Plug and Prey? Measuring the Commoditization of Cybercrime via Online Anonymous Markets

van Wegberg, Rolf; Tajalizadehkhoob, Samaneh; Soska, Kyle; Akyazi, Ugur; Hernandez Ganan, Carlos; Klievink, Bram; Christin, Nicolas; van Eeten, Michel

Publication date
2018

Document Version
Final published version

Published in
Proceedings of the 27th USENIX Security Symposium

Plug and Prey? Measuring the Commoditization of Cybercrime via Online Anonymous Markets

Rolf van Wegberg and Samaneh Tajalizadehkhoob, ***Delft University of Technology;*** **Kyle Soska,** ***Carnegie Mellon University;*** **Ugur Akyazi, Carlos Hernandez Ganan, and Bram Klievink,** ***Delft University of Technology;*** **Nicolas Christin,** ***Carnegie Mellon University;*** **Michel van Eeten,** ***Delft University of Technology***

https://www.usenix.org/conference/usenixsecurity18/presentation/van-wegberg

This paper is included in the Proceedings of the 27th USENIX Security Symposium.

August 15–17, 2018 • Baltimore, MD, USA

ISBN 978-1-939133-04-5

Open access to the Proceedings of the 27th USENIX Security Symposium is sponsored by USENIX.

Plug and Prey? Measuring the Commoditization of Cybercrime via Online Anonymous Markets

Rolf van Wegberg[1], Samaneh Tajalizadehkhoob[1], Kyle Soska[2], Ugur Akyazi[1], Carlos Gañán[1], Bram Klievink[1], Nicolas Christin[2], and Michel van Eeten[1]

[1]Department of Multi-Actor Systems, Delft University of Technology
{R.S.vanWegberg, S.T.Tajalizadehkhoob, U.Akyazi, C.HernandezGanan, A.J.Klievink, M.J.G.vanEeten} @tudelft.nl
[2]CyLab Security and Privacy Institute, Carnegie Mellon University
{ksoska, nicolasc} @cmu.edu

Abstract

Researchers have observed the increasing commoditization of cybercrime, that is, the offering of capabilities, services, and resources as commodities by specialized suppliers in the underground economy. Commoditization enables outsourcing, thus lowering entry barriers for aspiring criminals, and potentially driving further growth in cybercrime. While there is evidence in the literature of specific examples of cybercrime commoditization, the overall phenomenon is much less understood. Which parts of cybercrime value chains are successfully commoditized, and which are not? What kind of revenue do criminal business-to-business (B2B) services generate and how fast are they growing?

We use longitudinal data from eight online anonymous marketplaces over six years, from the original Silk Road to AlphaBay, and track the evolution of commoditization on these markets. We develop a conceptual model of the value chain components for dominant criminal business models. We then identify the market supply for these components over time. We find evidence of commoditization in most components, but the outsourcing options are highly restricted and transaction volume is often modest. Cash-out services feature the most listings and generate the largest revenue. Consistent with behavior observed in the context of narcotic sales, we also find a significant amount of revenue in retail cybercrime, i.e., business-to-consumer (B2C) rather than business-to-business. We conservatively estimate the overall revenue for cybercrime commodities on online anonymous markets to be at least US $15M between 2011-2017. While there is growth, commoditization is a spottier phenomenon than previously assumed.

1 Introduction

Many scientific studies and industry reports have observed the emergence of cybercrime-as-a-service models, also referred to as the "commoditization of cybercrime." The idea is that specialized suppliers in the underground economy cater to criminal entrepreneurs in need of certain capabilities, services, and resources [23, 33, 39, 42]. Commoditization allows these entrepreneurs to substitute specialized technical knowledge with "knowing what to buy" - that is, outsourcing parts of the criminal value chain. The impact of this trend could be dramatic: Commoditization substantially lowers entry barriers for criminals, which is hypothesized to accelerate the growth of cybercrime. Prior work found strong evidence for specific cases of commoditization: booters offering DDoS services [29], suppliers in "pay-per-install" markets distributing malware [13], and exploit kit developers supplying "drive-by" browser compromises [22]. The overall pattern is much less clear, however, as not all cybercrime components are equally amenable to outsourcing [21].

This paper answers two core questions: Which parts of cybercrime value chains are successfully commoditized and which are not? What kind of revenue do these criminal business-to-business services generate and how fast are they growing? Addressing these questions requires that we properly define and scope the concept of commoditization. To do so, we turn to transaction cost economics (TCE). We argue that the characteristics of commodities are highly congruent with the characteristics of online anonymous marketplaces. More precisely, the one-shot, anonymous purchases these markets support require suppliers to offer highly commoditized offerings. Conversely, if cybercrime offerings can be commoditized, online anonymous markets should be a highly attractive place to sell them. Indeed, these platforms can reach a large audience and provide risk management services for criminals, e.g., by protecting their anonymity, and featuring reputation systems to root out fraudulent sales and shield sellers from risky interactions with buyers.

While data from online anonymous marketplaces provides a unique opportunity to track the evolution of

commoditization, we are not arguing that these marketplaces provide a complete picture. They do not have a monopoly, of course. In fact, certain types of commoditized offerings are not suited for trading on these marketplaces, e.g., affiliate programs, subscription-based offerings, or services requiring a rich search interface may be better served by alternative distribution channels [26,48]. Yet, on balance, the congruence of commoditized forms of cybercrime and online anonymous markets means that the evolution of commodizitation should be clearly observable on those markets.

We analyze longitudinal data on the offerings and transactions from eight online anonymous marketplaces, collected between 2011 and 2017. We first present a conceptual model of the value chain components in dominant criminal business models, and develop a classifier to map cybercrime-related listings across all markets to these components. This allows us to track trends in vendors, offerings and transaction volumes. We then discuss the type of offerings to assess to what extent each component can be outsourced - i.e., to what extent it is successfully commoditized. In short, we make the following contributions:

- We present the first comprehensive empirical study of the commoditization of cybercrime on online anonymous markets. We analyze 44,000 listings and over 564,000 transactions across eight marketplaces. We draw on data from prior work [40] and newly collected data on AlphaBay.
- We find commoditized business-to-business offerings for most value chain components, though many of them are niche products with only modest transaction volumes. Cash-out services contain the most listings and generate the largest revenue. We estimate the lower bound of overall B2B revenue to be around $2 million in 2016 and over $8 million for the whole period.
- We also uncover a surprising amount of revenue in retail cybercrime – that is, business-to-consumer sales rather than business-to-business, similar to the patterns observed for drug sales. The lower-bound estimate for 2016 is over $1 million and nearly $7 million for the whole period.
- We demonstrate that commoditization is a more spotty phenomenon than previously assumed. The lack of strong growth in transactions suggests that bottlenecks remain in outsourcing critical parts of criminal value chains.

The rest of this paper is structured as follows. Section 2 defines transaction cost economics, and discusses how the concept applies to cybercrime commoditization.

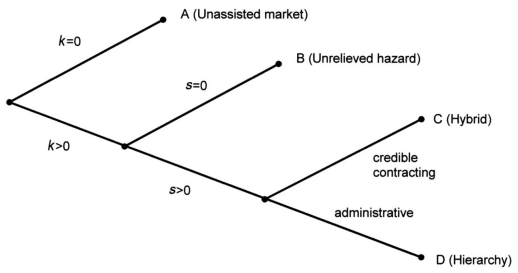


Figure 1: **Contracting scheme in the TCE framework.**

Section 3 describes the demand of cybercrime outsourcing. Section 4 presents our measurement methodology. Section 5 lays down our classification analysis, and section 6 identifies the best-selling clusters of cybercrime components. Section 7 discusses our findings, and Section 8 connects our work to earlier contributions. Section 9 concludes.

2 Commoditization and anonymous marketplaces

With outsourcing, entrepreneurs can decide to either "make" or "buy" each component of the value chain. Transaction cost economics (TCE) is a mature economic theory that seeks to explain under what conditions economic activity is organized in markets (buy) and when it is vertically integrated (make) – i.e., the entrepreneur develops the component himself or brings someone with that capability into the enterprise. Here, we apply TCE to the context of cybercrime to predict if and when outsourcing takes place.

Williamson [47] distinguishes several asset characteristics that determine if and how outsourcing will occur, as shown in Figure 1. A, B, and C are various forms of outsourcing and D is vertical integration. Factors such as asset specificity, frequency and uncertainty separate the underlying transactions [45]. k is a measure of asset specificity, referring to the degree to which a product or service is specific to e.g., a vendor, location, control over resources, etc. A key characteristic of commodities is that they are "fungible", meaning that different offerings of it are mutually interchangeable ($k = 0$) – i.e., a booter is a booter [29, 31] – and subject to vendor competition [18]. In commodity markets, buyers can easily turn to other suppliers, and suppliers can sell to other buyers, reducing possible hazards. The more specific an asset is ($k > 0$), the more investments are specialized to a particular transaction.

The second factor, s, refers to contractual safeguards. Transactions where investments are exposed to unrelieved contractual hazards ($s = 0$) will not be traded pub-

licly (i.e., anonymous online marketplaces such as Silk Road or AlphaBay are a poor fit), but on smaller, "invite-only" markets where trust relations are forged among specialized insiders, anonymity is not absolute, and escrow services are less prominent [36]. When $s > 0$, contracts with transaction-specific safeguards are in place.

Commodities are sold via unassisted markets (A). These markets incentivize sellers to reduce asset specificity as much as possible, hence commoditizing the offering. The efficiency gains also work in the other direction: those who offer goods or services that can be commoditized would use these markets to sell them and benefit from the wide reach and high frequency of transactions, without being exposed to risky direct interaction and coordination with buyers.

In terms of TCE, online anonymous marketplaces are unassisted markets – i.e., they are the place to go for commoditized cybercrime. Anonymous markets reduce uncertainty risks through escrow mechanisms, review systems and strict rules enforced by a market administrator [15,40]. For transactions where $k = 0$, "no specific assets are involved and the parties are essentially faceless" [46, p. 20], which is precisely the case for anonymous markets. Complex components such as highly customized malware are more likely to be self-supplied or delivered under special contracts, while frequently used, standardized components, like DDoS-services, would be supplied more efficiently by the unassisted market. TCE tells us that the organization of criminal activities will be guided primarily by the relative costs of completing illegal transactions within the market [19, p. 28].

Similar to the prominent drugs-trade on anonymous online markets, we expect two type of commodities on these markets: business-to-business (B2B), e.g., wholesale quantities of credit card details, and business-to-consumer (B2C), e.g., a handful of Netflix accounts. We are primarily interested in B2B, as that is the form of commoditization that is the most worrying and speculated to cause a massive growth in cybercrime, though we will also report the main findings for B2C. To assess the degree to which B2B services are commoditized, the next section develops a framework to identify the different value chains where there is demand for commoditized cybercrime.

3 Demand for cybercrime outsourcing

To empirically assess the commoditization of cybercrime, we first need to establish what capabilities, services and resources criminal entrepreneurs actually need. This provides us with a framework against which to evaluate where commodities are available to meet this demand and where they are not – as measured through listings on anonymous marketplaces. Of course, entrepreneurs might demand an endless variety of goods and services. For this reason, we use as our starting point the dominant criminal business models that were identified in prior work. We look at the value chain underlying each business model and synthesize them in a common set of components that entrepreneurs might want to outsource. Our point of departure is Thomas et al. [42]'s inventory of criminal business models. We update and extend this set with models discussed in related research. Table 1 shows this updated overview.

First, we look into the value chain behind spamvertising, which is driven by three resources: a) advertisement distribution b) hosting and click support and c) realization and cash-out [34,42].

Second, extortion schemes, for instance ransomware or fake anti-virus [17] have a value chain that consists of four distinctive resources: a) development of malware b) distribution, by either exploits or (spear)phishing e-mails, c) take-over and "customer service" and d) cash-out [30,42].

Third, click fraud is supported by four similar, general resources: a) development of a website, malware or a JavaScript, b) distribution through botnets, c) take-over by either malware or JavaScript and d) cash-out [32,42].

Fourth, the criminal business model in social engineering scams, such as tech support scams [35], or one-click fraud [16] leans on: a) (optional) development of malware or a malicious app, b) distribution by phishing e-mail or website, or through social engineering, c) take-over and setting-up "customer service," and d) cash-out [35,42]. The boundary between extortion and social engineering scams is fuzzy. Both could well be categorized in the same family. For now, we take the view that extortion (e.g., ransomware) requires development of malware, where social engineering scams do not necessarily rely on anything being installed on the victim's machine (e.g., one-click frauds [16]).

Fifth, cybercriminal fraud schemes, e.g. those enabled by financial malware, build on four general, main resources: a) development and b) distribution of malware or a malicious app, c) take-over, for instance by using web-injects or a RAT,[1] and d) cash-out [42,44].

Sixth, cryptocurrency mining relies on near-similar resources as click fraud: a) the development of malware or JavaScript, b) distribution of malware by botnets or the injection of a JavaScript in a compromised websites, c) the take-over, i.e. mining, and d) cash-out [27,42].

Seventh, the criminal business model that profits from selling stolen credit card details makes use of: a) development of a phishing website, malware or a malicious

[1] Remote Access Tool, i.e., malware that allows a miscreant to remotely access a victim's machine.

Table 1: **Overview of present-day cybercriminal business models**

Business model	Example Modus Operandi	Source
Spamvertised products	Selling knock-off products	Levchenko et al. [34], Thomas et al. [42]
Extortion	Ransomware	Kharraz et al. [30], Thomas et al. [42]
Clickfraud	Hijacked traffic	Kshetri et al. [32], Thomas et al. [42]
Social engineering scams	Customer support scams	Miramirkhani et al. [35], Christin et al. [16], Thomas et al. [42]
Fraud	Financial malware	Thomas et al. [42], Van Wegberg et al. [44]
Mining	Cryptocurrency mining	Huang et al. [27], Thomas et al. [42]
Carding	Credit card reselling	Holt [23], Thomas et al. [42]
Accounts	Reselling credentials	Holt [23], Thomas et al. [42]

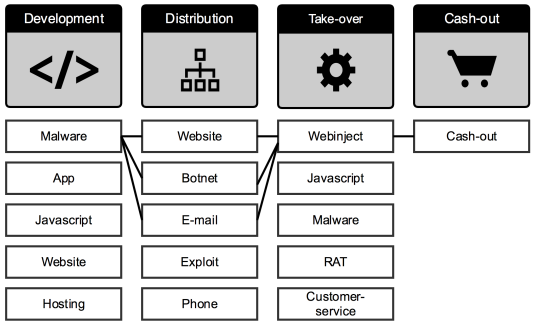


Figure 2: **Conceptual model of value chains, showing a representation of the financial malware value chain**

apps, b) distribution, c) take-over, i.e. the logging of information, and d) reselling and cashing-out [23, 42].

Last, the resale of non-financial accounts leans on the exact same resources as carding [23, 42].

Looking at these value chains, we can see that some components are common among them. All models relate to at least four main resources: development, distribution, take-over and cash-out. We merge these into a single component that belongs to two or more value chains. We can synthesize all value chains in a overall set of 13 components. Some components, e.g., malware, can be used for more than one main resource. Figure 2 summarizes our conceptual model and the overall demand for B2B services in cybercrime.

4 Measurement methodology

Our measurement methodology consists of 1) collecting and parsing data on listings, prices and buyer feedback from eight prominent online anonymous markets, 2) implementing and applying a classifier to the listings to map them to cybercrime components from our conceptual model of value chains (Figure 2) as well as to additional categories of B2C cybercrime, and 3) using Latent Dirichlet Allocation (LDA, [10]) to identify the best-selling clusters of listings and compare their offerings to the capabilities, resources and services needed for each component of the conceptual model.

4.1 Data collection

We first leveraged the parsed and analyzed dataset of Soska and Christin [40] to obtain information about item listings and reviews on several prominent online anonymous marketplaces. For each of the over 230,000 item listings, the data include (but are not limited to) titles, descriptions, advertised prices, item-vendor mapping, category classification, shipping restrictions and various timestamps. Additionally, each item listing contains feedback that has been proven to be a reasonable proxy for sales [15, 40]. Each piece of feedback contains a message, a numerical score, and a timestamp.

We then extended this data with an additional 16 complete snapshots of AlphaBay that we collected from May 30, 2016 to May 26, 2017, just two months before its closure in July 2017 [4]. Table 2 summarizes the dataset. We merged the new AlphaBay scrapes with the existing dataset by first parsing out the same supported fields and then running a compatible analysis using the categorical classifier from Soska and Christin [40].[2] AlphaBay is important since, according to the FBI [4], by the time of its closure, it had featured over 100,000 listings for stolen and fraudulent documents, counterfeits, and malware in particular. The US Department of Justice (DoJ) also claims that AlphaBay was the largest single online anonymous marketplace ever taken down [3].

As an important data processing note, some vendors set "holding prices" to their listings when the product or service they are selling is out of stock. Instead of removing the listing, these vendors increase the price (astronomically) to prevent buyers trying to buy their product. Soska and Christin [40] developed a heuristic that corrects these holding prices, which we applied in the preprocessing of the parsed and labeled dataset. This limits

[2] Soska and Christin's dataset included 17 snapshots of AlphaBay, dating back to December 2014, that they did not use in their published analysis [40].

Table 2: **Markets crawled**

Market	First seen	Last seen	# Snapshots
Agora	2013-12-24	2015-02-11	161
Alphabay	2014-12-31	2017-05-26	33
Black Market Reloaded	2012-11-21	2013-12-04	25
Evolution	2014-01-13	2015-02-18	43
Hydra	2014-04-14	2014-10-26	29
Pandora	2013-11-02	2014-10-13	140
Silk Road 1	2011-06-21	2013-08-19	133
Silk Road 2	2013-11-27	2014-10-29	195

the potential for errors stemming from falsely assuming a certain holding price was associated with a buy.

4.2 Classifying cybercrime listings

Most listings on these marketplaces are related to drugs and other non-cybercrime activities [15, 40]. Our aim is to classify each item listing into one of the 10 categories of cybercrime components from the conceptual framework (Figure 2). Unfortunately, the labels provided by Soska and Christin are not expressive enough to capture these nuanced categories, so we begin by using their labels as a pre-filter and retain only item listings that were identified as being either "Digital goods" or "Miscellaneous" (19% of all listings).

Next, we implemented a Linear Support Vector Machine (SVM) classifier. Manual inspection confirmed our suspicion that the markets also contain retail (B2C) cybercrime offerings, next to wholesale cybercrime offerings. For this reason, we added six product categories to distinguish supply in that part of the market: accounts, custom requests, fake documents, guides and tutorials, pirated goods, and vouchers. A final category, namely, "other", captures the listings that did not fit anywhere else (e.g., scanned legal documents). The classifier is initially trained and evaluated on a sample of listings ($n = 1,500$) from all the markets, where ground truth is created via manual labeling.

Table 3: **"Digital Goods" & "Miscellaneous" Listings**

Market	# Listings	# Vendors	Total revenue
Agora	3,240	526	$ 1,818,991
Alphabay	21,350	3,055	$ 13,471,406
Black Market Reloaded	2,069	386	$ 685,108
Evolution	9.551	1.002	$ 6,125,136
Hydra	377	28	$ 242,230
Pandora	1,204	169	$ 394,306
Silk Road 1	4,053	645	$ 2,239,436
Silk Road 2	2,734	441	$ 4,455,339

4.3 Ground truth

For labeling the ground truth, we randomly selected 1,500 items from all listings classified as either "Digital Goods" or "Miscellaneous" ($n = 44,060$), or approximately 3.5% of the data. Only around 30% of the listings in the random sample belonged to one of the ten B2B cybercrime components. Around 45% belonged to one of the B2C categories and the remaining 25% were labeled as "other." Those were comprised of drug listings that were misclassified as "miscellaneous," as well as luxury items and other physical goods. We also found some incomprehensible listings, which might be test entries by vendors. Labeling the ground truth yielded four more observations. First, we identified listings that contain more than one cybercrime component, e.g., offering both a piece of malware and (access to) a botnet. Second, we identified *package listings*, such as complete cryptocurrency mining schemes. Third, we observed that some vendors add unrelated keywords to their listings, presumably in a marketing effort similar to search engine optimization. Fourth and last, we observed *custom listings*, i.e., listings that are specifically created to be sold only once to one specific buyer. Custom listings contain bespoke products or services ranging from custom quantities to a completely custom-made product such as pre-booked plane tickets.

After labeling our random sample of listings, we can assess whether each category meets our criteria for accurately classifying listings to categories of cybercrime components. To avoid overfitting to a specific component, we ensure the training set for our classifier holds at least 20 listings per category of cybercrime components. Because of the highly skewed distribution of listings in our random sample, we were forced to increase our ground truth by manually adding listings to the following categories: app, botnet, e-mail, exploit, hosting, malware, phone, RAT and website. To that end, we operated a manual search in the filtered portion of data using up to three keywords on those cybercrime components. We manually verified whether the listings with the keyword in the title or description advertise the actual product or was a false positive – e.g., a vendor using the word "malware" in a listing of lottery tickets.

4.4 Training and evaluation

Before training the classifier, we excluded three categories of cybercrime components from the classification: JavaScript malware, webinjects, and customer support. For these, we found no listings in our random sample.

The classification phase itself consists of three steps: (i) data cleaning, (ii) tokenizing, (iii) training and evaluation of the ground-truth samples which are the concatenation of the title and description of the item listings. In

data cleaning, we removed all English stop words, punctuations, numbers, URLs and accents of all unicode characters. We then lemmatized the words in order to group together the inflected forms of a word so they can be analyzed as a single item, identified by the word's lemma, or dictionary form before being trained and tested. We tokenized each item (assuming all items are in English) and computed a *tf-idf* (term frequency inverse document frequency) value for each of the resulting 9,629 unique tokens or words. To calculate the *tf-idf*, we used a *max-df* (maximum document frequency) equal to 0.7 – this discards words appearing in more than 70% of the listings. In the classification phase we then used these values as an input for an L2-Penalized SVM under L2-Loss. We implemented this classifier using Python and *scikit-learn*.

The reported imbalance in the distribution of listings among categories causes an imbalance in the labeled categories of our ground truth. On the one hand, we have nearly 25% of listings labeled as "other" and around 45% labeled as one of B2C products or services. On the other hand, we have a large portion of the rest of our ground truth listings (30%) that are labeled as "cash-out" listings (25%). We mitigate the negative impact of this imbalance on our classification results by re-sampling our ground truth listings by the SMOTE (Synthetic Minority Over-sampling Technique) method, thereby increasing the cardinality of each category to match the size of the largest labels; this is a standard technique towards improving algorithmic fairness. Due to the implicit optimization of our classifier, this over-sampling method allows the model to carve broader decision regions, leading to greater coverage of the minority class [14].

Because of the nature of listings that cover multiple categories, e.g. bundled goods, we anticipate some classification errors. It is however important to distinguish between errors where the item listing is classified as "other" (false negative) from acceptable approximations, e.g., a listing that includes access to a botnet bundled with malware and is classified as a botnet. The first example denotes a classification error, while the second is a listing that truly is a combination of multiple cybercrime components. Our main goal is therefore to prevent cybercrime component listings, like malware, from ending up in "other" and vice versa.

We evaluate the performance of our classifier in Figure 3. In this normalized confusion matrix, each row represents the instances in an actual category while each column represents the instances in a predicted category. All correct predictions are in the diagonal of the table (numbers denote recall). The average precision is 0.78 and the average recall is 0.76, denoting some confusion between cybercrime components categories. However, the classifier meets our goal of avoiding confusion between cybercrime components and "other" listings.

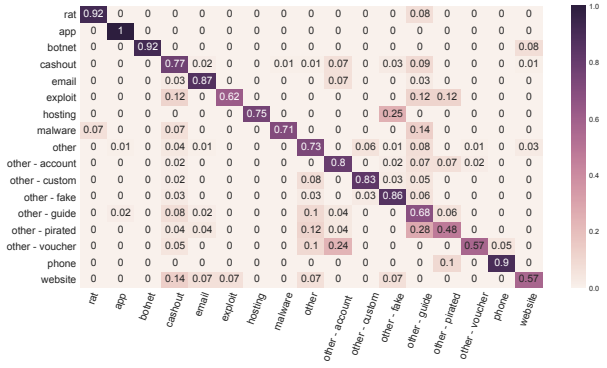


Figure 3: **Classifier normalized confusion matrix**

4.5 Post-processing

The heuristic for dealing with holding prices [40] used in pre-processing does not correct situations where all instances of a listing among our snapshots were either only seen with a holding price, or in some cases do not exceed a set maximum of $10,000. To get an idea of how frequently this happens, we looked into items priced above $5,000. We manually identified 12 listings which received a total of 118 pieces of feedback at holding prices. In one case we found the correct price from a customer commenting "good product for $10". The remaining 11 listings seemed clear instances of holding prices, and were removed, as we had no information about the true sales price.

After examining holding prices, we found some instances of misclassified drug listings in categories of cybercrime components (false positives). To correct this, we first removed 12 Xanax listings that we encountered when inspecting the holding prices. To find additional misclassified drug listings, we leveraged the distinctive features of drug listings, namely the unique terminology used to list the quantity of drugs offered, e.g., "grams," "mg," "ug," "lbs," "ml," "pills," etc. Following this process, we automatically identified and removed 82 misclassified drug listings.

5 Results

In this section we present the results of the classified listings. At first glance, we can observe the differences in number of listings between the categories. Just over 30% of the listings are in the B2B categories of our conceptual model, listed in the top half of Table 4. The lower half of the table covers B2C cybercrime (around 36% of listings), custom orders (14%) and others (20%).

We primarily focus on the B2B categories, though we do report on the B2C categories later in the section. Before we turn to B2B offerings, we take a closer look at

Table 4: **Listings per category.** The top half represents B2B listings, the bottom half, B2C.

Category	# Listings	# Vendors	Total revenue
App	144	75	$ 12,815
Botnet	125	79	$ 46,904
Cash-out	12,125	2,076	$ 7,864,318
E-mail	550	216	$ 97,280
Exploit	115	75	$ 17,603
Hosting	20	15	$ 1,182
Malware	310	162	$ 57,598
Phone	261	148	$ 74,587
RAT	105	65	$ 16,070
Website	664	293	$ 286,405
Accounts	3,759	577	$ 598,491
Fake	3,386	815	$ 2,877,184
Guide	5,049	1,020	$ 2,620,635
Pirated	1,420	338	$ 129,961
Voucher	1,293	386	$ 753,116
Custom	6,310	1,887	$ 5,793,064
Other	8,424	2,652	$ 7,749,788
Total	44,060	5,552	$ 28,997,006

the large category of custom listings. These listings are a bit counter-intuitive to the market structure as they concern one-time, buyer-specific products or services. For instance, stolen credit card details from Norway, a modified type of keylogger, or compromised hosts from the Netherlands. Although some of these listings are in fact B2B cybercrime services, they are not fully commoditized, as the listing reflects a one-time sale and a non-standardized product or item.

There are large differences across the categories of B2B offerings. Cash-out stands out: In terms of the number of listings, active vendors, and in total revenue, this category is by far the largest. It also stands out in other ways. Table 5 reports the median and mean number of listings for each vendor per category, which reflects the degree in which different products need to be differentiated. We see most products offered do not need differentiation. More specific requests might be handled with custom listings, but are not enough to merit a more permanent listing. Cash-out offerings, on the other hand, contain many more relevant distinctions. A vendor can split up its stock of stolen credit card details into smaller sets of details, for instance differentiated to type of credit card.

The second column in Table 5 shows median revenues per listing. Cash-out listings have the highest median revenue. RATs and exploits exhibit, counterintuitively, a similar median revenue. This is a consequence of the generally low-value exploit listed in anonymous marketplaces, e.g., run-of-the-mill Office exploit macros. Rare, high-value exploits, such as iOS or Chrome exploits, would be sold through specialized white or black markets or through private transactions [7]. Other categories have a median between $15 and $34 revenue per listing. As the median revenue is a simple summary of the underlying distribution, we also show the price range – in terms of median, mean, min-max and standard deviation (SD) – for listings in the B2B categories. We see, again, that the cash-out category contains the most expensive set of offerings with very diverse pricing. This diversity in price can also be observed in other categories – in fact, the overall shape of the price distribution function remains relatively unchanged across categories. Moreover, the lifespan of a listing also tells us something about the standardization of the product. A listing that receives instances of feedback over multiple months denotes that the associated product remains valuable and has not become outdated or unrecognizable. Like an ecstasy tablet, a RAT will hold its value over time in terms of being a functional solution. In contrast, stolen credentials "go bad" after some time. The first buyer who uses these credentials will in all likelihood set off red flags at the credit card company for irregular spending, making a subsequent purchase of the same credentials worthless. Curiously, the median lifespan of cash-out listings is above average, which could be due to vendors updating the specific product listed, or persistently selling unusable credit card details, or to a slower-than-expected detection of suspicious transactions by credit card companies.

Looking into median lifespan of listings reveals little differences as all but three categories have a median listing lifespan of close to one month. Both exploit and hosting listings have a low median lifespan of around 0.3 months – approximately 10 days. At the other end of the spectrum, we see that RAT listings have a median lifespan of 1.44 months – approximately 40 days. So, a RAT listing has a significant longer lifespan than an exploit listing. The distribution of cybercrime listing lifespan is heavy-tailed and on average, a cybercrime component is offered for 2.7 months. In short, vendors have one or two listings, except for cash-out listings, where that number is higher. Turnover is between $15 and $60 dollars per listing and lifespan is typically less than a month.

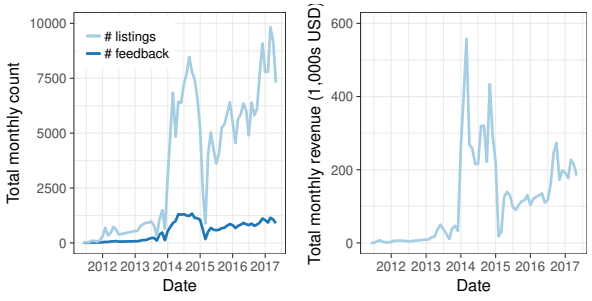
5.1 Listings and revenue over time

The claim that cybercrime is commoditizing also implies a growth in transactions and revenue. Figure 4(a) shows, per month, the unique number of listings and number of feedback. Figure 4(b) shows the corresponding projected revenue. The number of feedback is a proxy for the minimum number of sales, as a buyer can only leave feedback

Table 5: **Vendors, revenue, and lifespan per category**

Category	Listings per vendor		Revenue per listing	Price per listing				Lifespan in months
	Median	Mean	Median	Median	Mean	Min–Max	SD	Median
App	1	1.97	$24.33	$5.70	$18.79	$0–$64	$40.89	0.91
Botnet	1	1.61	$34.44	$14.73	$106.89	$0–$2,475	$341.13	0.60
Cash-out	2	5.88	$60.00	$14.85	$72.42	$0–$9,756	$280.20	0.72
E-mail	1	2.58	$22.85	$7.34	$42.14	$0–$1,606	$139.17	0.52
Exploit	1	1.56	$15.57	$5.26	$28.64	$1–$500	$80.09	0.36
Hosting	1	1.33	$31.60	$16.40	$25.14	$3–$99	$25.47	0.32
Malware	1	1.95	$22.90	$5.45	$37.96	$0–$1,984	$133.68	0.98
Phone	1	1.80	$30.00	$9.90	$45.13	$0–$3,200	$221.99	0.79
RAT	1	1.66	$20.00	$5.41	$38.35	$0–$919	$126.78	1.44
Website	1	2.28	$29.80	$8.72	$51.58	$0–$1,695	$146.42	0.83

when she buys a product. Feedback does not however yield a one-to-one mapping to sales as customers may leave a single piece of feedback after purchasing a high quantity of an item. Anonymous marketplaces depend on effective reputation mechanisms to mitigate uncertainty in transactions.



(a) Listings and feedback amount (b) Revenue

Figure 4: **Number of unique listings, feedback and revenue in categories of B2B cybercrime components per month**

Figure 4 shows a growth in listings, amount of feedback and revenue for cybercrime components between 2012 and 2017. The drop at the end of 2013 and the beginning of 2014 is partly due to the take-down of Silk Road 1 and Black Market Reloaded. The steep increase thereafter is distributed over four new markets (Agora, Evolution, Hydra and Silk Road 2), but shows that the aggregate pattern is clearly one of rapid growth. The next drop, around the end of 2014, is caused by a combination of the law-enforcement operation against Silk Road 2, the exit scam of Evolution and the sudden disappearance of Agora. Right after this volatility, the AlphaBay market emerged, and subsequently became the largest to date. Their operation halted suddenly in July of 2017, when the FBI together with the Dutch Police shut down AlphaBay (and Hansa Market, which we do not report on here). Still, the overall pattern clearly is one of growth. The trade in cybercrime commodities seems resistant to the turbulence across marketplaces.

Figure 5 shows that the upward trend in feedback instances is not only caused by an increase in listings, but also to the increase of amount of feedback per listing. In 2011, a listing on average received around five pieces of feedback per month. Over time, this ascended to around eight pieces of feedback per listing in 2017, with intermediate spikes to over ten pieces of feedback in 2012. Those spikes coincide with the period of time in which Silk Road 1 became known by the general public due to extensive coverage by news and media over the course of 2011 [1]. Conversely, the trough at the end of 2013 is primarily due to the Silk Road 1 takedown and the chaotic few weeks that ensued [40]. Overall, we see that the average amount of feedback per listing stabilizes halfway through 2012 and from that moment onwards seems to follow a slow rise.

Essential to the understanding of the ecosystem is identifying which categories can be attributed to most of the growth in sales and revenue. For each item listing, revenue is calculated by multiplying each feedback specific to a listing with the dollar-price of that listing at the moment the feedback was generated. The revenue from these listing is then aggregated per month and per category. Figure 6 shows revenue per category. The spikes and troughs are, again, the result of marketplace turbulence.

The category of cash-out listings is by far the biggest cybercrime component, in terms of listings, revenue and vendors. We take a closer look to see whether this rev-

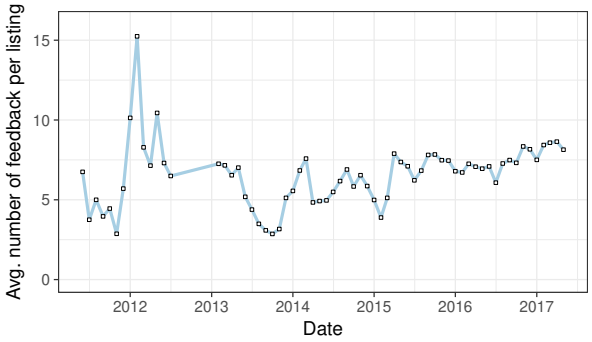


Figure 5: **Feedback per listing in categories of B2B cybercrime components per month**

enue is driven by a small fraction of listings or whether it represents a broader volume of trade. It turns out the a large portion of the increase between 2014 and 2015 is driven by feedbacks on CVV listings. More specifically, one listing offering "US CVVs" received nearly 700 feedbacks in the first quarter of 2014. From the beginning of 2015 onwards we see a steady growth in revenue alongside the growth of AlphaBay market as a whole. In the early days of the ecosystem we see an increase in cash-out revenue which was primarily driven by a listing offering "10,000 USD CASH," which can be seen as typical money laundering – the customer pays in bitcoin and receives cash.

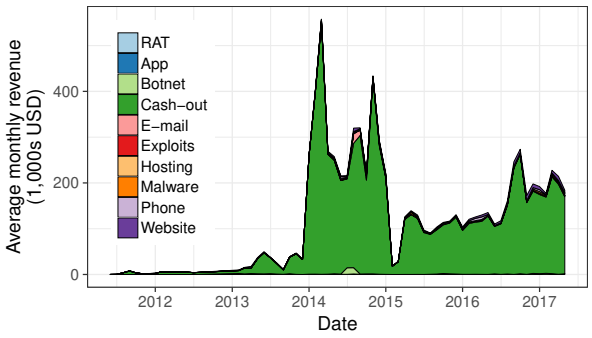


Figure 6: **Total revenue per category per month**

The revenue of cash-out listings is obscuring the other categories. When we omit it in Figure 7, we see that the trend of increasing revenue between 2012 and 2017 becomes apparent yet again. In the second half of 2014, listings in e-mail distribution such as spam tutorials, spam runs or large databases of e-mail addresses generate very high revenue numbers. Similarly, we see a spike in botnet-related sales driven by a mysterious listing titled "source," receiving 10 – rather negative – feedbacks in the summer of 2014. The average of $5000 per month in 2013 grows to $15,000 per month in late 2017. Compared to the average monthly revenue of the entire market ecosystem however – nearing $600,000 per month in late 2014, mostly generated by drugs [40] – this is just a fraction.

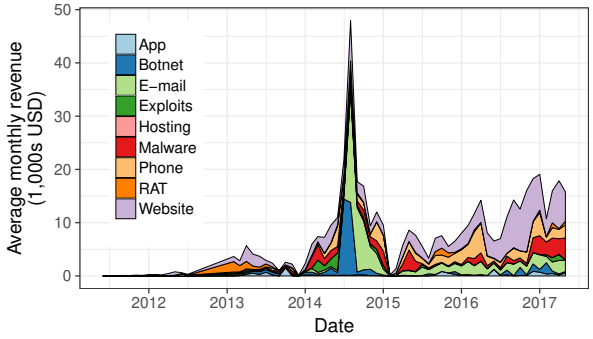


Figure 7: **Total revenue per category per month, excluding cash-out category**

5.2 Vendors over time

Another element in the assessment of commoditization is the level of vendor competition. Figure 8 shows the number of vendors per category over time. A vendor is defined to be active if she has at least one active item listing and may be instantaneously active in multiple categories.

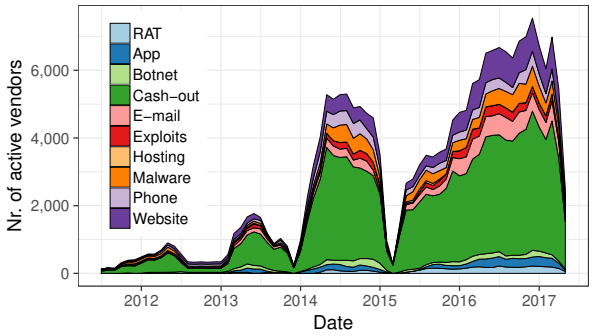


Figure 8: **Number of active vendors per month**

As with the revenue per listing, the number of unique vendors per category is generally increasing over time, however the increase in vendors is steeper than the increase in listings. Figure 9 clearly shows that the increase in vendors from 2014 onwards is due the Evolution and AlphaBay marketplaces. Soska and Christin showed that in the contemporary ecosystem (i.e., after the Silk Road take-down), it is common for each vendor to maintain more than one alias on different marketplaces which may be partially responsible for this observation.

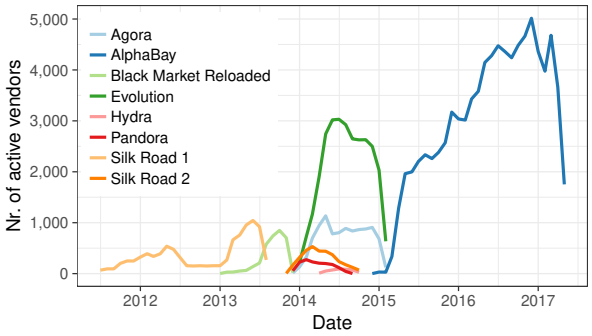


Figure 9: **Vendors over time across markets**

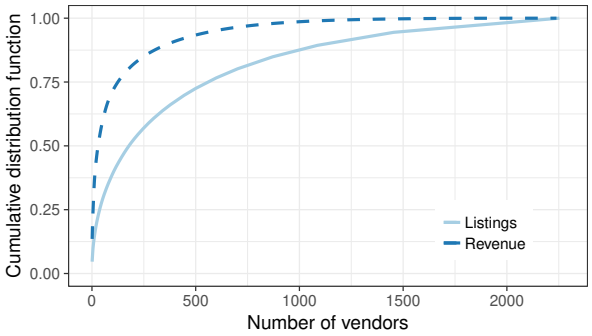


Figure 10: **Cumulative distribution function of listings and revenues across vendors**

Listings and revenue are not distributed normally across vendors. As in many markets, there are big players and small players. Figure 10 plots the cumulative percentage of listings and revenue of cybercrime components over vendors. A small portion of vendors are responsible for a large fraction of the listings. To be more precise, around 30% of vendors are responsible for 80% of all listings. More interestingly, just under 10% of vendors are responsible for generating 80% of the total revenue. That means that around 174 vendors have sold for nearly $7 million worth of cybercrime components. This translates into an average revenue per vendor of around $40,000, but the distribution is wide and skewed. The 174 vendors range from a minimum revenue of $7,355 to a maximum of $1,148,403.

5.3 Marketplaces

Different marketplaces might develop different profiles or specialties in terms of what they sell – i.e., they attract a different set of vendors, offerings or buyers. To compare the product portfolio of different markets, Figure 11 displays the distribution of offerings across different categories. To deal with the large differences in size of the categories, we first take the logarithm of the number

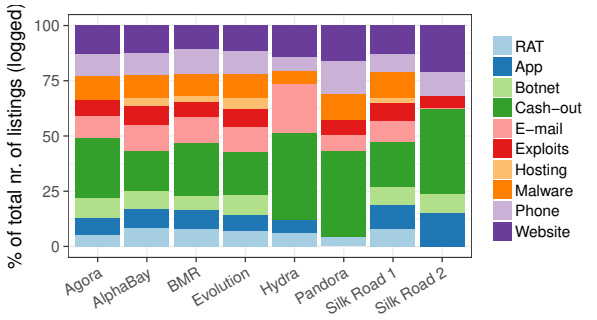


Figure 11: **Percentage of total number of listings per market per category (numbers of listings are logged)**

of listings in each category and then calculate the percentage of each category in this log-scaled total count of listings. There are minor variations visible, but the more obvious pattern is the similarity between most markets. All except two markets, namely Hydra and Pandora, contain listings in each of the categories. Hydra and Pandora are relatively small markets, with a shorter life-span and the absence of listings in some categories is probably due their comparatively modest size and short existence. In terms of commoditization, all categories of the criminal value chains are consistently offered across markets and time. Moreover, these components see vendor competition.

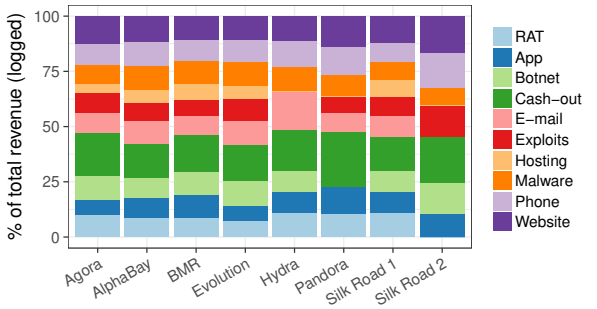


Figure 12: **Percentage of total revenue per market per category (revenue is logged)**

Another way to evaluate market specialization is by categorical revenue. In Figure 12, we show the percentage of revenue – after log transformation – per category of cybercrime component per market. The story is unchanged: there are no major differences between markets. If anything, the picture painted by looking at revenues is even more uniform across marketplaces.

5.4 B2C listings

Finally, we take a look at the listings in retail cybercrime. This covers the categories of "accounts," "fake," "guide," "pirated," and "voucher." We briefly describe the type of listings assigned to those categories. "Accounts" denote listings advertising small batches of accounts from services like Netflix and Spotify. "Fake" contains offerings of fake IDs, counterfeit documents or money. Listings that sell mere instructions or tutorials, are categorized as "guides." The "pirated goods" category encompasses listings that offer pirated movies, software or e-books. Last, the "voucher" category comprise listings that offer discounts at numerous places, ranging from discounted airline tickets to pizza shop gift cards[3]. The retail cybercrime offerings are also forms of commoditization, albeit a slightly atypical one. Indeed, these B2C products are meant to be used or consummed by the buyer, and are not parts for some large value chain with another profit center at the end of it.

The large portion of retail cybercrime is in line with what has been observed on the drugs side of these markets; B2C transactions for consumers of drugs, along with more modest amounts of B2B transactions with larger quantities for lower-level dealers [8].

We do not know however what type of listings within one category are the driving forces for these growing number of listings, feedbacks, vendors and revenue. To understand how commodization of cybercrime components really takes place, we have to look at finer grained information. To do so, we next cluster listings within cybercrime component categories and characterize the supply by analyzing the best-selling clusters within each category.

6 Characterizing supply

We now want to delve deeper into what is actually being offered in each category and how this supply compares to the overall demand for criminal capability, resources and services in that category. We apply unsupervised clustering to the listings in each category and then interpret the three best-selling clusters.

6.1 Clustering listings

The detailed sub-classification is created by identifying clusters within our categories of listings using Topic Modeling. We rely on the Latent Dirichlet Allocation (LDA) [10] clustering algorithm to determine the main topics from a text corpus. We cleaned the data (removing broken fragments and correcting egregious errors) and lemmatized the words before clustering.

Our goal is to extract and analyze the three clusters which represent the "main themes" in each category. A natural choice might be to select the three clusters whose items collectively generate the largest amount of revenue. However we observed that a small fraction of very expensive items tends to obfuscate this analysis, thus we instead opted for identifying these "main theme" clusters based on the number of unique feedbacks. LDA is parameterized by a hyper-parameter that upper bounds the number of clusters to identify. Motivated by the expected heterogeneity of listings in the categories of cybercrime components, combined with the assumed homogeneity in other categories, we set this parameter to 10. As a consequence, it may be the case that LDA will not generate clusters for small categories of listings (when the true number of clusters exceeds 10); and those will instead be projected into larger clusters.

6.2 Best-selling clusters

We identified the three best-selling clusters per category by summing the number of feedbacks of all listings in a specific cluster. We then compute the total revenue generated by the item listings in each cluster. The results are shown in Table 6. We excluded three categories from the classification, as explained in Section 4.4. For all categories, the three best-selling clusters contain more than 46% of all feedbacks, and in many cases more than 60% of all feedbacks. Looking at revenue, we observe a diffused pattern. The categories "botnet," "website," and "RAT" show lower revenue numbers. Upon manual inspection, we could identify a very small cluster with only a few feedback that was dominated by a few very expensive items.

The second part of this clustering approach aims to understanding which type of products and/or services are transacted in these main clusters. To that end, we can use the output features of our LDA clustering algorithm to label the prominent clusters, sometimes assisted by manual inspection. In the next two sections, we present our findings and elaborate on whether the identified main topic clusters fit the overall demand for criminal capability, resources and services following our conceptual model.

6.3 Clusters in cash-out offerings

The main clusters of the cash-out category in descending order of size are 1) credit card details, more specifically "bins" - i.e., computer-generated credit card numbers that pass simple verification, but are not actually issued by banks, 2) so-called "fullz," stolen credit cards, includ-

[3] Interestingly, in underground slang, "pizza" may also denote credit card listings—which are sold in "slices." While this vernacular could *a priori* be confusing to an automated classifier, manual inspection suggests misclassification is very rare, as we will discuss later.

Table 6: **Best-selling clusters per category**

Category	# Feedback	Top3 Feedback	Top3 Revenue
App	1,175	784 (67%)	$ 7,083 (55%)
Botnet	968	657 (68%)	$ 8,995 (19%)
Cash-out	236,566	164,124 (69%)	$ 4,991,272 (63%)
E-mail	4,684	2,605 (56%)	$ 64,642 (66%)
Exploit	1,335	936 (70%)	$ 11,514 (65%)
Hosting	120	97 (81%)	$ 829 (70%)
Malware	2,446	1,127 (46%)	$ 30,806 (53%)
Phone	2,731	1,851 (68%)	$ 48,154 (65%)
RAT	768	501 (65%)	$ 4,887 (30%)
Website	8,586	5,044 (59%)	$ 65,111 (23%)
Account	75,469	47,149 (62%)	$ 316,851 (53%)
Fake	34,341	20,568 (60%)	$ 1,386,363 (48%)
Guide	57,361	38,586 (67%)	$ 2,397,006 (91%)
Pirated	11,242	6,093 (54%)	$ 55,864 (43%)
Voucher	22,769	13,643 (60%)	$ 441,572 (59%)

ing its full details, such as the CVV number. We can also identify a cluster pertaining to 3) guides on "making money," or money mule recruitment. Next to these three prolific clusters, we explore the seven other clusters in cash-out offerings, ordered by there relative feedback volume. We observe clusters with distinct offerings in 4) carding tutorials, 5) PayPal accounts, 6) Visa and Mastercard card details[4] , 7) "bitcoin deals," 8) bank account credentials, 9) Amazon refund guides and 10) Bitcoin exchanges, specialized in cash pay-outs. All in all, we can observe a broad spectrum of cash-out solutions being offered. They range from guides, to actionable solutions, like PayPal or bank account access. Next, we can discern services aimed at cashing out cryptocurrencies, more specifically Bitcoin, through dedicated exchange services. Consistent with what previous studies showed for cybercrime forums [23], carding makes up a big part of cybercrime components transacted on online anonymous markets as well.

6.4 Clusters in other B2B offerings

In this section we present the best-selling clusters in the other categories of B2B cybercrime components.

App. Prominent clusters of the App category include offers for Android loggers, i.e., malicious keylogger apps, Android bank apps, i.e., malicious banking apps, and Dendroid, a RAT for Android.

Botnet. Prominent clusters of botnet listings feature products and services revolving around Zeus botnets, varying from tutorials, to source-code, to "turn key" setups. We also identified offers on C&C servers and DDoS services.

E-mail. The prominent clusters in the e-mail category contain two types of spam lists, namely basic lists of e-mail addresses, as well as complete databases, including personal details to create personalized (spear) phishing mails. In addition we find a cluster of offerings on spam-related services.

Exploit. Within the exploit category, the two main themes are 1) Microsoft Office exploits, e.g., malicious macros, and 2) browser exploits. We also recorded a non-trivial set of sales for Mac exploits.

Hosting. The prominent "hosting" clusters include hosting through VPS or CPanel-listings. We also find a prominent cluster on hosting of Tor-based websites.

Malware. Within the malware category, ransomware stands out by featuring two prominent clusters. One cluster revolves around the Stampado ransomware, the other on Philadelphia ransomware. We also observed a prominent cluster on miscellaneous (assistive) software tools such as keyloggers or portscanners.

Phone. In the category of phone listings, one prominent cluster comprises listings on bypassing security features on phones. The other two prominent clusters offer respectively hacked Vodafone accounts and lists of usable phone numbers.

RAT. Two out of three prominent clusters in RAT listings contain generic RATs. The third cluster specifically deals with Mac OS RATs.

Website. The website category is composed of three distinct, prominent clusters. One cluster contains website development listings. The second is predominantly VPN-connections and/or SOCKS proxies. The third cluster consists of compromised RDP-servers/hosts listings.

Our analysis suggests that nearly all prolific clusters supply a component that matches B2B demand, but that this supply is incomplete, in that the observed supply fulfills only a niche demand in each category. For instance, we see ransomware dominating the malware category, whereas domain expertise suggests there are, in general, other types of malware in demand. This demand remains mostly unfulfilled in online anonymous marketplaces.

One exception to the aforementioned trend is in the "phone" category, where supply differs from the B2B demand. Research suggests that the actual latent demand is for using phones and social engineering to trick victims into falling for a scam [11]. Yet, the supply is only oriented towards setting-up the necessary phone lines. We observed that guides and tutorials are among the prominent clusters in the botnet and cash-out categories. We however note that selling a guide is not the same as outsourcing a cybercrime component.

In summary, the demand for cybercrime components is frequently met on online anonymous markets in our

[4]This cluster ressembles 1) and 2) but with a focus on Visa and Mastercard brands. It could a priori also include gift cards.

dataset, but the supply is highly restricted to specific niches and the accompanied revenue is generally modest.

6.5 Clusters in B2C offerings

In this section we briefly present the prominent clusters in the B2C categories – i.e., retail cybercrime.

Account. In listings that sell accounts, we observed two main clusters that revolve around offerings for single accounts to pornography websites. Next, we see a cluster of listings selling Netflix and Spotify accounts, in quantities between two and ten per listing.

Fake. The three prominent clusters are respectively offering fake passports, fake IDs and counterfeit money.

Guide. The clustering process revealed guides in a) bitcoin ("deals"), b) "making money" or starting a business, and c) "scamming."

Pirated. Miscellaneous pirated software, like the entire Adobe software suite or pirated adult videos, and pirated Microsoft software, e.g. Windows 7, are the prominent clusters in pirated products.

Voucher. In the category of voucher-related listings, we see offers for: a) Tesco vouchers, b) lottery tickets and c) "free" pizzas, of which most are indeed discount vouchers or gift cards for various pizza chains, but a few are in fact credit card offerings, where "slices" refer to groups of accounts.

The nature of products and services in all of the best-selling clusters tells us that we are observing transactions of retail cybercrime. We see that the best-selling clusters within accounts are listings in smaller quantities, ranging from single hacked accounts on a pornography website, to up to ten Netflix or Spotify accounts. It may at first appear to be curious why a single user would want 10 Netflix accounts, but when considering the inherent unreliability (and short lifespan) of stolen accounts, it becomes clear that this demand is plausible for personal use.

7 Discussion

In this section, we discuss our approach and results in light of our theoretical assumptions and research design.

7.1 Validation

In earlier work, Soska and Christin [40] discuss the validation of measurements on online anonymous markets. They find support for using feedback instances as a proxy for sales by looking at three specific cases where ground truth is available (due to arrests or leaks). However, the online anonymous marketplace ecosystem has grown quite significantly since then - in particular, in 2017, AlphaBay itself grossed, on a daily basis, more than the entire online anonymous marketplae ecosystem did in 2014.

The criminal complaint for forfeiture against the alleged AlphaBay founder and operator [5] estimates that "between May 2015 and February 2017, Bitcoin addresses associated with AlphaBay conducted approximately 4,023,480 transactions, receiving approximately 839,087 Bitcoin and sending approximately 838,976 Bitcoin. This equals approximately US$450 million in deposits to AlphaBay."

The estimates coming from our scrapes yield US $222,932,839 (and 2,223,992 transactions) for the entire time interval (including, this time, all of the goods sold on the marketplace). We believe the $450 million dollar from the complaint is a slight overestimate, due to currency mixing that could result in double-counting.

On the other hand, our own estimates are on the conservative side. In particular, we have to ignore a small fraction of credit card sales, due to a quirk in the way certain purveyors of credit card numbers do their business: A few stolen credit card number vendors list their items in generic form, with a price of zero, instead leaving the specifics in the shipping costs - presumably to obfuscate their stocks and possibly to reduce the commissions imposed by the marketplace operator. For instance, a listing would be for "credit card dumps," with a price of zero, but with shipping options for various types of cards at various prices. Because we cannot determine which cards are purchased, we simply conservatively ignore such sales.

More importantly, as Soska and Christin point out, it is important to repeatedly scrape online anonymous marketplaces to ensure adequate coverage [40]. This is particularly true when a marketplace is large, as the population of items is more likely to change over small time intervals. Our density of scrapes is lower in mid-2016, meaning that we might have missed a number of transactions occurring then.

All in all, we might be missing a non-negligible number of transactions occurring on AlphaBay; data for the other marketplaces is more complete, as validated in the original paper [40]. We point out, however, that these misses are unlikely to change our analysis beyond underestimating absolute sales volumes: indeed, with the small exception of the vendors using shipping costs for pricing, there are no specific biases in the missing data, so that the items we have in our corpora can be taken as a representative random sample.

7.2 Limitations

We next discuss the limitations of our study in two main areas: first, to what extent our data captures the commoditization of cybercrime and, second, the way we mapped the offerings on these markets onto categories of demand.

Observing cybercrime commoditization starts with knowing where to look. Building on transaction cost economics, we have argued that online anonymous marketplaces are the most logical place to trade cybercrime commodities due to the nature of these transactions. However, what seems logical from a TCE perspective does not necessarily seem logical to the criminal entrepreneur. Trust in a market is to a large extent subjective. This might mean that cybercriminals turn to other platforms with less safeguards to trade commoditized cybercrime. Even when criminals do follow TCE, some forms of commoditized cybercrime do not fit well with online anonymous markets: subscription models, affiliate programs, services requiring a rich search interface, or non-English offerings [26, 48] are all ill-suited to the type of markets we are investigating here. Since we did not study these forms of trade, our picture of commoditization is incomplete. To some extent, the same holds for underground hacker forums, though we would argue that many of the transactions on those forums are not actually commoditized, but forms of contracting (see Section 2).

Another limitation relates to how we mapped criminal demand. Successful commoditization is not just a matter of products and services being offered. These offerings also need to meet a demand, as observed in actual sales. To understand the potential demand of cybercriminals, we worked with a scope of known business models. Building on the work of Thomas et. al. [42], we have limited ourselves to cybercriminals who aim at making a profit. In other words, there may be cybercrime components that are being offered and that do match cybercriminal demand (e.g., for ideological or tactical purposes, rather than financial pursuits), yet are outside the identified value chains.

8 Related work

Core elements of our paper build on or benefit from recent progress in related research, which we discuss here.

Different researchers have tried to grasp the evolution of criminal activity in the underground economy. Initial work focused on underground forums [24, 36]. After the infamous Silk Road market came into existence, researchers looked closer at online anonymous markets [8, 15] and investigated the evolution of listings and revenue on these markets. Our study is among the first to explicitly leave the predominant drug listings out of scope and focus on a different product type (cybercrime). Most closely connected to our work is the first longitudinal study on the evolution of volumes in products transacted across multiple online anonymous markets by Soska and Christin [40]. Other studies focused on specialized markets or forums, for instance the stolen data and exploit market [9, 23]. They investigated the market for exploits - which turned out to be moderate in size - and the cybercrime-as-a-service market, where growing numbers of new services types were discovered. Furthermore, researchers investigated the increase in online drugs trade, specifically the B2B side of Silk Road 1 drugs offerings, and what factors determine vendor success [8].

In addition to quantitative studies of the evolution of online anonymous markets, our work is related to qualitative studies on buyers and sellers (vendors) on markets and forums. For instance, Van Hout and Bingham [25] looked into the buyers of drugs, and inspected the retail side of the market, as we did. Van Buskirk et al. [43] specifically focused on the motivation of drug buyers in Australia to turn to online anonymous markets instead of street dealers. They found that a cheaper price and higher quality of the drug are important.

Earlier research into the commoditization of cybercrime found evidence of commoditization of a number of specific products and services. Prominent examples are booters [29], the Pay-Per-Install (PPI) market [13], and exploit kit developers supplying drive-by browser compromise [22]. Thomas et al. [42] provided an overview of the prominent cybercriminal profit centers, based on multiple individual value chains such as spam [34], and clickfraud [32]. We can further identify earlier work on the value chains behind malware [38, 44] and carding [41].

Finally, our work can be tied to studies that aim to understand how and where cybercriminals collaborate. Leukfeldt et al. [33] investigated 40 cybercriminal networks using European and American police cases and interviews, Soudijn and Zegers [41] use data from a seized carding forum to unravel the collaboration between involved actors and Hutchings [28] studied the concept of co-offending in cybercrime and more specifically knowledge transmission amongst cybercriminals and identified distinct typologies of collaboration, ranging from fluid networks to real co-offending. In most cases, they found online meeting places, such as dedicated fora and markets, as the places where to buy tools or to collaborate with co-offenders.

9 Conclusions

We identified key value chain components that criminal entrepreneurs might want to outsource (i.e., purchase on

the market) and ordered them in ten categories. In three of them ("javascript," "customer service," and "web inject"), we found no offerings in the large random sample for the ground truth, not even when we searched the whole data with specific keywords. We assume this means there is very little, if any, commoditization of these value-chain components. In the other categories of cybercrime components, we found growing commoditization in terms of listings, vendors and revenue. Cash-out is by far the largest category. Some categories see only modest offerings and transaction volumes. Furthermore, not all offerings reflect the breadth of the demand. In some categories, only niche offerings are available.

In line with what other researchers have observed for the drugs trade on these markets, we see both B2B and B2C transactions in the cybercrime categories. B2B and B2C, a.k.a. retail cybercrime, turns out to be comparable in revenue. Between 2011 and 2017 the revenue of B2C cybercrime was around US $7 million, where B2B cybercrime generated US $8 million in revenue.

In conclusion, we find that, at least on online anonymous marketplaces, commoditization is a spottier phenomenon than was previously assumed. Within the niches where it flourishes, we do observe growth. That being said, there is no supply for many of the capabilities, systems and resources observed in well-known value chains. There is also no evidence of a rapid growth, and thus of a strong push towards commoditization, contrary to the somewhat alarmist language found in industry reporting and elsewhere.

In terms of generalizability of our findings, we have measured and explained the trends in commoditization of cybercrime on online anonymous markets. Beyond this, our findings only speculatively suggest that the trend toward commoditization might not be as comprehensive as has been claimed elsewhere. Perhaps less commoditized forms of B2B transactions - e.g., collaboration emerging out of forums - are important in the areas absent from the anonymous markets. Also, vertical integration probably remains important for more complex and dynamic forms of cybercrime.

Still, this casts an interesting perspective on the "theory of the commoditization of cybercrime." There is a huge discrepancy between the reported profitability of criminal business models like ransomware (over $1 billion in 2016, according to the FBI [20]) or DDoS-services (one youngster making $385,000 with his booter-service according to local British police [2]) and the marginal markets for cybercrime commodities. The commodities for a ransomware operation seem available in these markets: malware, PPI, cash-out. The huge profits would surely draw in new entrepreneurs to assemble this value chain based on components they can just buy on the anonymous markets. But if that would be the case, should that not cause a more observable rise in the commodities trade on these markets? The lack of strong growth suggests that there are still bottlenecks in outsourcing critical parts of criminal value chains. Entry barriers for would-be criminal entrepreneurs remain. The services that are highly commoditized, like booters, seem to draw in mostly B2C activities – i.e., consumers going after other consumers, as was the dominant finding in a victimization study of commoditized DDoS [37]. A recent takedown of a RAT operation also suggested consumer consumption, rather than B2B transactions [6].

This should not be read to downplay the relevance or danger of commoditization. A better understanding of where commoditization succeeds and fails helps to identify which capabilities, services and resources are still hard to come by, which supports designing better disruption strategies for criminal business models. The absence or scarcity of certain commoditized cybercrime components suggests these are either harder to produce or that they cannot function on their own after a single-shot sale. B2B services that require ongoing coordination among the criminals fall short of full-fledged commoditization. In other words, the scarcity of supply suggests less-scalable and potentially vulnerable components in criminal value chains. These might be targeted by interventions. Earlier work on interventions that target choke points shows that they can have measurable impact, not via a wholesale shutdown of the business model, but by raising transaction costs [12, 29]. For instance, we found virtually no offerings for customer support services. For a ransomware scheme, the customer service component to guide inexperienced victims through the steps to complete the ransomware payment might be the most vulnerable. Contrast this approach to the series of police actions aimed at the shutdown of whole markets: from our data, these operations seemed to have had only relatively modest effects on the overall trading of commoditized cybercrime. Understanding where commoditization is lagging behind points to alternative disruption strategies.

10 Acknowledgments

This research was partially supported by the MALPAY consortium, consisting of the Dutch national police, ING, ABN AMRO, Rabobank, Fox-IT, and TNO. This paper represents the position of the authors and not that of the aforementioned consortium partners. Kyle Soska and Nicolas Christin's contributions were partially sponsored by DHS Office of Science and Technology under agreement number FA8750-17-2-0188. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

References

[1] From marijuana to LSD, now illegal drugs delivered on your doorstep. *http://www.ibtimes.com/marijuana-lsd-now-illegal-drugs-delivered-your-doorstep-290021* (2011).

[2] Student pleads guilty to mass cyber attack. *https://www.bedsalert.co.uk/da/158731* (2016).

[3] AlphaBay, the Largest Online 'Dark Market', Shut Down. *https://www.justice.gov/opa/pr/alphabay-largest-online-dark-market-shut-down* (2017).

[4] Darknet Takedown, Authorities Shutter Online Criminal Market AlphaBay. *https://www.fbi.gov/news/stories/alphabay-takedown* (2017).

[5] United States of America vs. Alexandre Cazes – verified compaint for foreiture in rem, July 2017. United States District Court, Eastern District of California. Case 1:17-at-00557.

[6] International Crackdown on Anti-spyware Malware. *https://www.europol.europa.eu/newsroom/news/international-crackdown-anti-spyware-malware* (2018).

[7] ABLON, L., LIBICKI, M. C., AND GOLAY, A. A. Markets for Cybercrime Tools and Stolen Data. *National Security Research Division* (2014), 1–85.

[8] ALDRIDGE, J., AND DECARY-HETU, D. Not an 'Ebay for Drugs': The Cryptomarket 'Silk Road' as a Paradigm Shifting Criminal Innovation. *SSRN Electronic Journal 564*, October (2014).

[9] ALLODI, L. Economic Factors of Vulnerability Trade and Exploitation: Empirical Evidence from a Prominent Russian Cybercrime Market. In *CCS'17* (2017), no. 2.

[10] BLEI, D. M., EDU, B. B., NG, A. Y., EDU, A. S., JORDAN, M. I., AND EDU, J. B. Latent Dirichlet Allocation. *Journal of Machine Learning Research 3* (2003), 993–1022.

[11] BOGGS, N., WANG, W., MATHUR, S., COSKUN, B., AND PINCOCK, C. Discovery of emergent malicious campaigns in cellular networks. In *Proceedings of the 29th Annual Computer Security Applications Conference on - ACSAC '13* (New York, New York, USA, 2013), ACM Press, pp. 29–38.

[12] BRUNT, R., PANDEY, P., AND MCCOY, D. Booted: An Analysis of a Payment Intervention on a DDoS-for-Hire Service. In *Workshop on the Economics of Information Security (WEIS)* (2017).

[13] CABALLERO, J., GRIER, C., KREIBICH, C., AND PAXSON, V. Measuring Pay-per-Install: The Commoditization of Malware Distribution. In *Usenix Security Symposium* (2011).

[14] CHAWLA, N. V., BOWYER, K. W., HALL, L. O., AND KEGELMEYER, W. P. SMOTE: Synthetic minority oversampling technique. *Journal of Artificial Intelligence Research 16* (2002), 321–357.

[15] CHRISTIN, N. Traveling the Silk Road: a measurement analysis of a large anonymous online marketplace. In *Proceedings of the 22nd international conference on World Wide Web* (2013), pp. 213–224.

[16] CHRISTIN, N., YANAGIHARA, S., AND KAMATAKI, K. Dissecting one click frauds. In *Proc. ACM CCS'10* (Chicago, IL, Oct. 2010), pp. 15–26.

[17] COVA, M., LEITA, C., THONNARD, O., KEROMYTIS, A., AND DACIER, M. An analysis of rogue AV campaigns. In *Proc. RAID 2010* (Ottawa, ON, Canada, Sept. 2010).

[18] DEEK, F. P., AND MCHUGH, J. A. M. *Open source: Technology and policy*. Cambridge University Press, 2007.

[19] DICK, A. R. When does organized crime pay? A transaction cost analysis. *International Review of Law and Economics 15*, 1 (1995), 25–45.

[20] FBI. 2016 Internet Crime Report. Tech. rep., FBI, 2017.

[21] FLORENCIO, D., AND HERLEY, C. Phishing and money mules. In *International Workshop on Information Forensics and Security (WIFS)* (2010).

[22] GRIER, C., BALLARD, L., CABALLERO, J., CHACHRA, N., DIETRICH, C. J., LEVCHENKO, K., MAVROMMATIS, P., MCCOY, D., NAPPA, A., PITSILLIDIS, A., PROVOS, N., RAFIQUE, M. Z., RAJAB, M. A., ROSSOW, C., THOMAS, K., PAXSON, V., SAVAGE, S., AND VOELKER, G. M. Manufacturing Compromise: The Emergence of Exploit-as-a-Service. In *ACM Conference on Computer Communications Security* (2012).

[23] HOLT, T. J. Exploring the social organisation and structure of stolen data markets. *Global Crime 14*, 2-3 (2013), 155–174.

[24] HOLZ, T., ENGELBERTH, M., AND FREILING, F. Learning More About the Underground Economy: A Case-Study of Keyloggers and Dropzones. In *Computer Security–ESORICS* (2009).

[25] HOUT, M. C. V., BINGHAM, T., VAN HOUT, M. C., AND BINGHAM, T. "Surfing the Silk Road": A study of users' experiences. *International Journal of Drug Policy 24*, 6 (2013), 524–529.

[26] HUANG, D. Y., ALIAPOULIOS, M. M., LI, V. G., INVERNIZZI, L., MCROBERTS, K., BURSZTEIN, E., LEVIN, J., LEVCHENKO, K., SNOEREN, A. C., AND MCCOY, D. Tracking Ransomware End-to-end. In *IEEE Symposium on Security and Privacy (S&P)* (2018).

[27] HUANG, D. Y., DHARMDASANI, H., MEIKLEJOHN, S., DAVE, V., GRIER, C., MCCOY, D., SAVAGE, S., WEAVER, N., SNOEREN, A. C., AND LEVCHENKO, K. Botcoin: Monetizing Stolen Cycles. In *Proceedings 2014 Network and Distributed System Security Symposium* (2014).

[28] HUTCHINGS, A. Crime from the keyboard: Organised cybercrime, co-offending, initiation and knowledge transmission. *Crime, Law and Social Change 62*, 1 (2014), 1–20.

[29] KARAMI, M., PARK, Y., AND MCCOY, D. Stress testing the booters: Understanding and undermining the business of DDoS services. In *Proceedings of the 25th International Conference on World Wide Web* (2016), International World Wide Web Conferences Steering Committee, pp. 1033–1043.

[30] KHARRAZ, A., ROBERTSON, W., BALZAROTTI, D., BILGE, L., AND KIRDA, E. Cutting the Gordian knot: A look under the hood of ransomware attacks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2015), vol. 9148, pp. 3–24.

[31] KOLLOCK, P., AND BRAZIEL, E. R. How not to build an online market: the sociology of market microstructure, 2006.

[32] KSHETRI, N. The Economics of Click Fraud. *IEEE Security & Privacy Magazine 8*, 3 (5 2010), 45–53.

[33] LEUKFELDT, R., KLEEMANS, E., AND STOL, W. The Use of Online Crime Markets by Cybercriminal Networks: A View From Within. *American Behavioral Scientist* (2017), 000276421773426.

[34] LEVCHENKO, K., PITSILLIDIS, A., CHACHRA, N., ENRIGHT, B., FELEGYHAZI, M., HALVORSON, T., KANICH, C., KREIBICH, C., LIU, H., MCCOY, D., WEAVER, N., PAXSON, V., VOELKER, G. M., AND SAVAGE, S. Click trajectories: End-to-end analysis of the spam value chain. In *IEEE Symposium on Security and Privacy* (2011), IEEE.

[35] MIRAMIRKHANI, N., STAROVXI, O., AND NIKIFORAKIS, N. Dial one for scam: A large-scale analysis of technical support scams. In *Proceedings of the 24th Network and Distributed System Security Symposium (NDSS)* (Feb. 2017).

[36] MOTOYAMA, M., MCCOY, D., LEVCHENKO, K., SAVAGE, S., AND VOELKER, G. M. An Analysis of Underground Forums. In *ICM* (2011).

[37] NOROOZIAN, A., KORCZYŃSKI, M., GAÑAN, C. H., MAKITA, D., YOSHIOKA, K., AND VAN EETEN, M. Who gets the boot? analyzing victimization by DDoS-as-a-service. In *International Symposium on Research in Attacks, Intrusions, and Defenses* (2016), Springer, pp. 368–389.

[38] ROSSOW, C., DIETRICH, C., AND BOS, H. Large-Scale Analysis of Malware Downloaders. In *Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer, 2013, pp. 42–61.

[39] SOOD, A. K., BANSAL, R., AND ENBODY, R. J. Cybercrime: Dissecting the State of Underground Enterprise, 2013.

[40] SOSKA, K., AND CHRISTIN, N. Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem. *24th USENIX Security Symposium*, August (2015), 33–48.

[41] SOUDIJN, M. R. J., AND ZEGERS, B. C. H. T. Cybercrime and virtual offender convergence settings. *Trends in Organized Crime 15*, 2-3 (2012), 111–129.

[42] THOMAS, K., HUANG, D. Y., WANG, D., BURSZTEIN, E., GRIER, C., HOLT, T. J., KRUEGEL, C., MCCOY, D., SAVAGE, S., AND VIGNA, G. Framing Dependencies Introduced by Underground Commoditization. In *Workshop on the Economics of Information Security (WEIS)* (2015).

[43] VAN BUSKIRK, J., ROXBURGH, A., BRUNO, R., NAICKER, S., LENTON, S., SUTHERLAND, R., WHITTAKER, E., SINDICICH, N., MATTHEWS, A., BUTLER, K., AND BURNS, L. Characterising dark net marketplace purchasers in a sample of regular psychostimulant users. *International Journal of Drug Policy 35* (9 2016), 32–37.

[44] VAN WEGBERG, R. S., KLIEVINK, A. J., AND VAN EETEN, M. J. G. Discerning Novel Value Chains in Financial Malware. *European Journal on Criminal Policy and Research 23*, 4 (12 2017), 575–594.

[45] WILLIAMSON, O. E. Transaction-Cost Economics: The Governance of Contractual Relations. *Journal of Law and Economics 22*, 2 (1979), 233–261.

[46] WILLIAMSON, O. E. Transaction cost economics and business administration. *Scandinavian Journal of Management 21*, 1 (2005), 19–40.

[47] WILLIAMSON, O. E. Transaction Cost Economics: An Introduction. *Economics Discussion Paper* (2007), 0–33.

[48] ZHUGE, J., HOLZ, T., SONG, C., GUO, J., HAN, X., AND ZOU, W. Studying Malicious Websites and the Underground Economy on the Chinese Web. In *Managing Information Risk and the Economics of Security*, M. E. Johnson, Ed. 2009, pp. 225–244.

Appendix

Algorithm 1: Classifier of cybercrime listings

Input: Listings from 8 marketplaces
Output: Set of listings per cybercrime categories

Select a random sample of 1,500 listings;
Manually classify random sample into cybercrime categories;
Split the random sample into a training (70%) and a testing set (30%);
forall *listings* **do**
    Remove English stopwords, URLs, punctuation and digits;
    Lemmatize;
    Tokenize;
foreach *category* $\in$ *training set* **do**
    Balance category using SMOTE method;
Train a Linear Support Vector Classifier using the listings in the balanced categories;
foreach *listing* $\in$ *testing set* **do**
    Classify according to the trained LinearSVC;
Compute confusion matrix;
forall *listings* **do**
    Classify according to the trained LinearSVC;