that prefer anonymity [71]. *Peer-to-peer* (P2P) exchanges facilitate transactions between peers directly. In practice, this means that one party transacts bitcoins to a counterparty directly. The counterparty then transfers the agreed amount of fiat currency back to the first party. Because no central authority is involved in this transaction, enforcing identity verification is challenging [78]. Additionally, this type of transactions limit the usefulness in monitoring the blockchain, because it is difficult to prove that the transactions are related to the same real-world identity [78].

Thus, cybercriminals may exhibit different types of security behaviour when choosing an OFSP: (a) they can minimise their data footprint by not registering any information, (b) perform trail obfuscation techniques by registering falsified information or by using P2P-exchanges or (c) cross jurisdictions by choosing the supporting infrastructures (i.e. OFSPs) for transactions [2].

Vendor characteristics & Security

A few indications are found in literature that relations between certain characteristics of vendors and security behaviour exist. Firstly, Van de Sandt [2, p.96] argues that the security practices of cybercriminals are positively influenced by the cybercriminals' experience. However, [18, p.161] found evidence that vulnerabilities, software updates or new security developments may be ignored by cybercriminals because of a 'status-quo bias'. The status-quo bias describes that one has the tendency to let things (in this case, security practices) remain the same [79]. The term is first coined by Samuelson and Zeckhauser [80], who argue that a status-quo bias is also present when an individual is not aware of the other, updated, options available to him. The behavioural pitfall of remaining the satus-quo with regard to security measures, might become a security risk and may consequently lead to deanonymisation [18, p.42]. Secondly, cybercriminals stemming from a traditional criminal background might make mistakes on the technical part of their security [p.154 18]. Thirdly, the business success is hypothesised to relate to security decisions, since the (opportunity) costs inflicted when getting caught or disrupted increase and the investments in security are spread out over larger revenues [2].

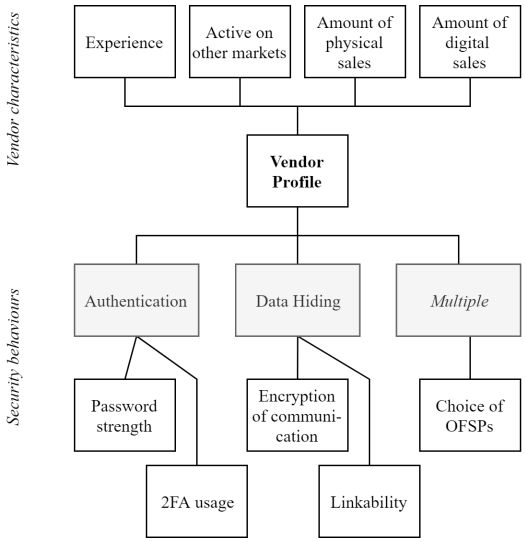


Figure 1: Conceptual model, assessing the relation between vendor profiles and security behaviours.

3 CONCEPTUAL MODEL

Not all security behaviours identified in section 2 are further analysed in the Hansa data. Firstly, pseudonyms can be easily imitated. Therefore only the reuse of PGP-keys over different markets is considered. Secondly, shipping internationally may increase or decrease security, thus is also not considered. The identified characteristics of vendors that may relate to security are *experience* and the amount of *physical sales* and *digital sales*. Because experience and sales can also be gained on other markets, *active on other markets* is added to the model. Similar to [24, 81] vendors are assigned 'vendor profiles', based on these characteristics.

4 METHODOLOGY

The vendor characteristic *experience* is defined as the amount of days between a vendors' first and last sale [24], *active on other markets* is measured by regarding whether a vendor used the Hansa functionality enabling vendors to import their reputation from other markets and *physical sales*/*digital sales* are the number of sales per vendor of which the market indicated that physical shipment/digital transaction applied. Vendors with similar characteristics are assigned vendor profiles using Latent Profile Analysis (LPA). An LPA maximises homogeneity within clusters and heterogeneity between clusters and takes data on any measurement level as input [82]. The LatentGOLD statistical software of Vermunt and Magidson [83] is used to achieve the clustering of resemblant vendors.

Authentication. This behaviour is regarded by analysing a vendors' password strength and whether a vendor uses 2FA. *Password strength* is measured by regarding password

complexity in entropy bits and by assessing password reuse. The amount of entropy bits is calculated via equation 1:

$$H_i = \log_2(R^{\#pw_i}), \quad pw_i = \{c_1, c_2, c_3, \ldots c_n\} \qquad (1)$$

Each vendor i has a password pw that consists of $\#pw$ unique characters. If R is the amount of all characters recognised by the system, the entropy H is described by taking the $\log_2$ of $R^{\#pw}$. R is set to the extended ASCII-charset ($R = 95$), which are all characters that can be created with common combinations of keys found on regular keyboards. Human-generated passwords follow certain trends, such as (re)using some characters more often than others, which greatly reduces entropy [84, 85]. Therefore, the number of unique characters are used to penalise the repetition of characters [86]. Password reuse is measured by matching passwords in the 'Have I Been PWND' database [87]. This database includes more than 10 billion leaked passwords, of which 573 million are unique. The SHA1-hashes of the passwords in the PWND database are matched with the SHA1 hashes of the passwords of Hansa users. *2FA usage* is indicated per vendor in the Hansa data.

Data Hiding. This behaviour assessed in two ways. First, through *Encryption of communication.* This security practice is measured by extracting key sizes of the public PGP-keys using a Python implementation of GNU Privacy Guard (GnuPG). Secondly, the *Linkability* of darknet market pseudonyms is measured by matching PGP-keys in the database of the *Grams* darknet market search engine.

Obfuscation, minimisation & crossing jurisdictions. To analyse to what extend these behaviours are observed among darknet market vendors, the payout addresses are queried to *Chainalysis*, which provides the much needed contextual information on the transactions performed by each vendor. Through a custom API-script, it is obtained whether a vendor (a) uses a payout address belonging to a known OFSP, i.e. payouts are directly transacted to a service wallet or (b) has his or her earnings transacted to a private wallet first. Unknown clusters of > 5 bitcoin addresses are considered to be private wallets. In case a private wallet is used, it is analysed whether this wallet has outgoing transactions to known OFSPs.

To analyse all security behaviours jointly, a simple scoring function is calculated. Vendors are awarded one point for each of the satisfied criteria:

- A higher password complexity than median password complexity;
- No match in the PWND password database;
- 2FA Enabled;
- PGP-keystrength of 2048+ bits;
- No direct transactions to central exchanges.

Note how the PGP-match in the Grams database is omitted from the scoring method, as this security behaviour only applies to vendors that are active on other markets. While this scoring method has its obvious limitations, it does separate the security conscious vendors from the vendors that show less secure behaviour.

5 RESULTS

Clustering vendors

In order to capture multiple vendor characteristics in one variable, the vendors are clustered into vendor profiles through LPA. No model with less than than 10 clusters showed a perfect *global* fit. Even the Bayesian Information Criterium (BIC), that tends to favour parsimonious and underfitting models [88], did not provide a definitive answer which model with $\leq$ 10 clusters to select. However, achieving full heterogeneity between all clusters is not the most important goal of clustering vendors into vendor profiles. The resulting clusters should be easy to interpret in the context of this research and the sample sizes should be sufficiently large [89, 90]. These considerations are important, especially when fitting criteria do not minimise [91]. The 5-cluster model is a parsimonious model that does differentiate between physical and digital sales. As assessed by a Wald-test, all vendor characteristics contribute significantly to the 5-cluster model thus should be retained. Moreover, while 5 bivariate residuals remain significant, 4 out of 6 variable pairs do show a reduction of approximately 90% when the 5-cluster model is compared to the 1-cluster model. Striving for such of BVRs instead of non-significance is valid when sample sizes are large [92]. Additionally, pairwise Kruskal-Wallis H tests are performed that indicate that all relevant medians/means of the vendor characteristics differ significantly between clusters. The distributions are visualised in Figure 2 and Figure 3. Based on the distributions of vendor characteristics and what products are sold within clusters, the following latent profiles, or 'vendor types' are assigned to each vendor belonging to a cluster.

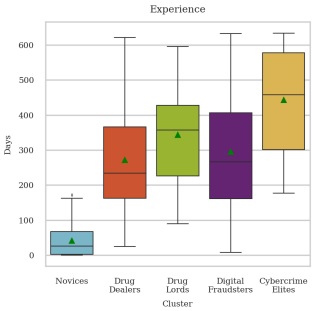


Figure 2: Distribution of experience per cluster

The biggest cluster, `Novices` ($n = 988$), distinguishes itself from the other clusters by a relatively low amount of physical and digital sales and the lowest number of days of experience. Only 40.2% of the users imported their reputation from other markets. This is the lowest compared to all other clusters and below the market average of 52.3%. About 80% of the products sold are drugs, but a few

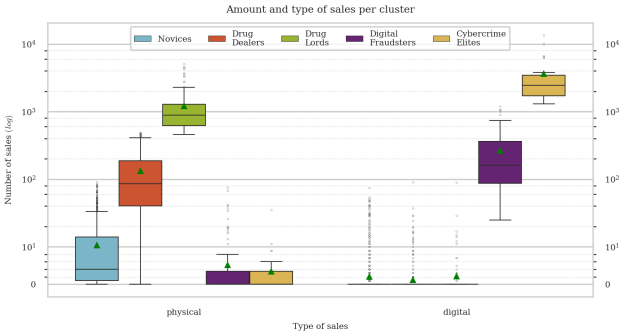


Figure 3: Amount of sales per sales type and cluster

vendors do have digital sales. No vendors with more than 100 physical or digital sales are included in the `Novices` profile.

`Drug Dealers` ($n = 509$) have more physical sales and experience compared to the `Novices`. In terms of being active on multiple markets, 69.0% of the `Drug Dealers` have their reputation imported, which is higher than the `Novice` cluster. More than half of the vendors clustered as `Drug Dealers` has been active for 230 days, similarly half of the cluster has more than 80 physical sales. Of the products sold, 98% are drugs.

The last cluster with mainly physical sales are `Drug Lords` ($n = 110$), who do not really differentiate in terms of experience, but do have extremely high amounts of physical sales. The average multihoming score is high: 78.2% indicated to be active on other markets. All sales (100%) are drugs related.

The next two clusters thrive in digital sales rather than physical sales. First, the `Digital Fraudsters` ($n = 103$) might have very few or very much days of experience. Theses vendors do have at least 15 sales in te digital domain. About 75% has more than 100 digital sales. Not much can be said on their preference to be active on other markets, since 58.3% showed to import their reputation. Some vendors with mainly digital sales also made a number of physical sales.

Lastly, `Cybercrime Elites` is a very small group ($n = 23$). It is chosen to accept this cluster with few vendors because indeed, very successful vendors of digital items are scarce and they clearly trump the `Digital Fraudsters` in terms of sales. A large amount of `Cybercrime Elites` is active on other markets (73.9%).

Security behaviours

Data hiding: authentication. The passwords of 1081 vendors ($\approx$ 62.4% of all vendors) and 85620 regular members ($\approx$ 20.5% of all members) are retrieved. These passwords are analysed in this section. First, the distributions of password complexity in entropy bits are regarded. To statistically determine whether there is any difference between the clusters' means, a one-way Analysis of Variance (ANOVA) is performed. After log-transforming the data and removing outliers $z \geq 5$, the data meets the assumptions of normality within clusters and homogeneity of variances. The ANOVA-test is significant ($F(4, 1068) = 5.89, p = 0.0001$). This shows that the means of password complexity differ significantly among the vendor profiles. A Tukey-Kramer HSD post-hoc test is performed to learn how the vendor profiles significantly differ from each other. Only the differences in password complexity between 4 pairs of vendor profiles are significantly different. No statistically significant differences were found for `Cybercrime Elites`. This may be due to a lack of power as a consequence of the small sample size. The tests show how `Drug Lords` > `Novices` and `Drug Lords` > `Digital Fraudsters`. In which '>' denotes a statistically significant higher average of password complexity. Likewise, `Drug Dealers` > `Digital Fraudsters`.

Lastly, it is taken into consideration that simpler passwords might be chosen when vendors make use of two-factor authentication. Generally, this does not seem to be the case. 2FA-usage correlates *positively* with password complexity, as assessed by a Spearman rank-order correlation ($r_s = 0.219, p < 0.0000$). This shows that vendors do not tend to compensate relatively poor passwords with the additional layer of security that 2FA adds. Rather, the positive correlation is an indication that the priority given to security truly differs between vendors.

Regarding password reuse, the passwords of 26540 Hansa Market users could be matched with the passwords in the PWND database. This amounts to 30.6% of the users whose plaintext password was available. Regarding vendor accounts, 185 passwords are matched (17.1%). A χ^2-test of the proportion password matches/no matches is performed. This proportion differs significantly between vendor profiles ($p = 0.0064$). From a pairwise post-hoc z-test of proportions with FDR-BH correction, the following is inferred. First, for most pairwise comparisons no significant difference in the proportion of non-matches/matches is found. Secondly, `Drug Dealers` > `Novices` and `Drug Dealers` > `Digital Fraudsters`. In which '>' indicates a statistically significant higher proportion of non-matches (i.e. more secure behaviour). Likewise, `Drug Lords` > `Digital Fraudsters`.

With respect to 2FA, 60.5% ($n = 1049$) of the vendor population enabled 2FA. Statistically significant differences between profiles exist, as assessed by a χ^2 test of homogeneity ($p < 0.0000$). A z-test with FDR-BH corrections shows that `Drug Lords` > `Drug Dealers` > `Novices` > `Digital Fraudsters` and `Drug Lords` > `Cybercrime Elites`. In which '>' indicates a statistically significant higher proportion of 2FA usage.

Data hiding: encryption of communication. The PGP-adoption among vendors is high. Only 5 vendors do not have a PGP-key listed: two `Novices` and three `Drug Dealers`. It could be, that they removed their PGP keys from their

accounts after they stopped trading. The adoption of PGP-keys among buyers is noticeably lower. Only 50,657 out of 415,703 buyers (12.19%) registered a PGP-key. The key size indicates how 'secure' the key is. Weak keys ($\leq$ 1024 bits) are observed for 9 vendors and 88 buyers. Even by 2015's standards, these key lengths are considered not to be sufficient [93, 94]. Currently, NIST recommends key sizes of asymmetric cryptosystems based on factorisation problems, such as RSA, to be at least 2048-bits [36]. The creation date of the keys is compared with their key size. For vendors, no trend is observed that key sizes increase over time. Considering that the security benefit of any key stronger than 2048-bits is negligible, it is expected that key sizes are chosen 'randomly' or according to whatever is recommended in one of the many PGP-tutorials found on dark web discussion fora. However, extremely secure keys (2048+ bits) are more often found among vendor profiles that proved to be more 'security aware' in the other analyses as well. No weak keys of `Drug Lords` are found and among the digitally focused clusters, very strong PGP-keys are observed notably less. To statistically determine which vendor types show higher proportions of extremely secure PGP-keys, a χ^2-test with a FDR-BH adjusted post-hoc z-test is performed on the proportion of 2048+ bits keys within the clusters. Thus, in this test all key sizes $\leq$2048-bits are grouped together. The clusters differ significantly, $p < 0.0000$. From the post-hoc test the following conclusions are valid: `Drug Lords` > `Digital Fraudsters` & `Cybercrime Elites` and `Drug Dealers` > `Novices` > `Digital Fraudsters` & `Cybercrime Elites`. In which '>' indicates a statistically significant higher proportion of extremely secure PGP-keys.

Data hiding: linkability. When active on multiple markets it is beneficial to use the same PGP-key on each market [44]. It does introduce potential security risks, because it allows LEA to link darknet market pseudonyms. This section shows which vendors are known to be active on other markets, i.e. they imported their reputation from other markets, but whose PGP-keys could not be matched via the database of the Grams search engine. If no match has been found, this indicates that vendors use different PGP-keys. Figure 4 shows how the following groups overlap: vendors with PGP key (n = 1728), vendors known to be active on other markets (n = 908) and vendors who are linked with Hansa and any other market (n = 902). From this figure is concluded there is a group (n = 265) who is known to be active on another market but whose PGP-keys could not be matched in the Grams data. However, no differences between clusters exist in terms of the proportion of vendors that are known to be active on other markets that presumably changed their PGP-keys as assessed by a χ^2-test (p = 0.8425).

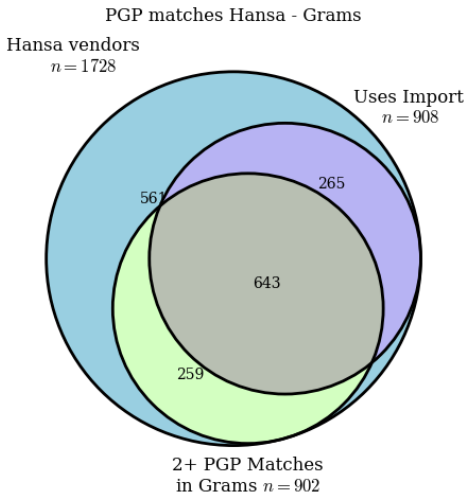


Figure 4: Percentage of grouped key sizes per cluster

Multiple: Choice of OFSPs

The use of intermediaries to facilitate the conversion of cryptocurrency earnings to spendable fiat currencies is inevitable. At the same time it introduces potential security risks. In the Hansa Market dataset 19.238 unique bitcoin payout addresses could be recovered. These addresses are vendors' own. From these payout addresses, 2680 ($\approx$ 14%) could be directly attributed to clusters that are identified with known service wallets, such as centrally organised exchanges, peer-to-peer exchanges and bitcoin mixers. In this thesis, payouts directly transacted to such known services are referred to as *direct links*. As expected, the majority of the bitcoin addresses cannot be directly linked with service wallets. This may be due to the fact that a) some service wallets are not identified by Chainalysis and b) vendors do *not* send their criminal proceeds directly to a service wallet. Instead, they first accumulate their earnings on a privately owned (hardware) wallet.

It is very likely that a high amount of vendors use private hardware wallets for storing their bitcoins (scenario b). Therefore, the heuristic presented in section 4 is applied to separate the vendors with hardware wallets (scenario b) from the Chainalysis shortcomings (scenario a). The heuristic results in the 16,564 payout addresses that are not directly linked to a service being categorised as follows. Assumed private wallets (n = 4165), private wallets with no exposure (n = 4037) and wallets that are either private wallets or service wallets (n = 8344). The analysis of 12 addresses threw an error, partly because of parsing issues or due to the address not being recognised by Chainalysis. These numbers are summarised in Figure 5.

Transacting directly to a known service does not equal a security risk per se. Transacting to mixing services and peer-to-peer exchanges is considered to be more secure compared to transacting to a regular exchange. Therefore, the types of services directly transacted to are shown in Figure 6. Here, non-normalised counts are visualised. The

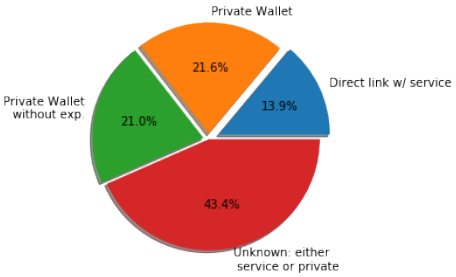


Figure 5: Types of wallets identified

relatively low amount of observations per category would otherwise overestimate the precision of the graph. The figure only displays services *directly* transacted to. Services indirectly transacted to cannot be plotted over time. When payouts are first accumulated on private wallets, the amount of transactions to services and their transaction dates are very much dependent on the preferences of the vendor. How frequently they move their funds from the cold storage to the services for further processing, is up to them.

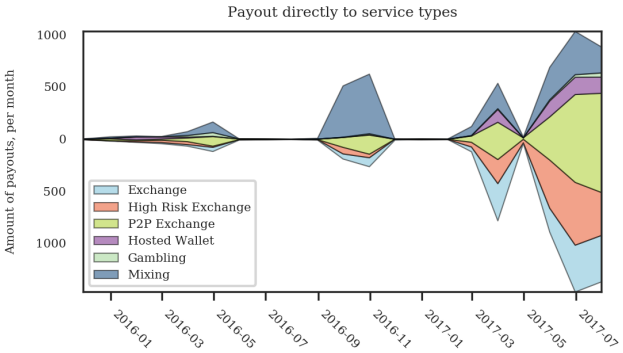


Figure 6: Payout transacted to

Direct payouts to exchanges are an ongoing event (Figure 6). From the perspective of Law Enforcement, it is good to see that over this extended period of time, this security risk keeps being created by vendors. The link that is irrefutable and that translates to a severe security risk is having the payouts directly transacted to a central exchange. The proportion of vendors of whom this behaviour is *not* observed is compared between clusters. Differences are significant, as assessed by a χ^2-test ($p < 0.0000$). The results of the post-hoc test actually contradict the observations made during analysis of the other security behaviours. `Novices` show the most secure behaviour regarding transacting directly to exchanges. In this case, `Drug Lords` have significantly lower proportions of secure behaviour compared to `Novices` and `Drug Dealers`. The proportion of secure behaviour among `Cybercrime Elites` is the lowest compared to all other vendor profiles, as confirmed by the post-hoc test.

6 JOINT ANALYSIS OF SECURITY BEHAVIOURS

In this section vendors the overall security behaviour of vendors is assessed by a simple scoring method. Per vendor one point is awarded for for each of the following security practices: the vendor has a higher password complexity than median password complexity of vendors (67-bits), there is no password match in the PWND databse, 2FA is enabled, an 'extremely secure' PGP-key of more than 2048-bits is chosen and no direct transactions to exchanges have been observed. This results in the distribution presented in Figure 7.

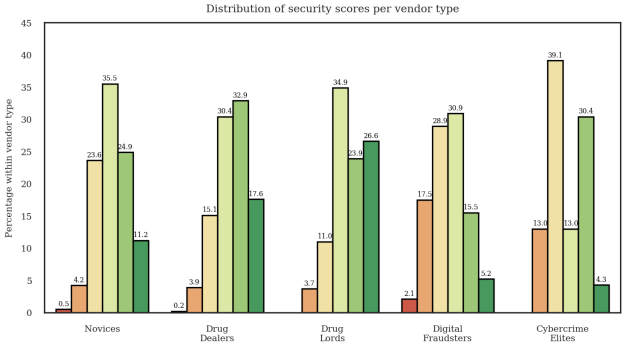


Figure 7: Distributions of security scores (0-5) per cluster

Only a few vendors managed to score 0 out of 5 security points. About half of the `Digital Fraudsters` (49%) and half of the `Cybercrime Elites` (52%) have a score below 3. This is in congruence with the analyses of individual security behaviours, which showed that these clusters consistently have low proportions of secure behaviour. Important to note is that 15% of `Drug Lords` have a security score below 3. This is in line with the observations made when examining the relation between physical sales and the individual security behaviours: while the tendency is that security behaviour scales with the amount of physical sales, for every security behaviour exceptions were observed.

Since the scoring method is a simple sum of dichotomous indicators, it is of no surprise that the score of 3 is common. The number of `Drug Lords` with a score of 3 is more than expected. However, not every security malpractice has an equal amount of 'impact' on the security of vendors. For example, the use of 2FA and extremely high PGP-key sizes should be interpreted as an indicator of security awareness or priority given to security rather than security mechanisms that prevent immediate de-anonymisation.

About a third of the `Cybercrime Elites` (35%) is scored with a 4 or 5. This shows that within this cluster of vendors there is actually a group of vendors that behaves securely (>3) and a group that does not (<3). However, the sample size within this cluster is not large ($n = 23$), this translates in relatively large differences increases or between security scores when these differences are expressed as percentages.

No poor security behaviours are observed for 27% of the `Drug Lords`. Given this observation, it would be concluded

that about a quarter of the Drug Lords are vendors that strive for 'maximum security'. When including vendors with slightly suboptimal security that score a 4, the Drug Lords (51%) do not differ much from the Drug Dealers (51%).

7 RELATED WORK

Van de Sandt [2, p.231] demonstrates the necessity for a new academic field of study demystifying the security practices of cybercriminals. This field of study awaits major contributions from the socio-technical disciplines known for combining social sciences with computer science research. The current academic works that arguably make a first contribution to shaping this field of study are discussed in this section. Van Hardeveld [18] elaborates on the decision-making of carders[1]. The author examines technical security mechanisms found in online carding tutorials and discusses cognitive biases that lead to suboptimal security. Expert interviews provided evidence that some of these biases apply to carders. Van de Sandt [2] lays a mostly theoretical foundation of how cybercriminals deploy technical computer security controls that aim to protect the criminal and the crimes he or she commits [2, p.7]. While Van de Sandt has a strong focus on conceptualising security practices of cybercriminals and his findings are predominantly of qualitative form, his research is not aimed at darknet market users specifically nor does it provide explicit definitions of security behaviours that might be observed on darknet markets. He does acknowledge that approaching the research on cybercriminal security behaviour in a quantitative manner, will produce more granular insights [2, p.232]. Lastly, the research of Wegberg and Verburgh [44] revolves around a single and specific security behaviour. The authors show that vendors attempt to reduce the linkability of their pseudonyms when migrating from one market to another. The security mechanism analysed is whether vendors stick with their PGP-key and/or username when switching markets.

Apart from research on the security behaviour of individuals, a significant amount of work focuses at understanding the security mechanisms of entities that facilitate the illegal transactions. These are efforts on a *platform-level*. For example, the self-regulation on darknet markets [12] through its reputation mechanisms [10, 11], the interaction between the popularity of anonymity enhancing cryptocurrencies and darknet markets [61, 58] or the forensic challenges and opportunities that the popularity of cryptocurrencies results in [95]. While important to the field of research of cybercriminals' security practices, these studies fail to generate insights on what additional security measures individuals on darknet markets take to safeguard their security.

There are numerous indications throughout literature that (cyber)criminals not always achieve maximum security. First, any criminal is economically incentivised, resulting in a trade-off between enhanced security and improved efficiency of operations [96]. Second, Holt et al. [13] observe that users actively trade-off between risks and rewards of a transaction on forums facilitating the trade of stolen data. Third, Van de Sandt [2] argues that 'perfect security' is not economically viable for cybercriminals. Fourth, in a study on online underground forums, Sundaresan et al. [26] show that vendors do not consistently use VPN services to hide their likely geolocation and that they are prone to use less secure communication methods. Fifth, Bradley [7, p. 195] observes that users on a darknet marketplaces use the auto-encryption features of the market, even though this feature poses a security risk.

Three academic works exist in which the security behaviour of larger populations darknet market users is analysed. Next to the before-mentioned Wegberg and Verburgh [44], who analysed the evasion measures of vendors upon switching markets, Soska and Christin [48] measure how the use of encrypted communication methods increases over time on darknet marketplaces. Décary-Hétu, Paquet-Clouston and Aldridge [54] make an effort in measuring security risk-taking behaviour of vendors. The authors operationalise 'security risk' by only taking into account the willingness of vendors to ship internationally. These three works that study security behaviour on darknet markets, tend to focus at only one aspect of security behaviour and do not consider characteristics of the vendors when drafting conclusions about their security behaviour.

8 CONCLUSION

Overall, Drug Lords and Drug Dealers show the most secure behaviour. About a quarter of the Drug Lords have 'maximum security'. Digital Fraudsters and Cybercrime Elites show poor security behaviour most often. However, still 15% of the Drug Lords exhibit relatively much less secure behaviour. Likewise, a third of the Cybercrime Elites does show mostly secure behaviour.

Therefore, an approximately causal relationships is inferred between on the one hand vendor types, that represent a combination of business success in terms of physical and digital sales, experience and activity on other markets and on the other hand security behaviour. Among vendors selling mainly physical items, an increasing amount of business success, experience and activity on other markets, results in an increase of more secure behaviour regarding simple security measures such as authentication related practices. These differences may be explained through analysing them from the perspective of cybercriminals conducting risk assessments.

[1] Carders trade stolen credit card and bank account details.

REFERENCES

[1] Y. Zhang et al. 'A survey of cyber crimes'. In: *Security and Communication Networks* 5.4 (2012), pp. 422–437.

[2] E. Van de Sandt. 'Deviant security: the technical computer security practices of cyber criminals'. PhD thesis. University of Bristol, 2019.

[3] P. Hartel and R. van Wegberg. 'Crime and Online Anonymous Markets'. In: *International and Transnational Crime and Justice* (2019), p. 67.

[4] Chainalysis. *Crypto Crime Report 2020*. 28th Jan. 2020. URL: https://go.chainalysis.com/rs/503-FAP-074/images/2020-Crypto-Crime-Report.pdf (visited on 16/03/2020).

[5] D. S. Dolliver, S. P. Ericson and K. L. Love. 'A geographic analysis of drug trafficking patterns on the tor network'. In: *Geographical review* 108.1 (2018), pp. 45–68.

[6] R. Van Wegberg et al. 'Plug and prey? measuring the commoditization of cybercrime via online anonymous markets'. In: *27th USENIXSecurity Symposium (USENIX Security 18)*. 2018, pp. 1009–1026.

[7] C. Bradley. 'On the Resilience of the Dark Net Market Ecosystem to Law Enforcement Intervention'. PhD thesis. UCL (University College London), 2019.

[8] D. A. Wheeler and G. N. Larsen. *Techniques for cyber attack attribution*. Tech. rep. INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA, 2003.

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

[10] J. Aldridge and D. Décary-Hétu. 'Not an'Ebay for Drugs': the Cryptomarket'Silk Road'as a paradigm shifting criminal innovation'. In: *Available at SSRN 2436643* (2014).

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

[12] F. Wehinger. 'The Dark Net: Self-regulation dynamics of illegal online markets for identities and related services'. In: *2011 European Intelligence and Security Informatics Conference*. IEEE. 2011, pp. 209–213.

[13] T. J. Holt et al. 'Examining the risk reduction strategies of actors in online criminal markets'. In: *Global Crime* 16.2 (2015), pp. 81–103.

[14] C. Bradley and G. Stringhini. 'A Qualitative Evaluation of Two Different Law Enforcement Approaches on Dark Net Markets'. In: *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE. 2019, pp. 453–463.

[15] R. Broadhurst, M. Ball and H. Trivedi. 'Fentanyl availability on darknet markets'. In: *Broadhurst R, Ball M & Trivedi H* (2020).

[16] R. Harris. 'Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem'. In: *digital investigation* 3 (2006), pp. 44–49.

[17] J. M. Bauer and M. J. Van Eeten. 'Cybersecurity: Stakeholder incentives, externalities, and policy options'. In: *Telecommunications Policy* 33.10-11 (2009), pp. 706–719.

[18] G. J. Van Hardeveld. 'Deviating from the cybercriminal script: Exploring the contextual factors and cognitive biases involved in carding'. PhD thesis. University of Southampton, 2018.

[19] C. S. Peron and M. Legary. 'Digital anti-forensics: emerging trends in data transformation techniques'. In: *Proceedings of*. 2005.

[20] J. Cox. 'Staying in the shadows: the use of bitcoin and encryption in cryptomarkets'. In: *The Internet and drug markets* (2016), pp. 41–48.

[21] A. Whitten and J. D. Tygar. 'Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0.' In: *USENIX Security Symposium*. Vol. 348. 1999, pp. 169–184.

[22] S. Ruoti et al. 'Why Johnny still, still can't encrypt: Evaluating the usability of a modern PGP client'. In: *arXiv preprint arXiv:1510.08555* (2015).

[23] G. J. van Hardeveld, C. Webber and K. O'Hara. 'Deviating from the cybercriminal script: exploring tools of anonymity (mis) used by carders on cryptomarkets'. In: *American Behavioral Scientist* 61.11 (2017), pp. 1244–1266.

[24] R. van Wegberg et al. 'Go See a Specialist? Predicting Cybercrime Sales on Online Anonymous Markets from Vendor and Product Characteristics'. In: (2020), pp. 816–826.

[25] J. Aldridge and R. Askew. 'Delivery dilemmas: How drug cryptomarket users identify and seek to reduce their risk of detection by law enforcement'. In: *International Journal of Drug Policy* 41 (2017), pp. 101–109.

[26] S. Sundaresan et al. 'Profiling underground merchants based on network behavior'. In: *2016 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE. 2016, pp. 1–9.

[27] S. Furnell. 'Assessing password guidance and enforcement on leading websites'. In: *Computer Fraud & Security* 2011.12 (2011), pp. 10–18.

[28] J. M. Stanton et al. 'Analysis of end user security behaviors'. In: *Computers & security* 24.2 (2005), pp. 124–133.

[29] S. Kankane, C. DiRusso and C. Buckley. 'Can We Nudge Users Toward Better Password Management? An Initial Study'. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. 2018, pp. 1–6.

[30] S. Furnell et al. 'Enhancing security behaviour by supporting the user'. In: *Computers & Security* 75 (2018), pp. 1–9.

[31] R. Shay et al. 'A spoonful of sugar? The impact of guidance and feedback on password-creation behavior'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 2015, pp. 2903–2912.

[32] S. Komanduri et al. 'Of passwords and people: measuring the effect of password-composition policies'. In: *Proceedings of the sigchi conference on human factors in computing systems*. 2011, pp. 2595–2604.

[33] M. Golla et al. '" What was that site doing with my Facebook password?" Designing Password-Reuse Notifications'. In: *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. 2018, pp. 1549–1566.

[34] R. Wash et al. 'Understanding password choices: How frequently entered passwords are re-used across websites'. In: *Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016)*. 2016, pp. 175–188.

[35] B. Ives, K. R. Walsh and H. Schneider. 'The domino effect of password reuse'. In: *Communications of the ACM* 47.4 (2004), pp. 75–78.

[36] E. Barker et al. *Recommendation for key management: Part 1: General, 5th Rev.* National Institute of Standards and Technology, Technology Administration, 2020.

[37] T. Carr et al. 'Into the Reverie: Exploration of the Dream Market'. In: *2019 IEEE International Conference on Big Data (Big Data)*. IEEE. 2019, pp. 1432–1441.

[38] G. Zhou et al. 'A Market in Dream: the Rapid Development of Anonymous Cybercrime'. In: *Mobile Networks and Applications* 25.1 (2020), pp. 259–270.

[39] A. Pfitzmann and M. Hansen. 'A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management'. In: (2010).

[40] X. H. Tai, K. Soska and N. Christin. 'Adversarial Matching of Dark Net Market Vendor Accounts'. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019, pp. 1871–1880.

[41] P. Christen. *Data matching: concepts and techniques for record linkage, entity resolution, and duplicate detection*. Springer Science & Business Media, 2012.

[42] R. Kumar et al. 'eDarkFind: Unsupervised Multi-view Learning for Sybil Account Detection'. In: *Proceedings of The Web Conference 2020*. 2020, pp. 1955–1965.

[43] T. N. Ho and W. K. Ng. 'Application of stylometry to darkweb forum user identification'. In: *International Conference on Information and Communications Security*. Springer. 2016, pp. 173–183.

[44] R. van Wegberg and T. Verburgh. 'Lost in the Dream? Measuring the effects of Operation Bayonet on vendors migrating to Dream Market'. In: *Proceedings of the Evolution of the Darknet Workshop*. 2018, pp. 1–5.

[45] D. Décary-Hétu and A. Leppänen. 'Criminals and signals: An assessment of criminal performance in the carding underworld'. In: *Security Journal* 29.3 (2016), pp. 442–460.

[46] J. Lusthaus. 'Trust in the world of cybercrime'. In: *Global crime* 13.2 (2012), pp. 71–94.

[47] T. J. Holt, O. Smirnova and A. Hutchings. 'Examining signals of trust in criminal markets online'. In: *Journal of Cybersecurity* 2.2 (2016), pp. 137–145.

[48] K. Soska and N. Christin. 'Measuring the longitudinal evolution of the online anonymous marketplace ecosystem'. In: *24th USENIX Security Symposium*. 2015, pp. 33–48.

[49] J. Van Buskirk et al. 'The recovery of online drug markets following law enforcement and other disruptions'. In: *Drug and alcohol dependence* 173 (2017), pp. 159–162.

[50] D. Décary-Hétu and L. Giommoni. 'Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of Operation Onymous'. In: *Crime, Law and Social Change* 67.1 (2017), pp. 55–75.

[51] J. Broséus et al. 'Studying illicit drug trafficking on Darknet markets: structure and organisation from a Canadian perspective'. In: *Forensic science international* 264 (2016), pp. 7–14.

[52] G. Branwen. *Darknet Market Archives*. Mar. 2020. URL: https://www.gwern.net/DNM-archives#grams.

[53] J. C. Sremack and A. V. Antonov. 'Taxonomy of Anti-Computer Forensics Threats.' In: *IMF* 103 (2007), e12.

[54] D. Décary-Hétu, M. Paquet-Clouston and J. Aldridge. 'Going international - Risk taking by cryptomarket drug vendors'. In: *International Journal of Drug Policy* 35 (2016), pp. 69–76.

[55] M. C. Van Hout and T. Bingham. 'Responsible vendors, intelligent consumers: Silk Road, the online revolution in drug trading'. In: *International Journal of Drug Policy* 25.2 (2014), pp. 183–189.

[56] J. Van Buskirk et al. 'Who sells what? Country specific differences in substance availability on the Agora cryptomarket'. In: *International Journal of Drug Policy* 35 (2016), pp. 16–23.

[57] M. Rogers. 'Anti-forensics: the coming wave in digital forensics'. In: *Retrieved September* 7 (2006), p. 2008.

[58] C. Janze. 'Are cryptocurrencies criminals best friends? Examining the co-evolution of bitcoin and darknet markets'. In: (2017).

[59] S. Kethineni, Y. Cao and C. Dodge. 'Use of bitcoin in darknet markets: Examining facilitative factors on bitcoin-related crimes'. In: *American Journal of Criminal Justice* 43.2 (2018), pp. 141–157.

[60] H. Sun Yin and R. Vatrapu. 'A first estimation of the proportion of cybercriminal entities in the bitcoin ecosystem using supervised machine learning'. In: *2017 IEEE International Conference on Big Data (Big Data)*. 2017, pp. 3690–3699.

[61] S. Foley, J. R. Karlsen and T. J. Putniņš. 'Sex, drugs, and bitcoin: How much illegal activity is financed through cryptocurrencies?' In: *The Review of Financial Studies* 32.5 (2019), pp. 1798–1853.

[62] R. Wegberg, J. Oerlemans, O. v. Deventer et al. 'Bitcoin money laundering: mixed results? An explorative study on money laundering of cybercrime proceeds using bitcoin'. In: *Journal of Financial Crime* 25 (2018), p. 17.

[63] A. Narayanan and M. Möser. 'Obfuscation in bitcoin: Techniques and politics'. In: *arXiv preprint arXiv:1706.05432* (2017).

[64] D. Moore and T. Rid. 'Cryptopolitik and the Darknet'. In: *Survival* 58.1 (2016), pp. 7–38.

[65] Y. Fanusie and T. Robinson. 'Bitcoin laundering: an analysis of illicit flows into digital currency services'. In: *Center on Sanctions and Illicit Finance memorandum, January* (2018).

[66] D. Ermilov, M. Panov and Y. Yanovich. 'Automatic bitcoin address clustering'. In: *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE. 2017, pp. 461–466.

[67] S. Athey et al. 'Bitcoin pricing, adoption, and usage: Theory and evidence'. In: (2016).

[68] M. Lischke and B. Fabian. 'Analyzing the bitcoin network: The first four years'. In: *Future Internet* 8.1 (2016), p. 7.

[69] M. Paquet-Clouston, B. Haslhofer and B. Dupont. 'Ransomware payments in the bitcoin ecosystem'. In: *Journal of Cybersecurity* 5.1 (2019), tyz003.

[70] S. Meiklejohn et al. 'A fistful of bitcoins: characterizing payments among men with no names'. In: *Proceedings of the 2013 conference on Internet measurement conference*. 2013, pp. 127–140.

[71] T. Moore and N. Christin. 'Beware the middleman: Empirical analysis of Bitcoin-exchange risk'. In: *International Conference on Financial Cryptography and Data Security*. Springer. 2013, pp. 25–33.

[72] M. Campbell-Verduyn. 'Bitcoin, crypto-coins, and global anti-money laundering governance'. In: *Crime, Law and Social Change* 69.2 (2018), pp. 283–305.

[73] P. Reynolds and A. S. Irwin. 'Tracking digital footprints: anonymity within the bitcoin system'. In: *Journal of Money Laundering Control* (2017).

[74] P. Van Valkenburgh. *Tracking Bitcoin regulation state by state*. 2015.

[75] M. del Castillo. 'The 'Great Bitcoin Exodus' has totally changed New York's bitcoin ecosystem'. In: *New York Business Journal* (2015).

[76] M. Möser, R. Böhme and D. Breuker. 'An inquiry into money laundering tools in the Bitcoin ecosystem'. In: *2013 APWG eCrime Researchers Summit*. Ieee. 2013, pp. 1–14.

[77] S. J. Boxerman and M. F. Schwerin. 'Its Bark is Worse than Its Bit (e): Regulatory and Criminal Law Implications of Virtual Currency'. In: *Crim. Just.* 31 (2016), p. 10.

[78] FATF. 'FATF Guidance for a Risk Based Approach'. In: (2015). URL: https://www.amlc.nl/wp-content/uploads/2018/07/FATF-Guidance-RBA-Virtual-Currencies-2015.pdf.

[79] D. Kahneman, J. L. Knetsch and R. H. Thaler. 'Anomalies: The endowment effect, loss aversion, and status quo bias'. In: *Journal of Economic perspectives* 5.1 (1991), pp. 193–206.

[80] W. Samuelson and R. Zeckhauser. 'Status quo bias in decision making'. In: *Journal of risk and uncertainty* 1.1 (1988), pp. 7–59.

[81] V. Grapperhaus. 'From Zero To Hero: Identifying Vendor Characteristics that Impact Vendor Performance on Darknet Markets'. In: (2019).

[82] J. Magidson and J. K. Vermunt. 'Latent class models'. In: *The Sage handbook of quantitative methodology for the social sciences* (2004), pp. 175–198.

[83] J. K. Vermunt and J. Magidson. 'Technical guide for Latent GOLD 5.0: Basic, advanced, and syntax'. In: *Belmont, MA: Statistical Innovations Inc* (2013).

[84] M. Burnett. *Perfect password: Selection, protection, authentication*. Elsevier, 2006.

[85] W. Burr, D. Dodson and W. Polk. *Information security: Electronic authentication guideline nist*. Tech. rep. Technical report, Tech. Rep. Special Rep. 800-63, 2006.

[86] X. D. C. D. Carnavalet and M. Mannan. 'A large-scale evaluation of high-impact password strength meters'. In: *ACM Transactions on Information and System Security (TISSEC)* 18.1 (2015), pp. 1–32.

[87] T. Hunt. *Have I Been PWND*. 2020. URL: https://www.troyhunt.com/10b/.

[88] J. J. Dziak et al. 'Sensitivity and specificity of information criteria'. In: *Briefings in bioinformatics* 21.2 (2020), pp. 553–565.

[89] W. Meeus et al. 'Personality types in adolescence: change and stability and links with adjustment and relationships: a five-wave longitudinal study.' In: *Developmental psychology* 47.4 (2011), p. 1181.

[90] K. E. Masyn. '25 latent class analysis and finite mixture modeling'. In: *The Oxford handbook of quantitative methods* (2013), p. 551.

[91] L. M. Collins and S. T. Lanza. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Vol. 718. John Wiley & Sons, 2009.

[92] G. Notelaers et al. 'Measuring exposure to bullying at work: The validity and advantages of the latent

class cluster approach'. In: *Work & Stress* 20.4 (2006), pp. 289–302.

[93] A. K. Lenstra and E. R. Verheul. 'Selecting cryptographic key sizes'. In: *Journal of cryptology* 14.4 (2001), pp. 255–293.

[94] A. K. Lenstra. 'Key length. Contribution to the handbook of information security'. In: (2004).

[95] G. Tziakouris. 'Cryptocurrencies—a forensic challenge or opportunity for law enforcement? an interpol perspective'. In: *IEEE Security & Privacy* 16.4 (2018), pp. 92–94.

[96] C. Morselli, C. Giguère and K. Petit. 'The efficiency/security trade-off in criminal networks'. In: *Social Networks* 29.1 (2007), pp. 143–153.