Risky Business? Investigating the Security Practices of Vendors on an Online Anonymous Market using Ground-Truth Data

van de Laarschot, J.W.; van Wegberg, R.S.

Publication date
2021

Document Version
Final published version

Published in
Proceedings of the 30th USENIX Security Symposium

Risky Business? Investigating the Security Practices of Vendors on an Online Anonymous Market using Ground-Truth Data

Jochem van de Laarschot and Rolf van Wegberg, *Delft University of Technology*

https://www.usenix.org/conference/usenixsecurity21/presentation/van-de-laarschot

This paper is included in the Proceedings of the
30th USENIX Security Symposium.

August 11–13, 2021

978-1-939133-24-3

Open access to the Proceedings of the
30th USENIX Security Symposium
is sponsored by USENIX.

Risky Business? Investigating the Security Practices of Vendors on an Online Anonymous Market using Ground-Truth Data

Jochem van de Laarschot
Delft University of Technology

Rolf van Wegberg
Delft University of Technology

Abstract

Cybercriminal entrepreneurs on online anonymous markets rely on security mechanisms to thwart investigators in attributing their illicit activities. Earlier work indicates that – despite the high-risk criminal context – cybercriminals may turn to poor security practices due to competing business incentives. This claim has not yet been supported through empirical, quantitative analysis on ground-truth data. In this paper, we investigate the security practices on Hansa Market (2015-2017) and measure the prevalence of poor security practices across the vendor population ($n = 1,733$).

We create 'vendor types' based on latent profile analysis, clustering vendors that are similar regarding their experience, activity on other markets, and the amount of physical and digital items sold. We then analyze how these types of vendors differ in their security practices. To that end, we capture their password strength and password uniqueness, 2FA usage, PGP adoption and key strength, PGP-key reuse and the traceability of their cash-out. We find that insecure practices are prevalent across all types of vendors. Yet, between them large differences exist. Rather counter-intuitively, Hansa Market vendors that sell digital items – like stolen credit cards or malware – resort to insecure practices more often than vendors selling drugs. We discuss possible explanations, including that vendors of illicit digital items may perceive their risk to be lower than vendors of illicit physical items.

1 Introduction

Cybercriminals deploy security mechanisms that are intended to hinder investigators in their attribution efforts, making it difficult to link cybercriminal activity in the underground economy to an identity, location or machine [14, 64]. Since 'operational security' (OPSEC) techniques are frequently shared in the underground community [4, 56] and given the increasing amount of law enforcement scrutiny [15], we should expect that among cybercriminal entrepreneurs on online anonymous markets, poor security practices are rarely present.

However, there are numerous indications in earlier work that cybercriminals do not always achieve maximum security. Due to competing business incentives, criminals may turn to insecure practices that ease transacting illegal products or services. Here, we witness an inevitable trade-off between enhanced security and improved efficiency of operations [42]. 'Perfect security' therefore, is not economically viable. Like in the legitimate economy [54], security in the underground economy comes at a cost [53]. This leads us to wonder how prevalent poor security practices (or: "insecure practices") among online anonymous market vendors actually are.

While in earlier work attempts have been made to quantify insecure practices of cybercriminals trading in the underground economy, these only focus on a single, specific mechanism – e.g., PGP-adoption [49], consistent VPN usage [50] and the reuse of usernames and/or PGP-keys across different markets [59]. Moreover, it remains unknown *who* are behaving insecurely most often and we are left completely in the dark regarding *why*. All of these security mechanisms are designed and implemented to compromise the availability or usefulness of evidence to the forensic process [24]. Here, we should acknowledge that some market-based security mechanisms apply to every vendor. Rules, policies, content moderation, account verification and the mandatory use of cryptocurrencies and Tor-routing are examples of mechanisms that are imposed and enforced by the market [2, 7, 22, 62]. These mechanisms make up a form of 'extra-legal governance' that contributes to a more secure and trustworthy trading environment [13, 34]. Still, not all security mechanisms are introduced by the market administrators.

In this paper, we focus on specifically these mechanisms, as only unimposed practices can differ between vendors. To be precise, we analyze their password strength, password uniqueness, 2FA-usage, PGP-key adoption and key-strength, reuse of PGP-keys over multiple markets and the traceability of their cash-out to bitcoin exchanges. We capture these practices on a single market – Hansa Market, which was active from late 2015 to mid 2017. Seized data originating from the web server that hosted the market, has been made available to us by Dutch law enforcement. We combine the back-end database with three other data sources to measure the prevalence of poor security practices across the vendor population. In short, we make the following contributions:

- We present the first empirical, quantitative analysis leveraging unique ground-truth data to investigate vendor security practices on an online anonymous market.
- We measure the prevalence of poor security practices across different types of vendors on Hansa Market. For instance, we uncover that almost 40% of all vendors ($n = 1,733$) did not enable 2FA and find that at least 10% of vendors cash-out directly to mainstream bitcoin exchanges. Poor practices are also observed among the most successful vendors.
- We demonstrate that poor security practices do not occur at random. Rather counter-intuitively, vendors on Hansa Market selling digital cybercrime items are more likely to have insecure practices than vendors selling physical items – e.g., drugs.
- We discuss possible explanations for our findings, including that the perceived risk of transacting illicit digital items may be lower than the perceived risk for illicit physical items.

We structure the remainder of this paper as follows. Section 2 identifies the security practices of vendors on online anonymous markets. Section 3 elaborates on the data we analyze and our approach to measure the prevalence of insecure practices. In Section 4, we identify characteristics of vendors that can relate to their security practices and we cluster vendors with similar characteristics into distinct 'vendor types'. Section 5 shows how we capture the identified security practices, then we apply these measurements on the data to investigate the security practices across vendor types. We discuss possible explanations for differences in vendors' security practices as well as limitations and implications of our work in Section 6.We show how our work connects to related work in Section 7. Section 8 concludes.

2 Security practices on online anonymous markets

Online anonymous marketplaces take a prominent place in today's cybercrime ecosystem [25]. The first successful online anonymous marketplace – also referred to as a dark(net) market or cryptomarket – was Silk Road, which opened shop early 2011 [49]. This market introduced pseudonymous trading through an innovative platform only accessible through onion routing (Tor) and on which solely cryptocurrencies were accepted as mean of payment. By the end of 2013, Silk Road was shut down by law enforcement agencies. In the short period of time that Silk Road was active, it made its mark on the ecosystem as other initiatives successfully copy its business model to this day [49]. A decade later, some industry reports estimate the yearly revenue of all online anonymous markets combined, to be more than $790 million worth in cryptocurrencies [6]. First, predominantly illegal narcotics and prescription drugs were transacted on these marketplaces [7]. Nowadays, they also serve as one-stop shops for digital items – ranging from stolen credit cards to ransomware toolkits [58].

For those offering illicit substances or cybercrime items, online anonymous markets are attractive platforms to conduct their business on. The platforms provide contractual safeguards – like an escrow and review system – and anonymity enhancing functionalities that are superior to their alternatives [49, 58]. On top of that, vendors can employ additional security practices – ranging from authentication mechanism to obfuscating cash-out techniques. But, which practice makes perfect?

In this paper we aim to investigate which types of vendors pay more attention to their security than others. Thus, the security practices that the market imposes on *all* vendors, are not of our interest. Rather, we focus on the security practices that may differ between individuals. Leveraging earlier advances into 'deviant security' [1], we take the following six practices that impact the security of vendors into account. Later, in Sections 5 and 7, we will elaborate on the earlier work identifying these practices, and report how we are able to capture them in the data available to us.

Password strength. Although password authentication has been around for decades, people still have a tendency to choose predictable passwords [16], criminals included. This leaves them open to brute-force attacks that can give third-parties – e.g., rivals or law enforcement – access to their accounts. Which in turn, may lead to irreparable harm to business continuity.

Password uniqueness. A theoretically complex, but non-unique password can also be easily breached [3, 31]. Research suggests that password reuse is common, even among those who are security-aware [18, 61]. Additionally, databases of leaked passwords may include usernames or email addresses. Thus, password reuse can also lead to compromisability of users that operate on online anonymous markets.

2FA usage. Some markets provide users with the ability to enable two-factor authentication [55]. A 2FA-enabled login uses PGP as a verification mechanism, in which the user is challenged to decrypt a ciphertext that is encrypted with their public key [66]. This can only be achieved when in possession of the secret private key, making it an extra lock on the door.

[1] A term introduced by Van de Sandt [53] to describe the security of attackers or criminals, in contrast to that of defenders.

PGP usage. On online anonymous markets, PGP-encryption is the most used encryption protocol for secure communication [9]. Estimations show that in 2015, approximately 90% of market vendors listed a PGP-key on their profile [49]. The procedure to set up PGP is infamously known to be difficult to understand for the layman [48]. However, tutorials are widely shared within the cybercriminal community [56]. PGP-keys are based on a factoring problem, thus any key with a length of 2048+ bits is considered secure until the year 2030 [3,32].

PGP-key reuse. Unlinkability is an attribute of confidentiality [45]. When multiple usernames belonging to a single real-world entity are linkable, a security risk arises. Law enforcement may accumulate advanced knowledge on a persons behaviour and identity, potentially resulting in bringing this person to justice [28,53]. Still, some vendors that are active on multiple markets knowingly increase the linkability of their pseudonyms. A clear link between user accounts enables acquired reputation to be transferred to other markets [59]. The PGP-key listed by a vendor can create such a link. PGP-keys are suitable for signalling trustworthiness and transferring reputation, because their legitimacy can be verified by asking the other party to decrypt a text [51]. Using PGP-keys to link pseudonyms over different markets has been successfully used in prior work [5,49,51,59].

Traceability of cash-out. Bitcoin exchanges facilitate the conversion of bitcoins to fiat currency. Because of their often mainstream nature, these intermediaries can be subjected to regulation and subpoenaed for information on their clients [39]. This information may include full names and IP-, email-, or even residential addresses. Therefore, we consider it an insecure practice when the cash-out is traceable, thus when criminal earnings can be easily linked to an exchange. Cash-outs can be traced by analyzing the public ledger of the bitcoin blockchain.

We also identified security practices that did not have an *indisputable* effect on security. These either increase, or decrease security risks. These practices are excluded from further analysis, as it remains ambiguous how these affect the security of a user. First, the use of the market's auto-encryption functionality increases the 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 [4]. The same rationale applies to alternative messaging platforms, such as Jabber, ICQ, or Skype [1,50,57]. Second, regarding the mentioning of data minimization and data destruction practices in profile descriptions and listings [1,47], it remains unknown whether these practices are actually applied or merely mentioned. Lastly, with regards to shipping physical items over jurisdictional borders, Decary-Hetu et al. [10] argue that such cross-border shipments pose a security risk, while Van de Sandt [53] demonstrates that doing so creates information asymmetries between jurisdictions that benefit security. Also, many non-security related factors influence the decision to ship internationally [12,52].

3 Methodology

Now that we have a robust overview of pervasive security practices, we can turn to how our data sources enable us to measure these practices. In this section we elaborate on our data, discuss the ethics of using seized data, provide descriptives and present our measurement approach.

3.1 Data

In this paper we leverage four data sources: 1) the Hansa Market back-end database, 2) the 'Have I Been PWND' password database, 3) the database of the *Grams* search engine and 4) the *Chainalysis* blockchain analysis service. We describe these data sources, one by one, below.

Hansa Market. There have been several law enforcement interventions that directly targeted online anonymous markets and resulted in take-downs. In *Operation Bayonet* (2017) two of the largest online anonymous markets were shut down [15]. First, the Federal Bureau of Investigation (FBI) took down AlphaBay on July 5th 2017. Thousands of AlphaBay users in search of a new platform to continue their business on, migrated to Hansa Market. However, those who did, fell right into a trap. As this market was already infiltrated and under full control of the High Tech Crime Unit of the Dutch Police (NHTCU). This unit operated – as they also had taken over the admin accounts – the market from June 20th until they shut down the market on July 21st. In this period of time, the NHTCU even turned off the encryption of personal messages, and the hashing of passwords [20]. The sting operation not only resulted in the collection of valuable data such as names and street addresses of buyers, it also disrupted the ecosystem by causing distrust [59].

When Hansa Market was infiltrated by NHTCU, they first migrated the web servers in order to operate the market themselves and thereafter seize all contents. After the market was shut down, Dutch law enforcement shared the back-end data with other law enforcement agencies [15] and allowed us restricted access for in-depth analyses. In Section 3.3 we will extensively report on the subsets of back-end data, as we – despite the nature of our access – want to be as transparent as possible. Next, we will discuss the ethics of using the seized back-end of Hansa Market in greater detail in Section 3.2.

Have I Been PWND. The 'Have I Been Pwnd?' (HIBP) service accumulates login credentials found in hundreds of breached databases. On this website, users may search whether their credentials were compromised in any (known) data leaks. The website is regularly updated with new data breaches. The full database of SHA1-hashed passwords is publicly available [29]. At the time of our analysis, the most recent version (v6) of the database contained more than 10 billion leaked passwords, of which 573 million are unique [30].

Grams. *Grams*, a "Google for darknet markets" [65], made it possible to search through various markets at once. The search engine indexed listings and vendors through custom API-calls to the most popular markets. In doing so, it allowed users to locate their favorite vendors on multiple markets using a vendor's public PGP-key. *Grams* shut down in December 2017. We were allowed to match records from the Hansa back-end with a copy of the database that was acquired by law enforcement before its administrator announced the shutdown of the search engine on Reddit [19].

Chainalysis. Raw bitcoin blockchain data consists of logs of transactions between bitcoin addresses. It does not include any context that can be used to make sense of this data. Commercial and non-commercial tools are available that do provide this context [27]. These tools enable researchers and law enforcement investigators to track monetary flows between distinct entities [26]. In this paper, we use the *Chainalysis* blockchain analysis service. This service mainly makes use of a co-spend clustering heuristic. Co-spending occurs when two addresses engage in a single outgoing transaction [23]. Two co-spending addresses are likely to belong to the same real-world entity. By monitoring co-spending, *Chainalysis* is able to estimate which bitcoin addresses are controlled by – for example – bitcoin exchanges.

3.2 Ethics

The Hansa Market back-end is similar in nature to that of seized data used in earlier work [21, 37, 44]. Operating in conjunction with applicable laws and regulations, the Dutch authorities were allowed to seize the Hansa Market infrastructure. Despite the legal nature of the seizure, using this data for research purposes raises some ethical issues, which we discuss below.

In order to protect the privacy of Hansa Market users, we took great care not to analyze personally identifiable information (PII). Our data subset was limited to contain only data vital to our research, and stripped of all PII – usernames were replaced with unique IDs, private message logs excluded and plaintext passwords hashed. When our analysis did involve PII – i.e., to measure password strength, as this can only be done using plaintext passwords – we asked law enforcement to run our code and return the output. With this approach, the data was cleared by law enforcement authorities for the purpose of this research in accordance with Dutch privacy law. To mine the data, whilst not compromising any present and future investigations, we only had controlled on-premise access to subsets of the data.

Next, we believe that our analysis does not create further harm as we did not partake in or stimulate any criminal business model – by purchasing criminal services, or in any other way contribute to the criminal enterprise.The authors and involved law enforcement professionals believe the benefits of a comprehensive understanding on 'deviant security', outweigh the potential cost of making this kind of knowledge more widely known [53]. More so, as the anatomy and economics of online anonymous market are already well-documented in earlier work [7, 49, 58].

Finally, this study has been conducted with the prior approval of, and in collaboration with, Dutch law enforcement and public prosecutors. Note however, as we will cover in Section 3.3, that over 87 percent of users were inactive buyers for whom we have no evidence of illegal behavior. One should not, therefore, conclude that the majority of subjects were engaged in illegal behavior or that this was a factor in deciding to use their data for our research. Yet, other information in the back-end data can be directly used in police investigations. However, due to the extensiveness of the data, it also provides unique, behind-the-scene insights into how market users operate. Note, that providing evidence of any kind for continued law enforcement efforts is not the purpose of this study.

3.3 Hansa Market descriptives

Back-end. The back-end database of Hansa Market consists of more than a hundred data tables. Jointly, these give an insight into the complete administration of the market. Due to the classified nature of the data, we can not disclose the data structure in detail. However, we can qualitatively describe the data tables that are used in this paper.

The first data table that is central in our analyses contains information directly related to the user administration. Here, we find the registration dates of users, which users are registered as vendors and a field in which the public PGP-key of each user is stored. This table also lists whether a user enabled two-factor authentication or not.

Second, we use a data table that stores information related to the advertisements, or 'listings'. It includes the product class or advertisement category – e.g., credit cards – to which a listing belongs, the description of the listing and the vendor ID – which links listings to vendors. Additionally, this data table contains a field indicating whether an item should be physically shipped or digitally delivered. We reason that the front-end of the market used this field to determine whether the option of shipping (costs) should be presented and whether a shipment ID should be generated in case the item is sold. Shipping costs and a shipment ID therefore do not apply to digital items.

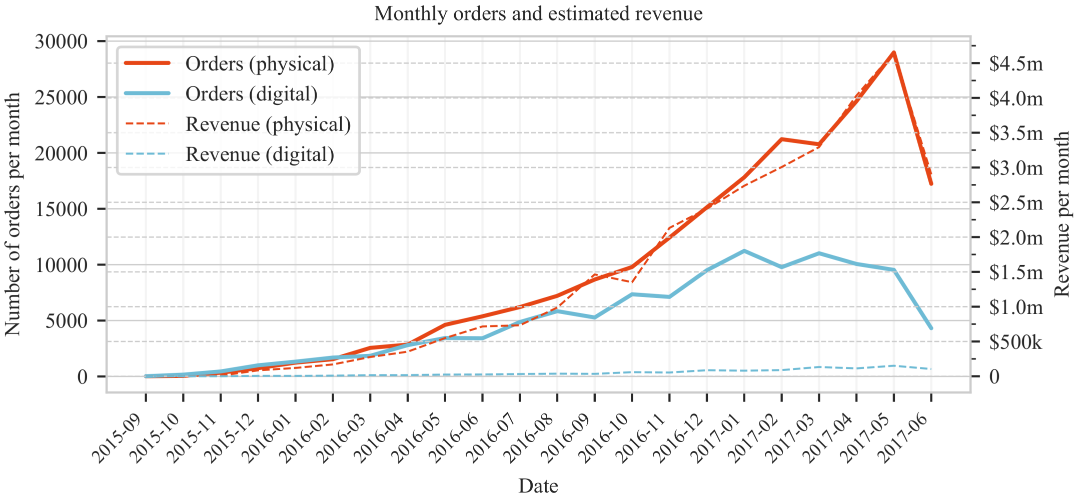

Figure 1: Monthly orders and estimated revenue, per type of product.

The third table that we make use of, entails data that keeps track of the orders that were placed by buyers. Information on order status, order date and also the bitcoin address to which the buyer's payment is transacted, is listed here. By default, Hansa Market purged orders older than 180 days. Despite the fact that vendors could extend this limit, this auto-purging feature results in missing data. Fortunately, Dutch law enforcement discovered old back-ups that the administrators of Hansa Market made. Consequently, parts of the presumed-to-be-deleted order data could be resurrected. Even then, not all orders could be recovered. We addressed this problem by reconstructing the lower-bound number of sales using the number of feedbacks given – as these were not purged. In earlier work, feedbacks proved to be an accurate proxy for transactions [49, 57, 58]. The feedbacks are stored in a different data table that includes an order ID, a rating and the price paid. As bitcoin payout addresses are only listed in the order table – and its back-ups – not all orders have complete payment information. Therefore, the data available to us does not include all bitcoin addresses of all vendors.

The last data table we utilized, comprised the connection logs that registered all logins of Hansa Market users. Since Dutch law enforcement gained complete and unrestricted access, they were able to alter the configuration of the market. They modified the market in such a way that passwords were saved as plaintext in the markets' connection logs [20]. As a result, plaintext passwords are available for all users that logged into their account during the last month the market was operational. Later on, as described above, these password were hashed so to exclude any PII in the subset of data we use in this paper.

Descriptive statistics. The first transactions on Hansa Market date from mid-September 2015. These early transactions entail dummy transactions between administrator accounts. We removed these transactions from the data and we consider the market to be publicly active from September 25th 2015 onwards. For most of the time that Hansa Market was in operation, no large scale user migrations – for instance as a result of law enforcement interventions – occurred. At the end of *Operation Bayonet* however, the coordinated shutdown of AlphaBay led to an enormous influx of new users on Hansa Market. We cannot directly compare these large amounts of migrated users to the existing Hansa user base. Since we can only analyze the Hansa data, all former AlphaBay users that fled to Hansa would seem 'inexperienced newcomers' to us. Their past 'career', including their reputation and experience, forms a blind spot. We therefore decide to discard all new users, orders and transactions made after June 20th 2017 – the day that law enforcement took over the administration of the market.

To illustrate the amount of funds that flowed through the market, we estimate the generated revenue. We convert the order price in bitcoins to dollars using the exchange rate at the time that the order was placed. In the defined time period – which excludes the last month in which the number of sales surged – we estimate over $33M is generated on Hansa Market. We plot the monthly revenue and the monthly number of orders in Figure 1.

Next, we distinguish three types of users. *(i)* Vendors – users that sold at least one item, or that have at least one feedback. Remarkably, this includes 160 vendors that are not registered as vendors in the Hansa administration. From this we conclude that some users are former vendors that decided to downgrade their accounts to regular member accounts, perhaps motivated to reclaim the vendor bond. *(ii)* Active buyers – users that bought at least one item, or that provided feedback to an item. *(iii)* Inactive buyers – users that registered an account on the market, but who did not buy or sell any items. Hopefully, this includes all security researchers.

In Table 1, we list the most important descriptive statistics of the market. We also describe how many orders are reconstructed using feedbacks, due to purged order-related data.

3.4 Approach

Our approach to measure the prevalence of poor security practices across different types of vendors consists of two steps:

1. We identify characteristics of vendors that can explain their security practices and cluster vendors that have similar characteristics into distinct ‘vendor types’ using latent profile analysis.
2. We capture the security practices we identify in Section 2 in our data and measure the prevalence of poor security practices across vendor types and compare these with the practices of buyers.

We visualize our approach in Figure 2.

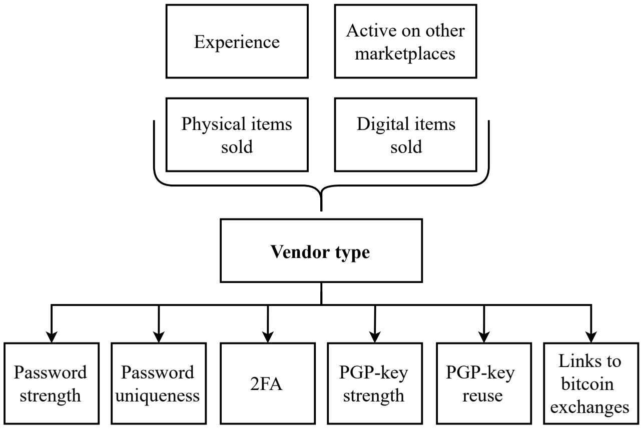

Figure 2: Research model

4 Vendor types

We now turn to identify characteristics of vendors that can explain their security practices and find latent groups of vendors based on these characteristics.

4.1 Vendor characteristics

Based one earlier work, we expect the following vendor characteristics to relate to their security behavior.

Experience. According to Van de Sandt [53], the security practices of cybercriminals are related to their experience. New security developments may be ignored by relatively inexperienced individuals, who do not become aware of these (updated) security mechanisms available [55]. For every vendor, we determine their *experience* on the market by calculating the amount of days between a vendor’s first and last sale [57]. To account for any experience gained on other markets, we also include the binary characteristic *active on other markets*. This characteristic is based on whether a vendor ‘imported’ a reputation, or rating, from another market through the reputation-import functionality of Hansa Market. Evidently, a vendor with an imported reputation must be active on at least one other market.

Table 1: Hansa market descriptive statistics

Users	Vendors	1,733
	Active buyers	50,433
	Inactive buyers	365,144
Listings	Physical items	67,905
	Digital items	38,729
Orders	Physical items	209,411
	of which reconstructed	130,420
	Digital items	112,046
	of which reconstructed	75,236
Est. revenue	Physical items sold	$32M
	Digital items sold	$1M

Next to experience, a relation between ‘business success’ and security is expected since *(i)* investments in security can be costly – in terms of time, knowledge, money – and *(ii)* increasing profits result in higher risks to security [53]. It is important to note that drug trade is set in a different criminal context than the trade of cybercrime items. Van Hardeveld [55] found indications that vendors with a traditional (offline) criminal background are more likely to make mistakes in their digital security. With regard to business success, we therefore differentiate between the number of physical items sold and the number of digital items sold.

Physical items sold. In the order data table, we count the amount of orders that are physically shipped for each vendor. Most physical items that are sold on online anonymous markets are types of drugs – e.g., cocaine, cannabis, MDMA, heroin or other psychoactive substances [7, 49].

Digital items sold. Likewise, we are able to count the digital items each vendor sold. Digital items include a great variety of products – e.g., botnet related items (tutorials, source codes, DDoS services), hacked accounts, fake IDs, databases of e-mail addresses, passwords and personally identifiable information, exploits and malware, ransomware, credit card details and listings that aim to recruit money mules [58].

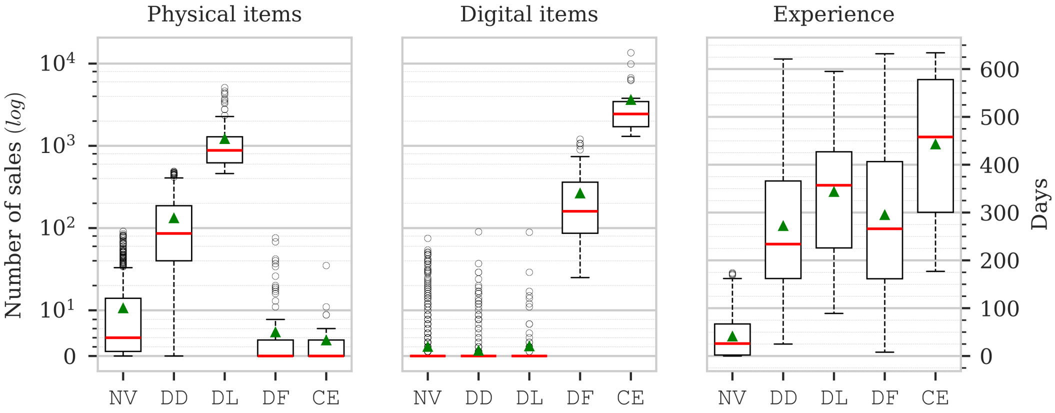

Figure 3: Distribution of *experience* and the number of *physical* and *digital sales* per vendor type. Abbr. and percentage active on other markets: NV = `Novices` (40.2%), DD = `Drug Dealers` (69.0%), DL = `Drug Lords` (78.2%), DF = `Digital Fraudsters` (58.3%), CE = `Cybercrime Elites` (73.9%).

4.2 Latent Profile Analysis

As shown in Figure 2, we grasp multiple vendor characteristics in a single variable by allocating each vendor a vendor type. A vendor type is a cluster of vendors with similar *experience*, *activity on other markets*, *physical items sold* and *digital items sold*. We create these clusters through a Latent Profile Analysis (LPA). This clustering algorithm maximizes the homogeneity within clusters and heterogeneity between clusters and takes data on any measurement level as input [35]. In recent work, Van Wegberg et al. [57] analyze data similar to ours. Here, the authors also create clusters of vendors using vendor characteristics. They show that vendors operating on AlphaBay are best clustered into five groups. We therefore estimate models with 1 to 5 clusters using the LatentGOLD statistical software [60]. Then, we select a fitting number of clusters through both evaluating the global fit via Bayes' Information Criterium (BIC) and the local fit through assessing the bivariate residuals (BVRs).

In Table 2, we present the results of the clustering algorithm. As indicated by the low BIC value, we conclude that the 5-cluster model has the best global fit to our data. The low total BVR value shows that the 5-cluster model provides the best local fit as well. The non-significant BVR-value indicates that there is no association between *physical items* and *digital items sold* in the 5-cluster model. Thus, the 5-cluster model separates these characteristics particularly well. Last, we perform pairwise Kruskal-Wallis H tests to assess whether the means and medians of the vendor characteristics are significantly different between vendor types. This is the case for all relevant [2] pairwise comparisons.

It is important though, not to evaluate the model only through numerical considerations [8, 36, 38]. We confirm that the 5-cluster model is a parsimonious model that clearly differentiates between vendors that specialize in selling physical or digital items. The 5-cluster model is easy to interpret in the context of this work – which we will do next – and the sample sizes are sufficiently large.

4.3 Resulting vendor types

Based on the distributions of vendor characteristics and what product categories are dominant within clusters, latent profiles emerge and each vendor is given a 'vendor type'. We visualize the clustering results in Figure 3.

The `Novices` ($n = 988$) have the lowest amount of physical and digital sales of all vendor types and have limited experience on the market. About 80% of the products they did sell are drugs, although some vendors made (few) digital sales as well. No vendors with more than 100 physical or digital sales are present in the `Novices` cluster. In contrast, `Drug Dealers` ($n = 509$) have far more physical sales, experience and activity on other markets compared to the `Novices`. More than half of the vendors identified as `Drug Dealers` have been active on Hansa Market for over 230

Table 2: Clustering fit

Model	$BIC(L^2)$	Non sig. BVR*	Total BVR
1 cluster	1761741	n/a	384060
2 clusters	951620	n/a	52944
3 clusters	570490	n/a	13138
4 clusters	385285	n/a	6806
5 clusters	294230	*physical-digital*	3861

* BVR > 3.84

[2] Clusters 1-3 describe vendors that specialize in selling physical items: vendors in these clusters do not differ significantly in terms of their digital sales. Likewise, clusters 4 and 5 consist of vendors that specialize in digital items, who do not differ significantly in terms of their physical sales.

Table 3: Security practices, earlier work identifying these practices, measurement and data source(s) leveraged.

Security practice	**Earlier work**	**Measurement**	**Data source(s)**
PW strength	Van de Sandt [53]	strength estimation using `zxcvbn` [63]	Hansa
PW uniqueness	Van de Sandt [53]	matching SHA1 hashed passwords	Hansa, HIBP database
2FA usage	Van Hardeveld [55]	observing binary indicator in data	Hansa
PGP usage	Soska & Christin [49]	PGP-key strength, $\leq$2048 or $>$2048-bits	Hansa
PGP-key reuse	Van Wegberg [59]	matching PGP-keys	Hansa, Grams
Traceability of cash-out	Van de Sandt [53]	analyzing transactions from payout addresses	Hansa, Chainalysis

days. Of the products they sold, 98% are drugs. The `Drug Lords` ($n = 110$) do not differ much in terms of experience in comparison to `Drug Dealers`, but have extreme amounts of physical sales and more activity on other markets. All sales (100%) are drugs related.

The following two clusters of vendors thrive in digital sales rather than physical sales. First, `Digital Fraudsters` ($n = 103$) have varying experience in selling fraud-related items. Yet, all vendors in this cluster have at least 15 sales in the digital domain and about 75% made more than 100 digital sales. Some vendors with mainly digital sales, also made a handful of physical sales. Second, `Cybercrime Elites` ($n = 23$). This small cluster of very successful vendors of digital items clearly trumps the `Digital Fraudsters` in terms of sales and are the most experienced groups of vendors on the market.

5 Security practices

Following our discovery of distinct vendor types, we can now investigate how each of them handle their security. In this section we first define how we capture security practices identified earlier. Then, we apply these measurements on the data and elaborate on the security practices for each vendor type. For the purpose of clarity, Table 3 provides an overview of the six security practices, earlier work identifying these, our measurement and data sources leveraged.

5.1 Measuring security practices

We capture the six security practices identified in Section 2 as follows.

Password strength. The strength of a vendor's password is captured by evaluating the estimated amount of guesses it will take to crack the password. `zxcvbn` [63] is a password strength estimation tool that outputs the estimated number of guesses, given a plaintext password. The `zxcvbn` tool recognizes common words and matches different types of patterns, such as repeated letters, word reversal, common substitutes of letters and keyboard sequences. The order of magnitude of the amount of estimated guesses it will take to brute force a password indicates password strength. When `zxcvbn` estimates that more than 10^{10} guesses are needed to crack a password, the password is considered 'very unguessable', $< 10^8$ equals 'somewhat guessable' and $< 10^6$ is 'very guessable' [67].

Only the passwords of the vendors that logged into the market during the last month it was operational are available. In total, we analyzed the passwords of 1,081 vendors ($\approx$ 62.4%) [3]. We find that on average, the password strength is $10^{14.7}$ estimated guesses, the median password strength is $10^{10.5}$ guesses.

Password uniqueness. We capture the uniqueness of users' passwords by matching the SHA-1 hashes of the available Hansa passwords with the SHA-1 hashes from the HIBP password database. Out of the 1,081 vendors of whom a password is available, 185 vendors (17.1%) logged in with a password that we matched in the HIBP database. Given the high security risks of using a non-unique password – i.e., access to user account(s) and potentially de-anonymization – this number is larger than we initially expected.

2FA usage. The data table that stores the user administration, includes a binary variable that indicates whether 2FA is enabled or not. Information on 2FA-usage is available for all vendors ($n = 1,733$). Of the total vendor population, only 60.5% ($n = 1,049$) protected their accounts with this additional layer of security.

PGP usage. Hansa vendors could publish their public PGP-key on their profile. From this key, we extract the creation date and the key-length using a Python implementation of GNU Privacy Guard (GnuPG). Some keys had peculiar lengths of 1023-, 2047- or 4095-bits. Such aberrant key sizes are the result of how RSA keys of length N are generated. Because N is generated by multiplying two randomly chosen primes $p \cdot q$ of length $N/2$, a small probability exists that a key of $N - 1$ is generated. Although not mandatory as per the RSA specification [41], some implementations of RSA correct for

[3] Note, like we stated in Section 3.2, we did not perform this analysis ourselves as we did not have any access to PII. To capture password strength, we asked law enforcement to run `zxcvbn` on the available plaintext passwords and return the output. We link this output to the unique ID's that were used to replace usernames, as this prevents us from analyzing any PII.

this. The atypical keys are – in practice – equally secure to their more common counterparts, so we replace all uncommon key lengths with the commonly found key lengths.

The PGP-adoption among vendors is high. Only 5 vendors do not have a PGP-key listed. It could be, that these vendors removed their PGP keys from their accounts after they stopped trading. Weak keys ($\leq$ 1024 bits) are observed for only 9 vendors. Even by 2015's standards, such key lengths are considered not to be sufficient [32]. We investigate the relation between the extracted creation date of the key and its key strength. No trend is apparent in which younger keys are stronger than older keys.

PGP-key reuse. To capture which Hansa vendors explicitly chose to use different PGP-keys on the markets they operate on, we focus on a subset of the Hansa data. We only consider the vendors that imported their reputation – of these vendors, we can be sure that they operated on more than one market. This method decreases the likelihood of including imposter accounts in our analysis. Using this subset, we investigate which PGP-keys are also listed in the database of the *Grams* darknet market search engine. If we match a PGP-key in both data sets, we check if – according to *Grams* – the PGP-key links to other markets than Hansa. If we find no match, or a match that links only to a Hansa account, we infer that a vendor explicitly chose to create new PGP-key(s) for its other account(s). A match that links to a non-Hansa account, shows us that this vendor reuses its key on at least one other market.

Figure 4 displays how the following groups overlap: vendors with a PGP-key listed ($n = 1,728$), vendors known to be active on other markets ($n = 908$) and PGP-keys that are listed on any other market than Hansa in the *Grams* search engine ($n = 902$). From this figure, we conclude that there is a group of vendors ($n = 265$) who are active on other market(s), but whose PGP-keys could not be matched. Surprisingly, there is also a group ($n = 259$) who did not use the import functionality but whose PGP-keys are matched in the Grams data.

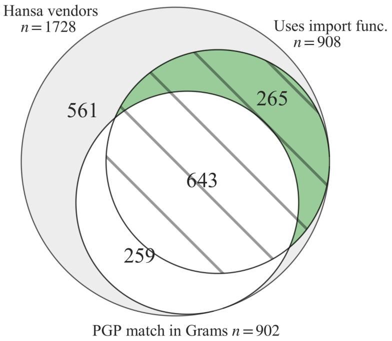

Figure 4: PGP-key matching.

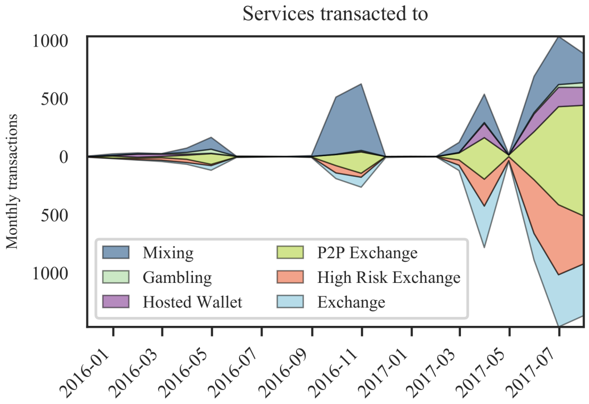

Figure 5: Type of wallets vendors transact their revenue to.

Traceability of cash-out. Using the *Chainalysis* blockchain analysis service, we capture the traceability of vendors' cash-outs. Specifically, we analyze which vendors transact their profits directly to mainstream bitcoin exchanges or hosted wallets – i.e., entities that can be subjected to subpoenas for information on their users. A vendor that transacts profits made by doing business on an online anonymous market directly to an exchange or hosted wallet, creates an indisputable transactional link between – most likely – criminal activities and PII collected by subpoenable entities. Thus, this practice is very insecure.

We analyze the 19,238 unique bitcoin payout addresses that are stored in the Hansa back-end database. Of these, 2,680 addresses ($\approx$ 14%) could be directly attributed to clusters of addresses that *Chainalysis* identifies with known service wallets, such as central exchanges, peer-to-peer exchanges and bitcoin mixers. Thus, the majority of the bitcoin addresses that vendors cash-out to, are unknown services. We expect that most vendors (first) have their payouts transacted to private (hardware) wallets or to mixing services that are not identified by *Chainalysis*.

We visualize the known services vendors directly cash out to in Figure 5. Exchanges that are reluctant in gathering data on its users, or those that do not perform any identity checks, are not likely to respond adequately to law enforcement subpoenas. *Chainalysis* labels such exchanges as 'high risk exchanges' (Figure 5). Next to this type of exchange, cybercriminals regard peer-to-peer (P2P) exchanges as safe-havens due to minimal identity verification [4].

5.2 Security practices across vendor types

As we have an overview of the prevalence of poor security practices in the total vendor population, we now analyze

[4] The data shows that LocalBitcoins.com is the most used P2P-exchange. In the 2015-2017 time frame, LocalBitcoins did not verify identities. At the time of writing, steps have been taken to adhere to AML regulations [33]

Table 4: Number of vendors within each vendor type that exhibits secure (y) or non-secure (n) behavior. For each security practice (unique pw, 2fa usage, *etc.*) applies: when, according to FDR-BH adjusted z-tests ($\alpha = 0.05$), two proportions of secure/non-secure behavior are significantly *different* between vendor types, this pair is annotated with the *same* sign.

	UNIQUE PW		2FA		2048+ PGP		NO KEY REUSE		NO BTC LINK	
	y/n	$sec.\%$	y/n	$sec.\%$	y/n	$sec.\%$	y/n	$sec.\%$	y/n	$sec.\%$
`Novices`	395/98	80.1^{*}	542/446	54.9^{*}	466/520	$47.3^{\dagger\circ}$	121/275	30.6	678/38	$94.7^{*\times}$
`Drug D.`	342/52	$86.8^{*\times}$	359/150	70.5^{*}	273/233	$54.0^{\dagger\circ}$	102/247	29.2	448/57	$88.7^{*\times}$
`Drug L.`	82/11	$88.2^{\dagger}$	90/20	$81.8^{*\times}$	62/48	$56.4^{*\times}$	22/64	25.6	86/23	78.9^{*}
`Dig. Frd.`	57/21	$73.1^{\dagger\times}$	45/58	43.7^{*}	30/73	$29.1^{*\dagger}$	15/45	25.0	78/20	$79.6^{\times}$
`Cyb. Elt.`	20/3	87.0	13/10	$56.5^{\times}$	5/18	$21.7^{\circ\times}$	5/12	29.4	12/11	$52.2^{*\times}$

the security practices across each vendor type. Because we face large differences in sample sizes – for example, 988 vendors are identified as `Novices`, while there are only 23 `Cybercrime Elites` – we perform extensive statistical testing. This ensures that the differences we observe are not a mere artefact of differences in sample size.

For each security practice, we first perform an *omnibus*-test to find out whether there are any differences between vendor types at all. If there are, we perform a post-hoc test. This test specifies *which* vendor types significantly differ from each other on security practices. Omnibus tests are more powerful compared to pairwise post-hoc tests. It is plausible that an omnibus test gives a significant result, while all pairwise post-hoc tests do not. Vice-versa, this is not the case.

Password strength. A password is not available for every vendor. Still, all vendor types remain well populated: we perform our analysis on 493 `Novices` (-50.1%), 394 `Drug Dealers` (-22.6%), 93 `Drug Lords` (-15.5%), 78 `Digital Fraudsters` (-24.3%) and all `Cybercrime Elites` ($n = 23$). We show the distribution of password strength per vendor type in Figure 6. A Kruskal-Wallis H test ($p < 0.0001$) indicates that there are significant differences in password strength between vendor types.

To learn which vendor types differ significantly in password strength, we perform a Dunn post-hoc test in which the significance levels are adjusted for multiple comparisons through FDR-BH adjustment [11]. With $\alpha = 0.05$, we find that only three pairs of vendor types differ significantly in password strength, so we interpret the results with care (Figure 6). We conclude that the password strengths of `Drug Lords` and `Drug Dealers` differ significantly with those of `Digital Fraudsters` and that there is a significant difference between `Drug Dealers` and `Novices`. Regarding the difference between `Drug Lords` and `Novices`, there is slightly more statistical uncertainty ($p = 0.0691$).

We take into consideration that simpler passwords might be used by those vendors that enabled 2FA. Generally, this does not seem to be the case. 2FA-usage correlates *positively* with password strength, as assessed by a Spearman rank-order correlation ($r_s = 0.219, p < 0.0001$). This indicates that vendors do not tend to compensate relatively poor passwords with the additional layer of security that 2FA adds.

Password uniqueness. In Table 4 we show the amount of vendors that made use of a unique password per vendor type. Since there are significant differences between vendor types, as confirmed by a χ^2-test ($p = 0.0064$), certain types of vendors make this security mistake more often than others. To find out how vendor types relate to each other in terms of password uniqueness, we performed a pairwise post-hoc z-test of proportions with FDR-BH correction (Table 4). It is evident that, again, the security practices of `Drug Lords` and `Drug Dealers` are better than those of `Digital Fraudsters`. The `Novices` perform relatively poor as well, although only the difference with `Drug Lords` is significant. While `Cybercrime Elites` score quite high, their score does not differ significantly from the other vendor types.

2FA usage. With respect to 2FA-usage, vast differences exist between vendor types (Table 4). We see that, again, `Drug Lords` and `Drug Dealers` chose this secure option the most often, whereas `Digital Fraudsters` go for

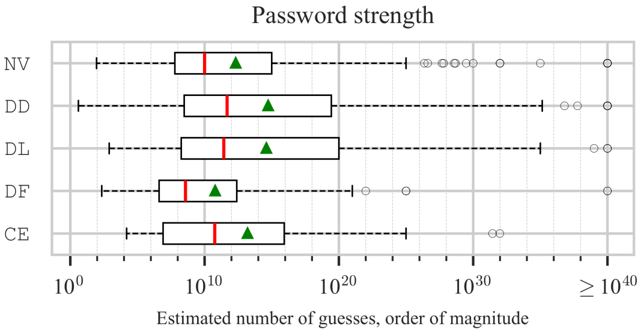

Figure 6: Distribution of password strength per vendor type. Medians are displayed in red, green triangle indicates the mean password strength.

security the least often. Within the groups of Novices and Cybercrime Elites, about half of the vendors enabled 2FA. So, it appears that especially experienced vendors that sell larger amounts of drugs, are willing to go through the hassle of verifying their log-ins through PGP.

PGP usage. Considering that the security benefit of any key stronger than 2048-bits is negligible until the year 2030, we initially expected that key sizes are chosen 'randomly' or according to one of the PGP-tutorials found on underground discussion fora. However, our analysis suggests the contrary (Table 4). Extremely secure keys are more often found among Drug Lords and Drug Dealers. The Cybercrime Elites and Digital Fraudsters have extremely secure keys the least often. Additionally, none of the Drug Lords use weak keys. Although careful interpretation of the results is necessary due to differences in sample sizes, these numbers suggest that chosen key strengths are not a coincidence. On Hansa Market, some types of vendors feel the need for 'extremely secure' keys, while other types of vendors tend to settle for regular 'secure keys'.

PGP-key reuse. We analyze the number of matched PGP-keys per vendor type, in which we only consider the vendors that are known to be active on multiple markets. It is clear that the differences between between vendor types are modest (Table 4). A χ^2-test confirms this. The proportion of vendors per cluster that could not be matched, does not significantly differ between vendor types ($\chi^2 = 1.409, p = 0.8425$).

Traceability of cash-out. Due to missing data, we did not find at least one payout address for every vendor. Only vendors associated with at least one payout address are included in this analysis. Therefore, this analysis entails 716 Novices (-27.5%), 505 Drug Dealers (-0.8%), 109 Drug Lords (-0.9%), 98 Digital Fraudsters (-4.9%) and all Cybercrime Elites ($n = 23$).

We analyze how many vendors per vendor type exhibit the insecure practice of transacting directly to mainstream bitcoin exchanges or hosted wallet. We observe large differences between vendor types regarding the proportion of vendors that are directly linked to these exchanges and hosted wallets (Table 4). This analysis yields two surprising conclusions. First, Novices show the most secure behaviour - that is to refrain from transacting their profits directly to exchanges or hosted wallets. Second, nearly half of the Drug Lords, who proved to be very security-aware otherwise, have transact their profits directly to mainstream exchanges or hosted wallets. This creates a serious risk to their security. Furthermore, we see that many Digital Fraudsters and especially Cybercrime Elites do not bother to obfuscate their criminal profits.

5.3 Security practices of buyers

To see if vendors are any different than other users on the market, we compare the security behavior of buyers and vendors regarding the strength and uniqueness of their passwords, 2FA usage, PGP-adoption and if applicable, chosen PGP-key lengths [5]. Data on PGP-key reuse and traceability of cash-out, are not available for buyers. From this comparison it becomes apparent that – on average – vendors do have better security practices than buyers (Table 5). Or put differently, buyers even have worse security practices compared to vendors. Note however, that a PGP-key is needed to be able to use 2FA. Since only 12.1% of the buyers have a PGP-key listed, the low 2FA usage is partly explained by the low PGP-adoption of buyers. When we only consider the users with a PGP-key listed, the proportion of users with extremely secure PGP-keys (2048+ bits) does not differ much between vendors and buyers.

Table 5: The security practices of buyers and vendors compared. 'PW strength' states the average/median of the estimated number of guesses a brute-force attack takes. '2048+' states the % of listed keys that are 2048+ bits.

	PW STRENGTH		PW U.	2FA	PGP	
	μ	*med.*	%y	%y	%y	2048+
Vendors	$10^{14.7}$	$10^{10.5}$	82.9	60.5	99.7	48.4
Buyers	$10^{9.8}$	$10^{8.1}$	69.2	3.4	12.1	47.2

6 Discussion

In this section, we will first discuss possible explanations of our findings. We next discuss the inherent limitations that arise from our measurement approach and the data sources we use. Last, we will touch upon the implications of our findings.

6.1 Possible explanations of our findings

We found latent groups of vendors that are similar regarding their experience on the market, amounts of physical and digital items sold and activity on other markets. Subsequently, we measured and compared the security performance of these vendor types. Doing so, we uncovered surprising patterns in security practices of the criminal entrepreneurs that operate on online anonymous markets. Clusters of vendors that specialize in selling digital items, such as hacked accounts, credit card details and databases of PII, make 'mistakes' in their digital security the most often. Counter-intuitively, successful drug dealers – i.e., Drug Dealers and Drug Lords – tend to have the best digital security. Especially Drug Lords use

[5] As we discussed in greater detail in Section 3.2, this data was gathered as part of a lawful investigation. Using it is in accordance with prior practice [18, 34, 40].

complex and unique passwords. Additionally, they tend to protect their accounts with 2FA and they encrypt their communications using extremely secure PGP-keys. How can we explain this pattern? Why do drug dealers have better digital security than cybercriminal entrepreneurs?

One possible explanation is that on Hansa Market, vendors of drugs perceived their risk to be higher than vendors of digital items. The former may anticipate a greater probability that their activities will draw law enforcement action. Punishment for drug vendors may also be more severe – drug offenses are punishable by death in at least 35 countries [17]. The nature of physical sales can also generate more evidence, such as addresses of buyers and shipping information stored in databases of postal services.

Assessment of risk may be subjective [53], which may explain some differences between and within groups. For example, even `Drug Lords` do not behave consistently with regards to password hygiene or preferred PGP-key lengths. Furthermore, misconceptions – such as a belief that bitcoin transactions are completely anonymous [46] – may impact decisions. Misconceptions or behavioral pitfalls [55] may have led `Drug Lords` to make choices like cashing out to well-known exchanges.

Possible explanations beyond risk perception exist for the practices and differences observed. For example, revenue for physical goods is much greater than for digital goods (see Figure 1). Vendors with higher earnings could be more likely to hire specialized experts to manage postings and security. A user study of vendors could help confirm the source of the differences in observed practices.

6.2 Limitations

First, our research focuses on a single online anonymous market: Hansa Market. Naturally, this is a limiting factor in our ability to generalize our findings. After the shutdown of Hansa Market and AlphaBay in Operation Bayonet, the continued scrutiny by law enforcement might have resulted in an increased security-awareness among those conducting business on online anonymous markets. This would mean that we observe a time frame wherein vendors are operating less securely, compared to today. Future research should try to replicate our analysis and see if security practices in the underground economy have evolved.

Second, parts of our analyses are hampered by missing data. We have addressed the issue of purged order data, by reconstructing the number of sales using the number of feedbacks per listing. We can expect that buyers are less likely to provide feedbacks on digital items, compared to providing feedback on physical items – given that feedbacks generally report on delivery and packaging. As a result, the amount of digital sales may be underestimated in some cases. Next, we did not have any data on plaintext passwords and bitcoin addresses for all vendors. Therefore, we performed the analysis of password strength and password uniqueness, and the analysis of the traceability of cash-outs on different subsets of data. Yet, we believe that these subsets contain most of the active vendors – since plaintext passwords are available for all vendors that logged on to Hansa in a time span of a month and bitcoin payout addresses are available for orders initiated up to 180 days before the infiltration of the market.

Third, the additional data sources we utilize, introduce some uncertainties. Although the *Grams* search engine was build on databases of online anonymous markets crawled via a special-purpose API, we cannot determine the accuracy or completeness of the *Grams* database. This may have resulted in unrecognized reuse of PGP-keys. Along this same vain, we are dependent on insights provided by the *Chainalysis* service. This tool uses proven heuristics to determine what bitcoin addresses belong to known intermediaries in the bitcoin ecosystem. As these heuristics still may fail, *Chainalysis* might falsely return that a certain bitcoin address does not belong to an exchange or hosted wallet, while it in fact does. Note, that untraceability of cash-out therefore, can be the result of vendors' behavior, limitations of blockchain analysis, or a combination of both. Either way – as law enforcement agencies face these limitations of blockchain analysis as well – investigators are equally hampered by any of the underlying causes for such untraceability. However, this means we have *underestimated* the amount of direct transactions to mainstream exchanges or hosted wallets.

Fourth, an inevitable limitation of our research is that some vendors may have additional security precautions in place outside the scope of the market to start investigators off on the wrong foot. For example, the more professional cybercriminal entrepreneur may use money mules to cash-out their profits. Instead of identifying the actual vendor, a subpoena might therefore result in identifying the individual that – perhaps unknowingly – aids the vendor in laundering its money by transferring funds between accounts. Still, we argue that leaving easy-to-trace transactional links to money mules is an insecure practice.

6.3 Implications

Although in the past anecdotal evidence was presented by law enforcement and industry reports on the failing security practices of cybercriminals, we did not know if this was just a lucky break or a pattern of poor security. We now know that at least on Hansa, the latter was the case. Given the shear amount of cybercrime to choice from to investigate, efficiency in the prosecution of cybercrime is key [53]. Our findings shine a light on exploitable security decisions cybercriminals make. Therefore, we are now better equipped to adequately understand and predict the insecure practices of a cybercriminal entrepreneur and how law enforcement can invest in these. Most notably, we show that among vendors that specialize in trading digital cybercrime items – of whom it would be assumed to have their digital security well organized – insecure practices are most prevalent. These findings may aid allocating the scarce capacity of law enforcement investigators more effectively. Instead of waiting for this one lucky break, or case of low-hanging fruit, it seems that even the most seasoned cybercriminals have at least one weak spot.

7 Related Work

Our paper builds on and benefits from recent advancements into a number of topics. First, our work relates to measurements of the anatomy and economics of online anonymous markets. Second, we can identify similar analyses compared to our investigation of 'deviant security' practices. Third and last, we benefit from and contribute to the research body on risk assessment in a criminal context. In this section, we discuss related work on these three topics.

Measuring online anonymous markets. Similar to our work, Christin [7], Soska et al. [49], and Van Wegberg et al. [58] perform large scale measurements on vendors, listings and transactions on online anonymous markets. Moreover, Van Wegberg et al. [57] similarly cluster vendors into 'vendor profiles' using LPA. In sharp contrast to that body of work, we base our analyses not on scraped, but unique ground-truth data. Security practices of users that operate on online anonymous market is also investigated in earlier work [59]. Here, the authors measure whether vendors stick with their PGP-key and/or username when switching markets. Likewise, Soska & Christin [49] assess the adoption of PGP among vendors active on markets between 2012-2015.

Cybercriminal security practices. Beneficial to our analyses, Van de Sandt [53] – through a grounded theory approach – develops a theoretical foundation on how cybercriminals deploy technical computer security controls. Additionally, he uses micro-economic theories to unravel the security practices of cybercriminals. Other work that connects to this paper are advancements made by Van Hardeveld [55, 56]. He discusses the cognitive biases that lead to insecure practices of carders [6]. Additionally, both authors conduct expert interviews and analyze technical security mechanisms mentioned in online carding tutorials.

Risk assessments. One of the possible explanations for the observed differences in security practices, is the perceived risks of certain illicit activities. Other work elaborates on the security practices of those who engage in the consumption or production of online child abuse material. These types of cybercriminals seem to prioritize their security practices based on the severity of potential punishment and the likelihood of law enforcement prosecution [43]. 'Simple downloaders' often lack technological knowledge and are easily identified, whereas producers of online child abuse material have very high security standards [40]. Similarly, Van de Sandt [53] reasons that cybercriminals who are not fully aware of the illegality of their acts, tend to have little to no security mechanisms in place.

8 Conclusions

In this paper, we measured the prevalence of poor security practices on Hansa Market across different types of vendors. We identified characteristics of vendors that can explain their security practices and clustered vendors that have similar characteristics into distinct 'vendor types' using latent profile analysis. We captured password strength, password uniqueness, the use of two-factor authentication, PGP-key strength, PGP-key reuse and the traceability of cash-outs of Hansa Market vendors. Then, we measured the prevalence of poor security practices across vendor types. We contrasted these findings with the practices of buyers. Finally, we explored possible explanations for the observed differences in security practices.

We found that security practices do not occur at random. There is a clear distinction in the security performance between the defined types of vendors. We observed a dichotomy in security practices between on the one hand two clusters of relatively experienced vendors that sold large amount of drugs and on the other hand two clusters of vendors that specialize in selling digital items. The former group prioritizes their security, while vendors belonging to the latter resort to insecure practices more often.

By comparing the security practices of buyers with that of vendors, we found that on average, buyers use less complex passwords, have less often an unique password and that very few buyers use 2FA as an additional security measure. The latter is partly due to the low PGP-adoption among buyers (12% *vs* almost 100%).

In conclusion, we found surprising patterns in the security practices of users that operate on online anonymous markets. Clusters of vendors that specialize in selling digital items make 'mistakes' in their digital security the most often, while vendors belonging to clusters of successful drug dealers tend to have the best digital security. Additionally, we conclude that many vendors – including the highly successful ones – make the mistake of initiating traceable cash-outs to mainstream bitcoin exchanges and hosted wallet providers.

Regarding the generalizability of our findings, it is important to note that we focused on only one market: Hansa Market (2015-2017). We also had to overcome some issues with partly missing data, through reconstructing orders using feedbacks, which may have led to an underestimation of the number of digital items sold.

[6] Carders trade stolen credit card and bank account details in the underground community.

Despite these limitations, we were able to perform the first empirical, quantitative analysis on cybercriminal security practices using ground-truth data of on online anonymous market. We are now able to better understand the patterns of (in)secure practices of cybercriminal entrepreneurs. Most notably, we show that among vendors that specialize in trading digital cybercrime items – of whom it would be assumed to have their digital security well organized – insecure practices are most prevalent. These findings may aid allocating the scarce resources of law enforcement investigators more effectively, as they now know that investing in building a case against seasoned cybercriminal entrepreneurs is anything but impossible.

9 Acknowledgments

First and foremost, we are grateful to Dutch law enforcement for their efforts and trust that ultimately resulted in the unique opportunity to use the Hansa back-end data for this paper. We thank our reviewers for their constructive feedback on this paper. Likewise, we thank Michel van Eeten and Gert-Jan van Hardeveld for their invaluable comments on earlier versions of our work.

References

[1] Judith Aldridge and Rebecca Askew. 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, 2017.

[2] Judith Aldridge and David Décary-Hétu. Not an ‘ebay for drugs’: the cryptomarket ‘silk road’ as a paradigm shifting criminal innovation. 2014.

[3] Elaine Barker, William Barker, William Burr, William Polk, Miles Smid, et al. *Recommendation for key management: Part 1: General, 5th Rev.* National Institute of Standards and Technology, Technology Administration, 2020.

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

[5] Julian Broséus, Damien Rhumorbarbe, Caroline Mireault, Vincent Ouellette, Frank Crispino, and David Décary-Hétu. Studying illicit drug trafficking on darknet markets: structure and organisation from a canadian perspective. *Forensic science international*, 264:7–14, 2016.

[6] Chainalysis. *Crypto Crime Report 2020*. *Accessed on* 2020-08-05 *at* https://go.chainalysis.com/rs/503-FAP-074/images/2020-Crypto-Crime-Report.pdf.

[7] Nicolas 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*, pages 213–224. ACM, 2013.

[8] Linda M Collins and Stephanie T Lanza. *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*, volume 718. John Wiley & Sons, 2009.

[9] Joseph Cox. Staying in the shadows: the use of bitcoin and encryption in cryptomarkets. *The Internet and drug markets*, pages 41–48, 2016.

[10] David Décary-Hétu and Anna Leppänen. Criminals and signals: An assessment of criminal performance in the carding underworld. *Security Journal*, 29(3):442–460, 2016.

[11] Alexis Dinno. Nonparametric pairwise multiple comparisons in independent groups using dunn’s test. *The Stata Journal*, 15(1):292–300, 2015.

[12] Martin Dittus, Joss Wright, and Mark Graham. Platform criminalism: The ‘last-mile’ geography of the darknet market supply chain. In *Proceedings of the 2018 World Wide Web Conference*, pages 277–286. International World Wide Web Conferences Steering Committee, 2018.

[13] Avinash K Dixit. *Lawlessness and economics: Alternative modes of governance*. Princeton University Press, 2011.

[14] Eurojust and Europol. Common challenges in combating cybercrime as identified by Eurojust and Europol, 2019.

[15] Europol. Massive blow to criminal dark web activities after globally coordinated operation, 2017. *Accessed on* 2020-08-05 *at* https://europa.eu/!mP49nj.

[16] Steven Furnell. *Cybercrime: Vandalizing the information society*. Addison-Wesley London, 2002.

[17] Giada Girelli. The death penalty for drug offences, 2019. *Accessed on* 2020-08-05 *at* https://www.hri.global/files/2019/02/22/HRI_DeathPenaltyReport_2019.pdf.

[18] Maximilian Golla, Miranda Wei, Juliette Hainline, Lydia Filipe, Markus Dürmuth, Elissa Redmiles, and Blase Ur. “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*, pages 1549–1566, 2018.

[19] GramsAdmin. Grams admin announcing shutdown on Reddit, 2017. *Accessed on* 2020-08-05 *at* https://archive.vn/20171215041614/https://www.reddit.com/r/grams/comments/7ikv9r/so_long_and_thanks_for_all_the_fish/.

[20] Andy Greenberg. Operation bayonet: Inside the sting that hijacked an entire dark web drug market, 2018. *Accessed on* 2020-08-05 *at* https://www.wired.com/story/hansa-dutch-police-sting-operation/.

[21] Shuang Hao, Kevin Borgolte, Nick Nikiforakis, Gianluca Stringhini, Manuel Egele, Michael Eubanks, Brian Krebs, and Giovanni Vigna. Drops for stuff: An analysis of reshipping mule scams. In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pages 1081–1092, 2015.

[22] Robert Augustus Hardy and Julia Ro Norgaard. Reputation in the internet black market: an empirical and theoretical analysis of the deep web. *Journal of Institutional Economics*, 12(3):515–539, 2016.

[23] Mikkel Alexander Harlev, Haohua Sun Yin, Klaus Christian Langenheldt, Raghava Mukkamala, and Ravi Vatrapu. 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*, 2018.

[24] Ryan Harris. Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem. *Digital investigation*, 3:44–49, 2006.

[25] Pieter Hartel and Rolf Van Wegberg. Crime and online anonymous markets. *International and Transnational Crime and Justice*, page 67, 2019.

[26] Bernhard Haslhofer, Roman Karl, and Erwin Filtz. O bitcoin where art thou? insight into large-scale transaction graphs. In *SEMANTiCS (Posters, Demos, SuCCESS)*, 2016.

[27] Abraham Hinteregger and Bernhard Haslhofer. An empirical analysis of monero cross-chain traceability. 2018.

[28] Thanh Nghia Ho and Wee Keong Ng. Application of stylometry to darkweb forum user identification. In *International Conference on Information and Communications Security*, pages 173–183. Springer, 2016.

[29] Troy Hunt. Have I Been PWND, 2020. *Accessed on* 2020-08-05 *at* https://haveibeenpwned.com/Passwords.

[30] Troy Hunt. Have I Been PWND, 2020. *Accessed on* 2020-08-05 *at* https://www.troyhunt.com/10b/.

[31] Blake Ives, Kenneth R Walsh, and Helmut Schneider. The domino effect of password reuse. *Communications of the ACM*, 47(4):75–78, 2004.

[32] Arjen K Lenstra. Key length. Contribution to the handbook of information security, 2004.

[33] LocalBitcoins.com. Aml regulation and new features update, 2019. *Accessed on* 2020-08-05 *at* https://localbitcoins.com/blog/aml-features-update/.

[34] Jonathan Lusthaus. Trust in the world of cybercrime. *Global crime*, 13(2):71–94, 2012.

[35] Jay Magidson and Jeroen K Vermunt. Latent class models. *The Sage handbook of quantitative methodology for the social sciences*, pages 175–198, 2004.

[36] Katherine E Masyn. 25 latent class analysis and finite mixture modeling. *The Oxford handbook of quantitative methods*, page 551, 2013.

[37] Damon McCoy, Andreas Pitsillidis, Jordan Grant, Nicholas Weaver, Christian Kreibich, Brian Krebs, Geoffrey Voelker, Stefan Savage, and Kirill Levchenko. Pharmaleaks: Understanding the business of online pharmaceutical affiliate programs. In *21st USENIX Security Symposium (USENIX Security 12)*, pages 1–16, 2012.

[38] Wim Meeus, Rens van de Schoot, Theo Klimstra, and Susan Branje. Personality types in adolescence: change and stability and links with adjustment and relationships: a five-wave longitudinal study. *Developmental psychology*, 47(4):1181, 2011.

[39] Sarah Meiklejohn, Marjori Pomarole, Grant Jordan, Kirill Levchenko, Damon McCoy, Geoffrey M Voelker, and Stefan Savage. A fistful of bitcoins: characterizing payments among men with no names. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 127–140, 2013.

[40] Michael Moran. Online child abuse material offenders. are we assigning law enforcement expertise appropriately. *Unpublished manuscript. Dublin, Ireland: University College Dublin*, 2010.

[41] Kathleen Moriarty, Burt Kaliski, Jakob Jonsson, and Andreas Rusch. Pkcs# 1: Rsa cryptography specifications version 2.2. *Internet Engineering Task Force, Request for Comments*, 8017, 2016.

[42] Carlo Morselli, Cynthia Giguère, and Katia Petit. The efficiency/security trade-off in criminal networks. *Social Networks*, 29(1):143–153, 2007.

[43] National Rapporteur on Trafficking in Human Beings. Child pornography – first report of the dutch national rapporteur. the hague: Bnrm, 2011.

[44] Arman Noroozian, Jan Koenders, Eelco Van Veldhuizen, Carlos H Ganan, Sumayah Alrwais, Damon McCoy, and Michel Van Eeten. Platforms in everything: analyzing ground-truth data on the anatomy and economics of bullet-proof hosting. In *28th USENIX Security Symposium (USENIX Security 19)*, pages 1341–1356, 2019.

[45] Andreas Pfitzmann and Marit Hansen. A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management, 2010.

[46] Fergal Reid and Martin Harrigan. An analysis of anonymity in the bitcoin system. In *Security and privacy in social networks*, pages 197–223. Springer, 2013.

[47] Damien Rhumorbarbe, Ludovic Staehli, Julian Broséus, Quentin Rossy, and Pierre Esseiva. 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, 2016.

[48] Scott Ruoti, Jeff Andersen, Daniel Zappala, and Kent Seamons. Why johnny still, still can't encrypt: Evaluating the usability of a modern pgp client, 2015.

[49] Kyle Soska and Nicolas Christin. Measuring the longitudinal evolution of the online anonymous marketplace ecosystem. In *24th USENIX Security Symposium*, pages 33–48, 2015.

[50] Srikanth Sundaresan, Damon McCoy, Sadia Afroz, and Vern Paxson. Profiling underground merchants based on network behavior. In *2016 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–9. IEEE, 2016.

[51] Xiao Hui Tai, Kyle Soska, and Nicolas Christin. Adversarial matching of dark net market vendor accounts. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1871–1880, 2019.

[52] Joe Van Buskirk, Sundresan Naicker, Amanda Roxburgh, Raimondo Bruno, and Lucinda Burns. Who sells what? country specific differences in substance availability on the agora cryptomarket. *International Journal of Drug Policy*, 35:16–23, 2016.

[53] Erik Van de Sandt. *Deviant security: the technical computer security practices of cyber criminals*. PhD thesis, University of Bristol, 2019.

[54] Michel Van Eeten and Johannes M Bauer. Economics of malware: security decisions, incentives and externalities. *OECD Science, Technology and Industry working papers*, 2008(1), 2008.

[55] Gert Jan Van Hardeveld. *Deviating from the cybercriminal script: Exploring the contextual factors and cognitive biases involved in carding*. PhD thesis, University of Southampton, 2018.

[56] Gert Jan Van Hardeveld, Craig Webber, and Kieron O'Hara. Deviating from the cybercriminal script: exploring tools of anonymity (mis) used by carders on cryptomarkets. *American Behavioral Scientist*, 61(11):1244–1266, 2017.

[57] Rolf Van Wegberg, Fieke Miedema, Ugur Akyazi, Arman Noroozian, Bram Klievink, and Michel van Eeten. Go see a specialist? predicting cybercrime sales on online anonymous markets from vendor and product characteristics. In *Proceedings of The Web Conference 2020*, pages 816–826, 2020.

[58] Rolf Van Wegberg, Samaneh Tajalizadehkhoob, Kyle Soska, Ugur Akyazi, Carlos Hernandez Ganan, Bram Klievink, Nicolas Christin, and Michel van Eeten. Plug and prey? measuring the commoditization of cybercrime via online anonymous markets. In *27th USENIX Security Symposium (USENIX Security 18)*, pages 1009–1026, 2018.

[59] Rolf Van Wegberg and Thijmen 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*, pages 1–5, 2018.

[60] Jeroen K Vermunt and Jay Magidson. Technical guide for latent gold 5.0: Basic, advanced, and syntax. *Belmont, MA: Statistical Innovations Inc*, 2013.

[61] Rick Wash, Emilee Rader, Ruthie Berman, and Zac Wellmer. Understanding password choices: How frequently entered passwords are re-used across websites. In *Twelfth Symposium on Usable Privacy and Security (SOUPS) 2016)*, pages 175–188, 2016.

[62] Frank Wehinger. The dark net: Self-regulation dynamics of illegal online markets for identities and related services. In *2011 European Intelligence and Security Informatics Conference*, pages 209–213. IEEE, 2011.

[63] Daniel Lowe Wheeler. zxcvbn: Low-budget password strength estimation. In *25th USENIX Security Symposium (USENIX Security 16)*, pages 157–173, 2016.

[64] David A Wheeler and Gregory N Larsen. Techniques for cyber attack attribution. Technical report, 2003.

[65] Kim Zetter. New ’google’ for the dark web makes buying dope and guns easy, 2014. *Accessed on* 2020-08-05 *at* https://www.wired.com/2014/04/grams-search-engine-dark-web/.

[66] Gengqian Zhou, Jianwei Zhuge, Yunqian Fan, Kun Du, and Shuqiang Lu. A market in dream: the rapid development of anonymous cybercrime. *Mobile Networks and Applications*, 25(1):259–270, 2020.

[67] zxcvbn. Github source, 2017. *Accessed on* 2021-05-14 *at* https://github.com/dropbox/zxcvbn.