RESEARCH-ARTICLE

Cheaper than you thought? A dive into the darkweb market of cyber-crime products

DIMITRIOS GEORGOULIAS, Aalborg University, Aalborg, Nordjylland, Denmark

RICARDO YABEN, Technical University of Denmark, Lyngby, Hovedstaden, Denmark

EMMANOUIL VASILOMANOLAKIS, Technical University of Denmark, Lyngby, Hovedstaden, Denmark

Open Access Support provided by:

Aalborg University

Technical University of Denmark

Published: 29 August 2023

Citation in BibTeX format

ARES 2023: The 18th International Conference on Availability, Reliability and Security
August 29 - September 1, 2023
Benevento, Italy

ARES '23: Proceedings of the 18th International Conference on Availability, Reliability and Security (August 2023)
https://doi.org/10.1145/3600160.3605012
ISBN: 9798400707728

Cheaper than you thought? A dive into the darkweb market of cyber-crime products

Dimitrios Georgoulias
Aalborg University
Copenhagen, Denmark

Ricardo Yaben
Technical University of Denmark
Kongens Lyngby, Denmark

Emmanouil Vasilomanolakis
Technical University of Denmark
Kongens Lyngby, Denmark

ABSTRACT

The darkweb is nowadays considered a very popular place to sell and buy illegal cyber-crime related content. From botnet services and malware, to user data such as credit card information and passwords, darkweb marketplaces offer ease of use, product variety, and most importantly effective anonymity to both buyers and vendors. In this paper, we crawl 8 popular darkweb marketplaces and perform a comprehensive quantitative analysis with a focus on cyber-crime related products. Moreover, we report some preliminary findings when examining the same marketplaces through their I2P mirrors. Our results suggest that overall there is a multitude of products that fall into the cyber-crime category, with products under the Fraud category dominating the market, and that the average cyber-crime products' price is relatively low. Furthermore, we explore how the vendors of this specific product group are distributed across platforms, utilizing harvested information such as usernames and PGP keys, and investigate how their reputation scores affect their operation.

KEYWORDS

darkweb, crawling, I2P, Tor

ACM Reference Format:
Dimitrios Georgoulias, Ricardo Yaben, and Emmanouil Vasilomanolakis. 2023. Cheaper than you thought? A dive into the darkweb market of cyber-crime products. In *The 18th International Conference on Availability, Reliability and Security (ARES 2023), August 29–September 01, 2023, Benevento, Italy.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3600160.3605012

1 INTRODUCTION

The darkweb has been serving as an effective means of providing anonymity to Internet users since the *Tor Project* [13] was founded in 2006, and especially after the Tor browser's release in 2008, which made the Tor network's usage more approachable to users. The implementation of hidden services by Tor, services only available inside the Tor network, raised its obfuscation capabilities to a much higher level. However, since this feature was made publicly available to all users, it meant also being open to utilisation by malicious parties.

The first darkweb marketplace surfaced in 2010, known as *The Farmer's Market*, followed by *Silk Road*, both serving as meeting places and selling platforms for users interested in acquiring illegal narcotics [17]. More recently however, there has been an increasing demand for products and services which can be utilized to compromise the privacy and security of digital systems and their users [2, 33]. These products and services, which in this paper we will be referring to as *Cyber-Crime Related Products (CCRPs)*, include stolen user account information (e.g., bank information, credit card details and online credentials), fraudulent documents (e.g., forged IDs and driver's licenses), malicious software (e.g., malware, password crackers and zero-day exploits), and other cyber-crime oriented services (e.g., background checks, booter services and phishing campaigns). In this ecosystem, cyber-criminals trade CCRPs anonymously thanks to the obfuscation properties of Tor, cryptocurrency usage (with *Bitcoin (BTC)* and *Monero (XMR)* at the forefront), and use of *Pretty Good Privacy (PGP)* encryption for the authentication and communication between the entities involved [18].

The main consequence of the commoditization of CCRPs is the reduction of costs and resources needed from a cyber-crime business model perspective [19, 29]; this leads to potentially more advanced threats and larger attacks. Continuously observing darkweb marketplaces provides valuable insight on the framework that surrounds this type of cybercrime, such as product and service details, trends, attack vectors, and vendor accumulation. Knowing the value of CCRPs traded in the darkweb market contributes towards assessing the development of cyber-criminal operations. Over the years, there have been efforts to explore the various types of products and services available on the darkweb market. Some aim at specific types of services, such as the underground drug [12] or firearm [27] trade, while others take a broader approach [18, 28], acquiring as much information from these platforms as possible. These approaches have been both quantitative (e.g., [6, 26, 30]), using crawlers to harvest data from the platforms, and qualitative (e.g., [5, 18]), based on the manual navigation of the sites. However, there is a significant lack of quantitative research on the specific trade of CCRPs.

In this paper, we present a quantitative analysis on the data gathered between January and June 2022 from 8 of the largest and most popular darkweb marketplaces containing CCRPs. Our dataset includes a total of 39,881 products and 1,202 vendors. The main contributions of this paper are the following: First, we present a crawler for darkweb marketplaces capable of circumventing many anti-crawling mechanisms. In addition, we conduct a preliminary analysis between Tor and I2P darkweb marketplace deployments, showing that crawling through I2P averages at a faster speed than Tor. Second, we utilize two Natural Language Processing (NLP) models to identify and classify CCRPs, followed by an analysis

ARES 2023, August 29–September 01, 2023, Benevento, Italy

ACM ISBN 979-8-4007-0772-8/23/08...$15.00
https://doi.org/10.1145/3600160.3605012

of the identified 19,175 Cyber-Crime Related Product (CCRP) listings, demonstrating that CCRPs are generally inexpensive, which increases their accessibility from potential buyers. Based on this analysis, we present the pricing details of each specific CCRP category for each of the eight crawled marketplaces. Lastly, we correlate vendor activity across these eight marketplaces using their usernames and PGP keys. Our data suggest that 78% of the vendor usernames appear in multiple platforms, while 12.98% of the PGP keys appear in more than one marketplace. Lastly, we also find that the majority of vendors present a higher than 65% positive feedback score.

The remainder of this paper is structured as follows. Section 2 discusses the related work. In Section 3 we go over the methods followed for the purposes of this work. In Section 4 we present and discuss our results, and finally, Section 5 concludes this paper.

2 RELATED WORK

Previous studies have emphasised on analysing the darkweb economy through dedicated crawling. In this section, we go over notable related works that perform both generic and product-specific studies, while following quantitative, as well as qualitative approaches.

2.1 Generic darkweb studies

Bayoumy et al. [16] shared their observations through a netnographic study on the darkweb marketplaces *Wall Street Market*, *Dream*, and the darkweb site *Intel Exchange*. The focus of their effort is the distribution of ransomware and the relationships between vendors and their products, reaching the conclusion that the reputation of most vendors has been built from selling products not necessarily related to ransomware.

Zhou et al. [33] collected data from the *Dream* marketplace. In their study, they estimate the income of the marketplace, and the type of products being sold by analysing almost 2M items. In addition, they describe the methods and challenges faced during the data collection process, including the development of a CAPTCHA solver for this specific marketplace.

Ball and Broadhurst [4] proposed a set of generic tools to crawl data from darkweb marketplaces with promising results. Moreover, *Yannikos et al.* [31] present a dedicated crawler for the *White House Market*. The crawler follows a similar structure to the crawler introduced in [4], iterating through the marketplace items found in each category page, and storing the raw and mined data in a separate database. The crawler includes a CAPTCHA solver that reduces the input needed to start the crawling session. CAPTCHA solving on the darkweb was also the topic of *Audran et al.* [3].

Benjamin et al. [5] focused on forums and other channels to identify potential cyber threats. For this purpose, they used a crawler capable of identifying potential threats using a dictionary of weighted keywords. Furthermore, they proposed a comprehensive analytical framework to evaluate threats from posts and messages, and the malicious actors involved.

In their recent work, *Cuevas et al.* [10] expand on some of the key points introduced by *Pastrana et al.* [26] (that focused on clearnet marketplaces and forums) to investigate the accuracy of crawling measurements. They conclude that a larger number of crawls, combined with higher crawling frequency, results in a more accurate overview of the marketplace's state. In this paper, we crawl each market continuously in a single snapshot to achieve a higher resemblance of the state of the markets.

Lastly, *Georgoulias et al.* [18] take a qualitative approach and map the key elements and properties that surround the operation of darkweb marketplaces, forums, and vendor shops, with the goal of gaining insight on the darkweb trade infrastructure.

2.2 Product-specific darkweb studies

There is a multitude of quantitative research covering vendors and products advertised on darkweb (and clearnet) platforms. However, most authors choose drug related products for their case study [22], as this type of product is the most predominant and most lucrative on the market.

Wang et al. [30], present an interesting investigation and comparison between English and Chinese marketplaces. For their study, they used a simple crawling setup to collect and analyse the listings from five different marketplaces. Their results show that Chinese marketplaces are generally less advanced compared to the English ones, both in terms of marketplace activity and policies (e.g., Chinese marketplaces are less restrictive).

Bracci et al. [7] used a dataset with close to 1M items from 30 marketplaces, looking for COVID-19 products, with the aim of uncovering the market for these products emerging in the darkweb.

Georgoulias et al. [20] investigate the market for COVID-19 vaccination certificates in marketplaces and vendor shops in the darkweb through qualitative methods. They showcase the availability and validity of these products, and also develop a taxonomy of certificate forging capabilities, from the side of the vendors.

2.3 Cyber-crime product-specific darkweb studies

The amount of work that studies CCRPs is limited and follows similar methodologies to identify CCRPs. The most common approach is to perform a sentiment analysis on the title and description of the products, resulting in a likelihood score that estimates the probability of a product being a CCRP [11, 25, 29]. In 2015, *Macdonald et al.* [23] performed the first sentiment analysis to identify conversations of hackers targeting critical infrastructure in darkweb forums. Their approach was to assign a sentiment score based on the repeated appearance of nouns from a dictionary of hacker jargon. *Deb et al.* [11] considered more than 100 forums to build a historical reference of events based on sentiment. *Nunes et al.* [24] replaced the dictionaries with semi-supervised *Support Vector Machine (SVM)* models to identify cyber threats in marketplaces and forums.

More recently, researchers have incorporated *Natural Language Processing (NLP)* models to improve the performance of their classification method. *Ebrahimi et al.* [15] expand on previous studies by incorporating non-English marketplaces. Their approach is based on *BiLSTM* [9], an NLP model capable of transferring contextual knowledge from multiple languages, without relying on translation tables. The authors published a similar work at a later date reusing *BiLSTM* models to classify products based on their description [14]. This time, they were able to classify more than 79,000 products from multiple English-based marketplaces. Their model reached an

F1-Score of 89.55%, using less than 3% human-labelled records to train the model.

However, sentiment analysis cannot be used to classify CCRPs (e.g., to differentiate malware from databases). To overcome this problem, previous work used the category names in which the products were found. This may result in redundant labels due to the different wording used on each market.

3 METHODOLOGY

In this section, we introduce our approach to collecting and processing data. Then, we continue with an overview of the data collection system, describing features that make this system viable for our use case. Finally, we describe the two NLP models used to label and categorise our dataset.

3.1 Heuristics

Overall, the following heuristics were used as a medium to simplify decision-making and general tasks.

Market selection. The criteria for choosing these specific eight marketplaces were their popularity, size, and promotion on several introduction points/onion service directory pages [1] . In the selection process, we decided not to include marketplaces that were unlikely to contain CCRPs (e.g., marketplaces dedicated to drugs), or based in languages other than English. We include the remaining of the most popular and largest darkweb marketplaces commonly known at the time of this study. The complete list of popular marketplaces can be seen in Table 1. The table shows whether the markets contain CCRPs and through which networks can be accessed.

Category selection. We intentionally targeted categories presumable to contain CCRPs (e.g., malware or digital products), crawling categories by their size in descending order. Larger categories tend to change faster than smaller ones; therefore, it is important to monitor larger categories more frequently. While we collected all of the products in the categories containing CCRPs, we did not collect products from the rest of the categories (e.g., drugs).

Holding prices. This term refers to techniques used by vendors to advertise their products differently than intended by the marketplace. For example, setting the price of the product to the maximum price allowed on the site to indicate the vendor is restocking, while setting advertisements without price implies the price is set upon agreement. The latest example is frequently used in combination with client-specific ads and tip-jar ads. To remove *holding prices* from our dataset, we use a similar heuristic to the one developed by [28] based on the probabilistic distribution of the observations. While they decided to remove larger prices than five times the median, our dataset contains interesting products with high prices, hence, we remove prices below the 25 percentile and greater than the 75 percentile.

3.2 Crawler

To the best of our knowledge, there is no crawler that can bypass CAPTCHA and login challenges. In addition, available darkweb crawlers [1, 8, 26, 33] are dedicated to specific marketplaces or forums, and present major design limitations that make customisation unfeasible. One major issue is that they are built to create snapshots of the target, instead of a live replica of the site. Moreover, these crawlers have only partially addressed issues regarding anti-crawling mechanisms, such as solving CAPTCHAs or mimicking human behaviour [32]. Cuevas et al. [10] voice these issues, showing that most of the research tends to focus on the results rather than on the methods. In this context, we developed a crawler for darkweb marketplaces capable of circumventing some anti-crawling mechanisms and tracking listings through continuous crawling.

Marketplace	I2P	Tor	CCRPs
ASAP	✗	✓	✓
Alphabay	✓	✓	✓
Cypher	✗	✓	✓
DarkFox	✗	✓	✓
Kingdom	✓	✓	✓
Tor2Door	✗	✓	✓
Versus	✗	✓	✓
Vice City	✗	✓	✓
Archetyp	✗	✓	✗
Bohemia	✗	✓	✗
Incognito	✗	✓	✗

Table 1: List of most popular darkweb marketplaces, accessible from Tor and/or I2p, and whether they include CCRPs.

This crawling system includes three services, one for each of the duties of a common crawler: crawling, scraping, and storing content. These services are written in Python to benefit from a large number of tools available for crawling and scraping web content. The crawler uses configuration files with validation rules to capture undesired HTML elements (e.g., CAPTCHA challenges, authentication inputs, or path-killed messages), elements always included (e.g., product advertisement cards in each category page), and other relevant elements (e.g., links to product pages). When an undesired element is found and the crawler is not capable of handling the situation, the crawler will prompt for human intervention. To mimic human behaviour, the crawler includes a budget that randomises requesting times and set a limit to the number of simultaneous connections with the market. Moreover, the content of each page is stored locally, which is scraped eventually for relevant data points and stored in a database. The crawler's design as well as the crawling process are illustrated in Figure 1 and described as follows.

Crawl pending vendors and product pages. The crawler starts two parallel processes to crawl pending vendor and product pages. This feature is meant to overcome accessibility issues and resume from a previous crawling point.

Crawl category pages with unseen products. The crawler iterates over category pages containing unseen products, starting at the last-crawled index page. Since the market activity may shift the index of this page, the crawler includes a window feature to continue crawling a fixed amount of pages before stopping. The number of pages is determined by the time gap between crawls, with a minimum of 2 pages considering busier categories, and a maximum of 5 pages for efficiency.

Scrape content. The scraper pre-processes and data mines raw vendor and product pages.

[1] These platforms index popular Tor and I2P hidden services, ranging from markets, to blogs and forums [18]. Some examples are *TorTaxi*, the *HiddenWiki*, *Recon*, and *Dark.fail*.

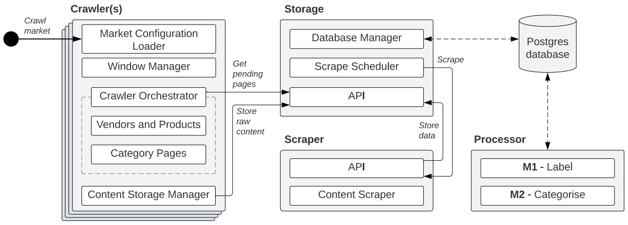

Figure 1: Darkweb crawler architecture design.

3.3 Identifying cyber-crime products

As previously seen in Section 2, recent studies use NLP models, which are better suited for text-based tasks that require an understanding of the context. Following this methodology, we trained two NLP models to identify (*M1*) and classify (*M2*) CCRPs using product descriptions instead of titles, which typically provide more information on the product. Table 2 shows the labels used for the two models and the criteria used for each label. The label and criteria used for *M2* follow a generalised naming convention among darkweb markets for categories and sub-categories.

Model	ID	Label	Content
M1	A1	Positive	Cyber-Crime Related Products
	A2	Negative	Unrelated Products
M2	B1	Media	Tutorials, E-Books and Videos
	B2	Malware	Keyloggers, RAT, Ransomware, Botnets and Exploits
	B3	Software	Tools and Security Applications
	B4	Fraud	PII, Credentials, Tokens, Documents and Databases
	B5	Services	Hosting, DDoS, Spam and Carding

Table 2: The two NLP models and their labels.

Following the criteria introduced in Table 2, we manually labelled 10% of the records (4,000 evenly divided among *M1* labels) to train and compare candidates for *M1*, where a record is the description of a product. Initially, we attempted using 3% and 5% manually labelled records following the methodology of [14]; however, we could not reach similar F1-Scores using less than 10% labelled records. For *M2*, we used the remaining 5% positive records (2,000), evenly divided among categories. Candidates were selected from the list of available NLP pre-trained models in "Simple Transformers" from [21], based on the best performance in the sentiment analysis (*M1*) and text classification (*M2*) tasks.

Table 3 includes the list of fine-tuning parameters used to train the candidates, which were chosen according to the size of the dataset and the maximum sample size (descriptions were not pre-processed). It is worth mentioning that non-deterministic parameters (e.g., early stopping) influence the result of the training. Thus, for a fair comparison, all models were trained twice with the same parameters and compared using their best performance run in terms of their *F1-Score*. The models were trained using K-fold ($K = 5$) cross-validation, each set divided as 80% training and 20% validation for both runs, randomly divided from the set of manually labelled records. The results from the top three candidates for both models can be seen in Table 4. Once the models labelled the dataset, we manually verified the results and applied fixes on mislabelled records.

Parameter	Value	Description
Training batch size	16	Batch size used for training
Evaluation batch size	64	Evaluates while training using a batch of the set
Warm up steps	600	Number of first steps using a lower learning rate
Learning rate	3e-5	Determines the impact of newly acquired information
Weight decay	0.01	Weight penalisation
Early stopping	True	Stops the learning on the current batch earlier than expected
Sliding window	True	Creates windows for inputs overly large

Table 3: Training parameters for the ML models.

For *M1*, the F1-Score variance between the models is not significant enough to declare a clear winner. While at a first glance *BERT base* appeared to be the better candidate, *DistilBERT* showed a smaller disparity between training and validation loss, suggesting that *DistilBERT* performed better at generalising. *DistilBERT* also outperformed other models in training duration and labelling speed. Therefore, we choose DistilBERT as the best candidate for the *M1* labelling model with approximately 93% F1-Score. It is important to mention that the models reached a point of almost zero learning after three epochs, suggesting that they adapted to the dataset rather quickly. We experimented with reducing the learning rate and increasing the number of epochs but did not yield significant improvements.

Model	Candidate	T. Loss	Accuracy	F1-Score	Duration
M1	DistilBERT base uncased SST-2	0.246	91.0%	92.9%	1h 37min
	BERT base	0.232	91.2%	93.3%	7h 50min
	RoBERTa base OpenAI	0.198	89.2%	91.0%	4h 4min
M2	BERT base	0.758	77.2%	76.1%	5h 24min
	XtremeDistil 16 h384 uncased	0.984	73.6%	70.9%	2h 51min
	MiniLM L12 H384 uncased	1.913	69.3%	63.0%	1h 41min

Table 4: Training results for M1 and M2 model candidates.

In comparison with the results from *M1*, *M2* F1-Scores are significantly lower. Moreover, *M2* models reach a point of zero learning after five epochs, while producing a much higher training loss. This means that it is much harder to find similarities between records of certain types, especially among those with fewer records and high entropy. We choose *BERT base* for *M2* because it outperformed the other candidates in terms of F1-Score by a large margin. However, the variance in training loss over time suggests that a larger dataset could improve the performance of the model.

4 RESULTS

In this section, we first cover our experimental setup to collect data from the eight selected darkweb marketplaces during the period of January-June 2022. Then, we introduce our preliminary crawling speed benchmark results between I2P and Tor. Next, we apply quantitative methods to analyse the collected CCRPs categorised into five groups. Finally, we cross-reference vendor information across darkweb marketplaces to study their presence in multiple marketplaces.

4.1 Experimental setup

For this study, the setup included 4 instances of the crawler service, one parser, and one storage service connected to a *Postgres* database. Due to the need for manual input to solve CAPTCHA challenges, we could handle up to 4 instances simultaneously without major interruptions. This also affects the completeness of our dataset, since products may be added and removed during the time between sessions and the time it takes for the crawler to map the market. Prior studies [10] consider this issue to be significant for general products; however, our targets are mostly digital products and services, therefore we consider such missing listings to be marginal. In addition, the availability of darkweb markets changes rapidly, directly impacting when we could crawl them. For example, a few markets were randomly unavailable during the day; ASAP closed for two weeks, and Versus closed permanently towards the end of the crawling phase.

Each crawler instance was allowed to maintain a maximum of 5 to 10 parallel connections with the same marketplace, and request delays lower than 5 seconds per request for uninterrupted sessions of 30 minutes in the most strict marketplaces. We note that marketplaces with I2P mirrors had a higher tolerance in terms of the frequency of requests per second before requiring input. Furthermore, we observed that marketplaces lowered their tolerance at certain times during the day and/or on specific dates. As an example, when a marketplace claimed to be under attack, it prompted for authentication more frequently.

4.2 I2P vs Tor: a preliminary study

From our previous work on darkweb platforms [citation anonymized], we were aware that marketplace administrators had started considering making their sites available also in the I2P network, for reasons such as availability, security, and service access speed. With a migration to the I2P network showing promise, we decided to dedicate part of this work to acquiring preliminary results on crawling the I2P mirrors of these marketplaces. We believe this effort offers insight and prepares researchers to adjust their methodology in case more I2P marketplaces appear in the future. However, a comprehensive comparison between I2P and Tor is rather difficult to achieve, since, at the time of our crawling, the only two marketplaces with I2P mirrors were *Alphabay* and *Kingdom*, and Tor mirrors up-time is much lower when compared to I2P mirrors.

First, analysis of our crawled data suggests that the content of the marketplaces is identical in both networks. We noticed this when we attempted to crawl the marketplaces from both networks; the crawler would skip all pages since they were already recorded. We also empirically noticed that the crawler instances using Tor were slower at crawling than the I2P counterparts. In addition, crawler instances using I2P mirrors could crawl for longer periods of time, while Tor mirrors were experiencing downtime. Overall, I2P mirrors were more reliable than Tor mirrors, thus, we could crawl faster using I2P than Tor.

To formalize the aforementioned findings, we conduct a benchmarking experiment to study the disparity between I2P and Tor crawls in a controlled setup. The aim is to record a crawling session on *Alphabay* from I2P and Tor before needing to re-authenticate in the marketplace, or for a maximum of 30 minutes, similarly to the average among markets. For this, we used a single crawler instance allowing for 10 simultaneous connections, and a maximum of 5 seconds of delay per request. In order to give a fair chance to both networks, we repeated this process 5 times per network. Figure 2 depicts the results of this experiment, showing the average for both networks. The response time was recorded in seconds, as the time between requesting an URL and receiving the response. The timestamp is represented as the minutes since the experiment started. In general, I2P connections averaged a lower response time (5.6 seconds) than Tor connections (9.1 seconds). None of the experiments triggered the need for re-authentication within the 30-minute window.

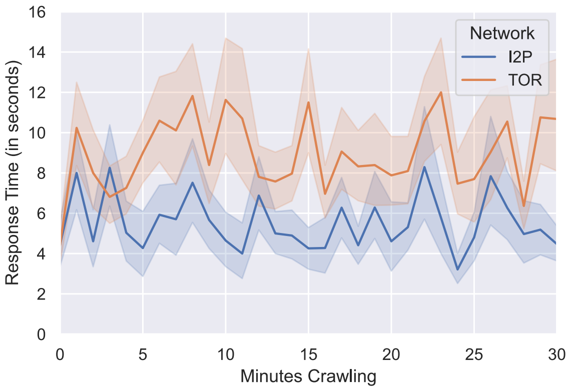

Figure 2: Benchmark experiment of the response times between I2P and Tor from Alphabay using an aggressive crawling setup.

4.3 Products

To locate CCRPs in the 8 targeted marketplaces, we deployed our crawler on 5 specific product categories, namely *fraud, malware, media, services,* and *software.* After applying the heuristics discussed in Section 3.1, we summarise our results on Table 5. It is evident that there is a large disparity between the total number of listings, and the positive results containing CCRP listings. This is mainly due to the fact that these categories also contain large quantities of products such as forged documents (e.g., passports, driving licences, IDs) and banknotes (e.g., dollar and euro bills).

As depicted on Table 5, the majority of listings were found in the *DarkFox* marketplace (6,604), which can be attributed to its larger size compared to the other 7 platforms. This applies to all listing types apart from *Services*, with the *Versus* marketplace occupying the biggest part of the market. The *Tor2Door* marketplace is the one with the highest ratio of CCRP listings, which take up 57.9% of the products in the 5 scraped categories, while *ASAP* presents the lowest number of CCRPs at a total of 32.3%. In terms of the types of

Marketplace	Fraud	Malware	Media	Services	Software	Total CCRPs	Total *Incl. non-CCRPs*
ASAP	630	52	104	36	93	915	**2,834**
Alphabay	2,275	125	232	134	176	2,942	**5,406**
Cypher	384	120	213	6	103	826	**1,790**
DarkFox	5,171	255	400	101	677	6,604	**13,436**
Kingdom	688	123	181	117	71	1,180	**2,685**
Tor2Door	1,743	197	370	119	280	2,709	**4,679**
Versus	1,683	69	102	139	251	2,244	**5,367**
Vice City	1,110	172	308	44	121	1,755	**3,684**
Total CCRPs	**13,684 (71.3%)**	**1,113 (5.8%)**	**1,910 (9.9%)**	**696 (3.6%)**	**1,772 (9.2%)**	**19,175 (100%)**	**39,881**

Table 5: Total amount of listings per market, total CCRPs, and CCRPs per category.

CCRPs discovered on these platforms, *fraud* listings constitute the overwhelming majority of the listings at 71.3%, while the *services* category comes in last at 3.6% (see Table 5). In the following sections, we go deeper into each of the five CCRP categories, namely *fraud*, *media*, *software*, *malware*, and *services*.

4.3.1 *Fraud.*

This category comprises the majority of our data with 71.3% of the grand total (see Table 5). Regarding the pricing across platforms, these listings range from 1.9 to 267 EUR, and present an average median of 15.1 EUR, which is the highest average among all product categories, declaring fraud products as the most expensive overall. More specifically, the most expensive fraud products were discovered in the *ASAP* marketplace, with the price median at 34.5 EUR (see Figure 3). The main product subcategories under fraud are *carding*, *hacked databases*, and *miscellaneous accounts*, with *carding* being by far the most popular product group, referring to stolen bank cards and bank account credentials, along with some instances of gift cards. As expected, we found that prices in this subcategory depended heavily on the card and account balance, as well as the country of origin. *Hacked database* listings included data on Internet services, voting procedures, and US bank records. Additionally, we notice that the price of database listings tends to be higher for database leaks/hacks that are more recent, rather than depending on the content. This leads to some isolated instances of databases being offered at more than 400 EUR (e.g., we found a private company [2] data leak from 2022, priced at 416 EUR). Lastly, the majority of products in the *miscellaneous accounts* category were harvested account login credentials from online streaming services (e.g., *Netflix*), or popular sites with pornographic content, and the prices depended on the platform and the subscription type of choice. The product distribution over price for the fraud category throughout the various platforms is illustrated in Figure 3.

4.3.2 *Media.*

This category takes up 9.9% of our dataset (see Table 5). The *Versus* and *Kingdom* marketplaces presented the highest price median in this category, at approximately 9.3 EUR per product, while in most cases the prices ranged from 1.9 to 79.4 EUR (see Figure 4). Furthermore, throughout our entire data sample, the median value of these products averaged 5.5 EUR per item. The most popular product subcategories under media are *hacking tutorials, Background and Personal Identifiable Information (PII) check guides*, and *carding methods*. *Hacking* and *cracking tutorials* mostly referred to guides on password cracking, information stealing, and exploitation of vulnerabilities in different systems, but also penetration testing from beginner to advanced level (e.g., Metasploit usage tutorials). *PII* guides are closely related to phishing attacks since they refer to user information-gathering methods. On the contrary to the hacking tutorials discussed above, these guides do not require technical expertise, and focus more on the methodology and strategic aspect of harvesting information on users through social engineering. Lastly, the *carding methods* subcategory is dedicated to specifically teaching users how to commit financial fraud, which among others includes credit card and bank fraud, exploiting the gift card mechanism in place by popular legitimate online stores, and refund scams (e.g., *Amazon*). The media product distribution over price and per platform is illustrated in Figure 4.

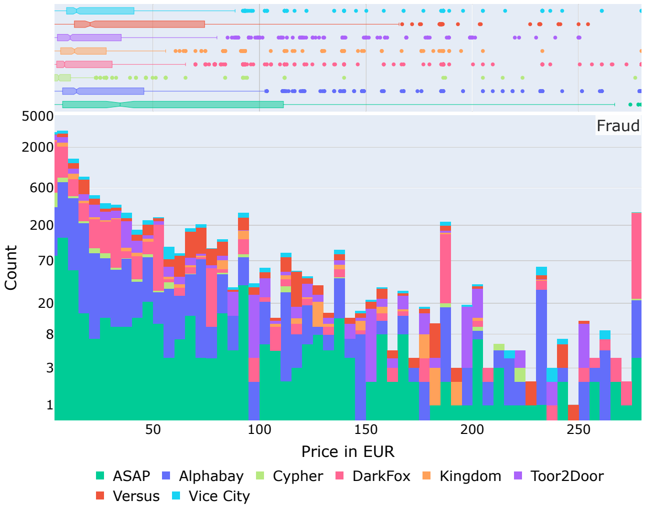

Figure 3: *Fraud* category boxplot (top) and price distribution (bottom).

4.3.3 *Software.*

These listings represent 9.2% of the gathered marketplace data, making it the third biggest category in our sample, after *fraud* and *media* (see Table 5). The prices for software products range from 1.9 EUR to 79.4 EUR across platforms, and the highest price median is that of the *Alphabay* marketplace at 10.7 EUR per product, which was found to be slightly higher than the average of all marketplaces, calculated at 7.1 EUR (see Figure 5). The products in this category can be further divided into two major subcategories: *hacking tools*, such as password crackers, crypto-wallet crackers,

[2] Company anonymized due to ethical reasons.

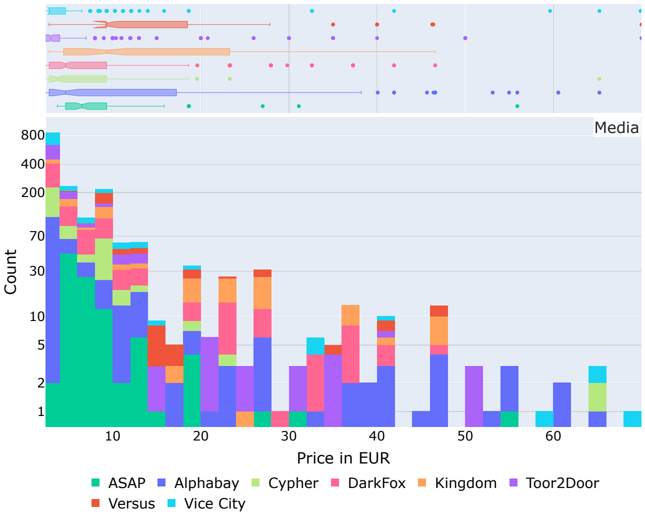

Figure 4: *Media* category boxplot (top) and price distribution (bottom).

phishing tools, and cloning software (e.g., credit card cloners), and *hacked software*, with *Adobe* or *Microsoft* products as two major examples. The distribution of software products depending on their price can be seen in Figure 5.

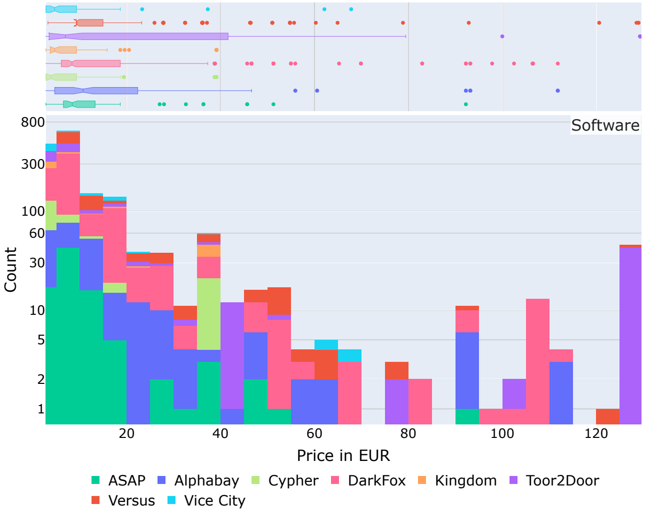

Figure 5: *Software* category boxplot (top) and price distribution (bottom).

4.3.4 *Malware.*

This category corresponds to 5.8% of all listings (see Table 5). Across platforms, malware products were mainly priced in the 1.9 to 93.2 EUR range, with an average median of 7.3 EUR, while the most expensive listings were located on the *Kingdom* marketplace, presenting the highest price median at 14 EUR (see Figure 6). The major types of products falling under this category are *crypto-miners* dedicated to mining the BTC, XMR, and Zcash (ZEC) cryptocurrencies, *Remote Access Trojans (RATs)*, *botnets*, and *ransomware*. In regard to RATs, we discovered products mainly targeting Windows, Linux, and Android systems, with some examples being the *Warzone*, *Orcus*, *BetterBackdoor*, *RedLine*, and *Dendroid* trojans. The botnet listings we documented presented great variety, ranging from older malware distributions like *Zeus* (2007), *Citadel* (2011), and *Mirai* (2016), to more recent ones such as *BlackNET* (2020). The same pattern was noticed with the ransomware listings, with a wide range of products dating back to 2013, such as CryptoLocker, but also more recent years such as *WannaCry* (2017), *Petya* (2017), and *Chaos (Ryuk)* (2021). Lastly, Figure 6 illustrates the distribution of products in the malware category across the 8 marketplaces, based on their pricing.

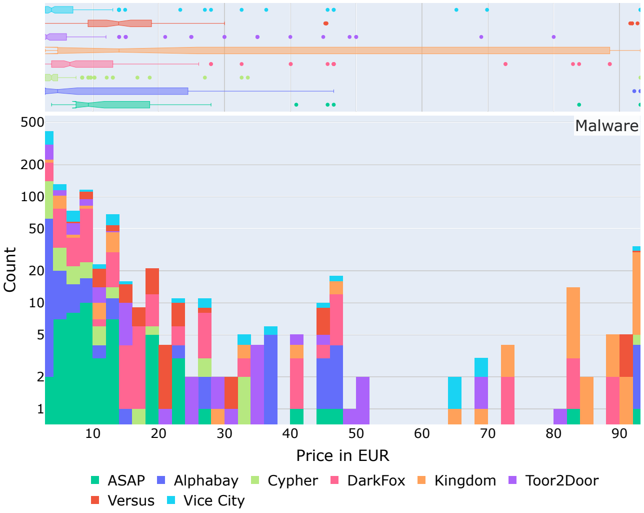

Figure 6: *Malware* category boxplot (top) and price distribution (bottom).

4.3.5 *Services.*

This is the smallest category in our dataset, corresponding to only 3.6% (see Table 5). The prices for services on all of the platforms varied between 1.93 and 93.2 EUR, with the average price median at 8.9 EUR, while the *ASAP* marketplace carried the most expensive products in the category, at a 18.6 EUR median (see Figure 7). The bulk of the listings in this group referred to hiring services such as *spam* and *Distributed Denial of Service (DDoS) attacks*, *review bombing*, *phishing email attacks*, *background checks*, and *hacking*. Additionally, in Figure 7 we present how the products in the services category are distributed on all of the marketplaces, depending on their price.

4.3.6 *Summary.*

Figure 8 presents an overview of the price value details described in the previous section. We illustrate the price range of CCRPs per category across all 8 marketplaces, including the minimum, maximum, and most importantly the median price value (indicated with a vertical line), with a 95% confidence interval. The category found to include the most expensive listings in all of the platforms is the *Fraud* category. The average median for the price of *Fraud* products is 15.1 EUR and was found on the *ASAP* marketplace, a value higher than the median of all category/marketplace combinations, which averages at approximately 8.8 EUR.

This fact suggests that most CCRPs are rather inexpensive, which is also evident in Figures 3 to 7, where in all of the cases the product concentration is higher on the lower end of the price value

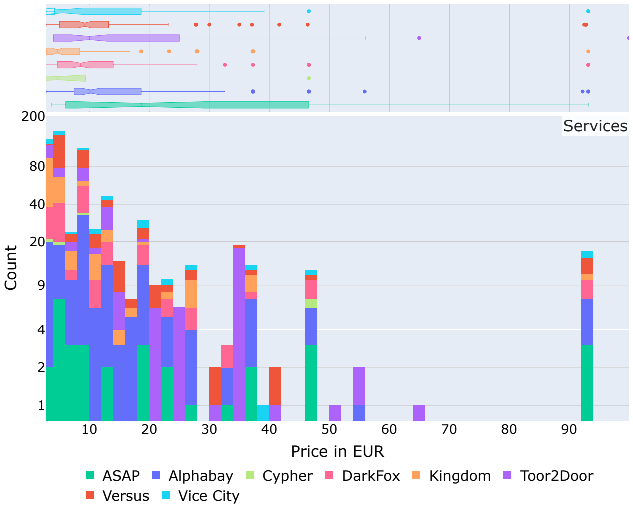

Figure 7: *Services* category boxplot (top) and price distribution (bottom).

Category	Avg Median (EUR)	Max Median (EUR)	
Fraud	15.1	34.47	(ASAP)
Malware	7.35	13.98	(Kingdom)
Media	5.49	9.32	(Kingdom)
Services	8.94	18.64	(ASAP)
Software	7.1	10.71	(AlphaBay)

Table 6: Average and maximum median values of product prices, for each product category across the 8 marketplaces.

axis. Additionally, *Fraud* products also present the highest price median across all categories, at 34.5 EUR, which was found on the *ASAP* marketplace. Table 6 illustrates the average and maximum price medians for each of the product categories, from all of the marketplaces examined.

4.4 Vendors

Since the list of registered vendors cannot be directly accessed in any marketplace, the vendors are derived from the information included in the product listings. Our crawler uses these data points to collect relevant information about vendors, such as trust level and PGP keys. The trust level is an internal metric that determines the degree of satisfaction of the vendor's customers. However, we did not collect the number of vendor sales, since we would have to assume that vendors only sell CCRPs. Moreover, only five of the selected marketplaces showed this information and not in every product. In this section, we go over the data gathered on the vendors corresponding to the CCRP listings that we crawled for the purposes of this work.

In total, we collected information from 1,202 unique vendor profiles. The level of detail of this information, however, varies from one marketplace to another. For example, the trust level in each marketplace may be calculated differently, and some marketplaces prefer not to take disputes into consideration. Therefore, the vendor's reputation has been normalised to use values from zero to one, representing the minimum and maximum values possible. Before moving forward, it is important to mention that the trust level could not be collected from *ASAP* and *Versus* vendors. The *Versus* marketplace closed down before we could collect its vendor profiles, while *ASAP* did not show this information.

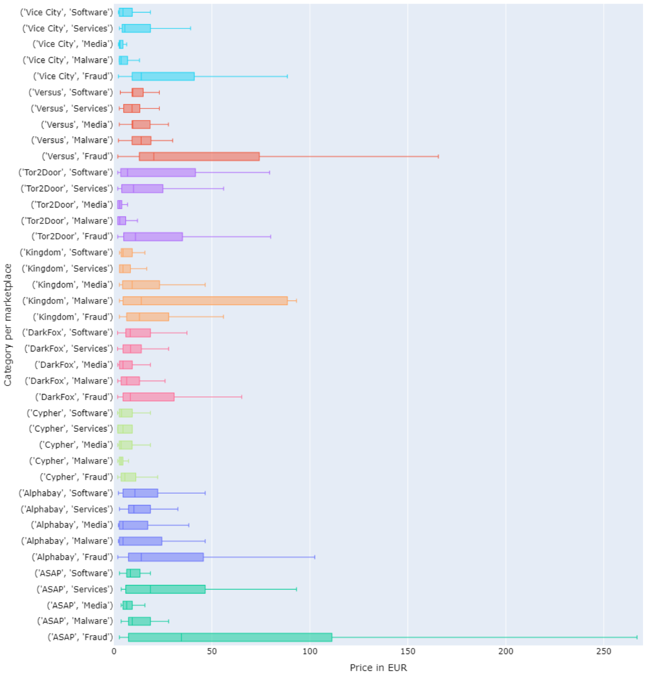

Figure 8: Price range of products per category and marketplace. For each of the marketplace/category combinations, the line inside the box illustrates the corresponding median value of the product price.

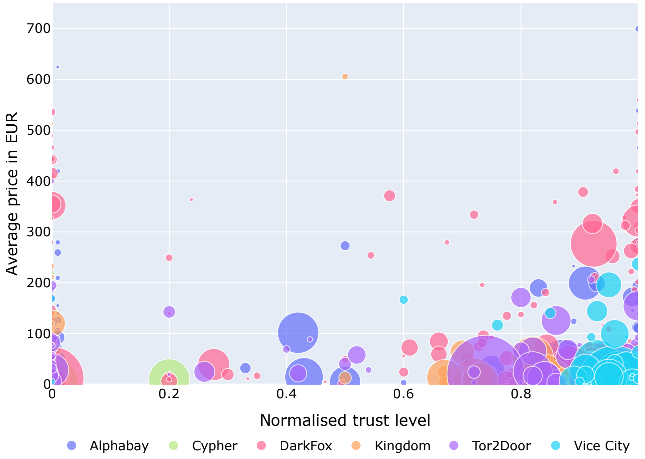

Figure 9: Bubble scatter plot of the vendors' trust levels, the average price of their catalogues and amount of listings offered by the vendors (bubble size) on 6 of the marketplaces.

Regarding the vendors' trust levels, Figure 9 illustrates the average price of the vendors' catalogues, the catalogue size (bubble size), and the normalised trust level per market, showing that the majority of the vendors on six of the marketplaces have a higher than 65% positive feedback from customers. The figure suggests a link between the trust level and the median price of each vendor's catalogue. Further analysis supports this theory, yielding a significant difference in prices for vendors with trust levels under 50%, and those above 95%. Vendors with a lower trust level than 50% average the lowest prices overall, with a median of 20.95 EUR per product and a catalogue of 6 products. Those between 50% and 95% average 21.39 EUR per product with a catalogue of 11 products. Furthermore, those with trust a level higher than 95% have on average a median of 29.21 EUR per product and a catalogue of 3 products, mostly composed of *Fraud* products. Besides *Fraud*, vendors with a trust level between 50% and 95% also sell a significant amount of media products (56.18% of the *Media* category), while those over 95% prefer *Software* products (44.46% of the Software category).

We observed cases in which the vendors' catalogues included listings that despite having different identifiers, they appeared identical in every way (e.g., pricing, title, description). Since these listings were registered separately on the sites and treated as unique listings, we chose to follow the same approach and include all of them when calculating the vendors' catalogue size. Furthermore, most vendors will specialise in one category (e.g., Fraud), though their catalogue may contain products from other categories (e.g., Malware).

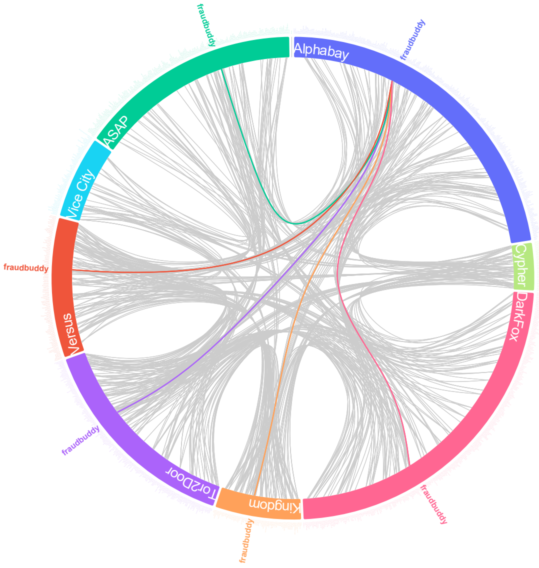

Figure 10: Hierarchical edge bundling of vendors appearing in multiple marketplaces with the same username.

According to our analysis, it is common for vendors to offer their catalogues in multiple marketplaces simultaneously. This relationship is shown in Figure 10, where we cross-referenced the vendors' usernames to visualise whether the same vendors may appear in the rest of the markets [3]. The data shows that 78% of the vendors' usernames appear on multiple marketplaces. After analysing the differences between the catalogues in multiple marketplaces, the most notable difference is the pricing, i.e., vendors will use different prices to advertise their products in different marketplaces. Nonetheless, product descriptions will remain the same, or closely similar across marketplaces. One reason for the difference in prices is that marketplaces have different fees per transaction. For instance, some are higher due to requiring additional services (e.g., escrow) or lower as a result of the lack of them (e.g., *Finalize Early*).

However, vendor usernames appearing in multiple marketplaces do not necessarily belong to the same identity. Out of the 782 PGP keys we discovered, 12.98% were found in more than one marketplace, while four keys were found in six different marketplaces. Figure 11 clusters the vendor accounts (coloured bubbles) using the same PGP key (outer circle) across marketplaces (inner circle). The figure presents the size of the coloured bubbles to determine the average price of the vendor catalogue. In the figure, it can be seen that vendors with a higher average catalogue price tend to appear in a single marketplace. In addition, the dataset contains many instances of unique pairs of PGP keys and usernames, contrary to the relationships seen in Figure 10. Vendors may simply decide to use different PGP keys for each marketplace to reduce traceability. Instead, markets include vendor verification programs that allow vendors to claim their accounts in multiple markets [18]. In line with this finding, we observed multiple instances of invalid public PGP keys. For example, the dataset contains multiple private PGP keys, empty fields, or random text.

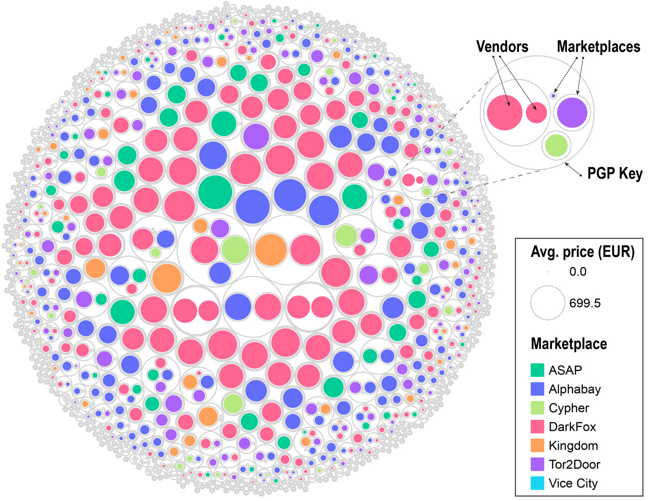

Figure 11: Circle packing of vendors' PGP keys used across 7 of the marketplaces.

5 CONCLUSION

In this paper, we harvest data from 8 darkweb marketplaces, specifically targeting 5 product categories that are bound to include CCRPs, in an effort to acquire an understanding of this particular type of trade, from product categories, to pricing and vendor availability. Additionally, we explore the option of crawler deployment on the I2P network, with the purpose of scraping platforms that offer availability on both Tor and I2P.

We find that the majority of these listings fall under the *fraud* category, which dominates the market at 71.3% of our entire dataset,

[3] The curious reader can interact with the plot here: https://observablehq.com/@anonymous-author/ccrp-hierarchical-edge-bundling

while also presenting the highest average price median at 15.1 EUR, making these products the most expensive. The overall median of CCRPs was calculated at 8.7 EUR, which is quite inexpensive. We also investigate the vendor landscape of the CCRP market, by acquiring information on 1,202 vendors, with special interest on usernames, reputation, and PGP keys. Our findings suggest that in the 6 marketplaces from which we were able to acquire reputation data, the majority of these vendors present a higher than 65% positive feedback score. Lastly, in terms of vendor distribution across the marketplaces, 78% of the harvested vendor usernames appear on multiple platforms, while out of the 782 PGP keys we harvested from 7 of the marketplaces, 12.98% were used on more than one marketplace. Our future work will focus on investigating how botnet-specific products are being distributed in the darkweb.

REFERENCES

[1] Bassel Alkhatib and Randa Basheer. 2019. Crawling the Dark Web: A Conceptual Perspective, Challenges and Implementation. *Journal of Digital Information Management* 17 (4 2019), 51. https://doi.org/10.6025/jdim/2019/17/2/51-60

[2] Luca Allodi. 2017. Economic Factors of Vulnerability Trade and Exploitation. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, Texas, USA) *(CCS '17)*. ACM, New York, NY, USA, 1483–1499. https://doi.org/10.1145/3133956.3133960

[3] David Audran., Marcus Andersen., Mark Hansen., Mikkel Andersen., Thomas Frederiksen., Kasper Hansen., Dimitrios Georgoulias., and Emmanouil Vasilomanolakis. 2022. Tick Tock Break the Clock: Breaking CAPTCHAs on the Darkweb. In *Proceedings of the 19th International Conference on Security and Cryptography - SECRYPT,*. INSTICC, SciTePress, Lisbon, Portugal, 357–365. https://doi.org/10.5220/0011273300003283

[4] Matthew Ball and Roderic Broadhurst. 2021. Data Capture and Analysis of Darknet Markets. *Available at SSRN* 1, 3344936 (Feb 2021), 25 pages. https://doi.org/10.2139/ssrn.3344936

[5] Victor Benjamin, Weifeng Li, Thomas Holt, and Hsinchun Chen. 2015. Exploring threats and vulnerabilities in hacker web: Forums, IRC and carding shops. In *2015 IEEE International Conference on Intelligence and Security Informatics (ISI)*. IEEE, Baltimore, MD, USA, 85–90. https://doi.org/10.1109/ISI.2015.7165944

[6] Tim M. Booij, Thijmen Verburgh, Federico Falconieri, and Rolf S. van Wegberg. 2021. Get Rich or Keep Tryin' Trajectories in dark net market vendor careers. In *European Symposium on Security and Privacy Workshops (EuroS PW)*. IEEE, Vienna, Austria, 202–212. https://doi.org/10.1109/EuroSPW54576.2021.00028

[7] Alberto Bracci, Matthieu Nadini, Maxwell Aliapoulios, Damon McCoy, Ian Gray, Alexander Teytelboym, Angela Gallo, and Andrea Baronchelli. 2021. Dark Web Marketplaces and COVID-19: before the vaccine. *EPJ Data Science* 10, 11 (12 2021), 6. https://doi.org/10.1140/epjds/s13688-021-00259-w

[8] Michele Campobasso, Pavlo Burda, and Luca Allodi. 2019. CARONTE: Crawling Adversarial Resources Over Non-Trusted, High-Profile Environments. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS PW)*. IEEE, Stockholm, Sweden, 433–442. https://doi.org/10.1109/EuroSPW.2019.00055

[9] Tao Chen, Ruifeng Xu, Yulan He, and Xuan Wang. 2017. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. *Expert Systems with Applications* 72 (2017), 221–230.

[10] Alejandro Cuevas, Fieke Miedema, Kyle Soska, Nicolas Christin, and Rolf van Wegberg. 2022. Measurement by Proxy: On the Accuracy of Online Marketplace Measurements. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, Boston, MA, 2153–2170.

[11] Ashok Deb, Kristina Lerman, and Emilio Ferrara. 2018. Predicting Cyber Events by Leveraging Hacker Sentiment. *Information* 9, 11 (11 2018), 280. https://doi.org/10.3390/info9110280 arXiv: 1804.05276.

[12] Jakob Demant, Rasmus Munksgaard, and Esben Houborg. 2018. Personal use, social supply or redistribution? Cryptomarket demand on Silk Road 2 and Agora. *Trends in Organized Crime* 21, 1 (2018), 42–61.

[13] Roger Dingledine, Nick Mathewson, and Paul Syverson. 2004. *Tor: The second-generation onion router.* Technical Report. Naval Research Lab Washington DC.

[14] Mohammadreza Ebrahimi, Jay F. Nunamaker Jr., and Hsinchun Chen. 2020. Semi-supervised cyber threat identification in dark net markets: A transductive and deep learning approach. *Journal of Management Information Systems* 37, 3 (2020), 694–722. https://doi.org/10.1080/07421222.2020.1790186

[15] Mohammadreza Ebrahimi, Mihai Surdeanu, Sagar Samtani, and Hsinchun Chen. 2018. Detecting cyber threats in non-english dark net markets: A cross-lingual transfer learning approach. In *2018 IEEE International Conference on Intelligence and Security Informatics (ISI)*. IEEE, Miami, FL, USA, 85–90. https://doi.org/10.1109/ISI.2018.8587404

[16] Yara Fareed Fahmy Bayoumy, Per Håkon Meland, and Guttorm Sindre. 2018. A Netnographic Study on the Dark Net Ecosystem for Ransomware. In *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. IEEE, Glasgow, UK, 1–8. https://doi.org/10.1109/CyberSA.2018.8551424

[17] European Monitoring Centre for Drugs and Drug Addiction. 2018. *Darknet markets ecosystem.* EUROPOL. https://www.emcdda.europa.eu/system/files/publications/8347/Darknet2018_posterFINAL.pdf

[18] Dimitrios Georgoulias, Jens Myrup Pedersen, Morten Falch, and Emmanouil Vasilomanolakis. 2021. A qualitative mapping of Darkweb marketplaces. In *2021 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE, Boston, MA, USA, 1–15. https://doi.org/10.1109/eCrime54498.2021.9738766

[19] Dimitrios Georgoulias, Jens Myrup Pedersen, Morten Falch, and Emmanouil Vasilomanolakis. 2023. Botnet Business Models, Takedown Attempts, and the Darkweb Market: A Survey. *ACM Comput. Surv.* 55, 11, Article 219 (feb 2023), 39 pages. https://doi.org/10.1145/3575808

[20] Dimitrios Georgoulias, Jens Myrup Pedersen, Morten Falch, and Emmanouil Vasilomanolakis. 2023. COVID-19 Vaccination Certificates in the Darkweb. *Digital Threats* 4, 1, Article 7 (mar 2023), 17 pages. https://doi.org/10.1145/3530877

[21] HF Canonical Model Maintainers. 2022. distilbert-base-uncased-finetuned-sst-2-english. https://doi.org/10.57967/hf/0181

[22] Víctor Labrador and Sergio Pastrana. 2022. Examining the trends and operations of modern Dark-Web marketplaces. In *2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, Genoa, Italy, 163–172.

[23] Mitch Macdonald, Richard Frank, Joseph Mei, and Bryan Monk. 2015. Identifying Digital Threats in a Hacker Web Forum. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015 (ASONAM '15)*. ACM, New York, NY, USA, 926–933. https://doi.org/10.1145/2808797.2808878

[24] Eric Nunes, Ahmad Diab, Andrew Gunn, Ericsson Marin, Vineet Mishra, Vivin Paliath, John Robertson, Jana Shakarian, Amanda Thart, and Paulo Shakarian. 2016. Darknet and deepnet mining for proactive cybersecurity threat intelligence. In *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*. IEEE, Tucson, AZ, USA, 7–12. https://doi.org/10.1109/ISI.2016.7745435

[25] George Pantelis, Petros Petrou, Sophia Karagiorgou, and Dimitrios Alexandrou. 2021. On Strengthening SMEs and MEs Threat Intelligence and Awareness by Identifying Data Breaches, Stolen Credentials and Illegal Activities on the Dark Web. In *The 16th International Conference on Availability, Reliability and Security (ARES 2021)*. ACM, New York, NY, USA, 1–7. https://doi.org/10.1145/3465481.3469201

[26] Sergio Pastrana, Daniel R. Thomas, Alice Hutchings, and Richard Clayton. 2018. CrimeBB: Enabling Cybercrime Research on Underground Forums at Scale. In *Proceedings of the 2018 World Wide Web Conference* (Lyon, France) *(WWW '18)*. ACM, Republic and Canton of Geneva, CHE, 1845–1854. https://doi.org/10.1145/3178876.3186178

[27] Giacomo Persi Paoli, Judith Aldridge, Nathan Ryan, and Richard Warnes. 2017. *Behind the curtain: The illicit trade of firearms, explosives and ammunition on the dark web.* RAND Corporation, Santa Monica, CA. https://doi.org/10.7249/RR2091 Citation Key: RR-2091-PACCS.

[28] Kyle Soska and Nicolas Christin. 2015. Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem. In *Proceedings of the 24th USENIX Conference on Security Symposium* (Washington, D.C.) *(SEC'15)*. USENIX Association, USA, 33–48.

[29] Rolf Van Wegberg, Samaneh Tajalizadehkhoob, Kyle Soska, Ugur Akyazi, Carlos Gañán, Bram Klievink, Nicolas Christin, and Michel Van Eeten. 2018. Plug and Prey? Measuring the Commoditization of Cybercrime via Online Anonymous Markets. In *Proceedings of the 27th USENIX Conference on Security Symposium* (Baltimore, MD, USA) *(SEC'18)*. USENIX Association, USA, 1009–1026.

[30] Yichao Wang, Budi Arief, and Julio Hernandez-Castro. 2021. Toad in the Hole or Mapo Tofu? Comparative Analysis of English and Chinese Darknet Markets. In *2021 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE, Boston, MA, USA, 1–13. https://doi.org/10.1109/eCrime54498.2021.9738745

[31] York Yannikos, Julian Heeger, and Martin Steinebach. 2022. Data Acquisition on a Large Darknet Marketplace. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (Vienna, Austria) *(ARES '22)*. ACM, New York, NY, USA, Article 53, 6 pages. https://doi.org/10.1145/3538969.3544472

[32] York Yannikos, Julian Heeger, and Martin Steinebach. 2022. Data Acquisition on a Large Darknet Marketplace. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (Vienna, Austria) *(ARES '22)*. ACM, New York, NY, USA, Article 53, 6 pages. https://doi.org/10.1145/3538969.3544472

[33] Gengqian Zhou, Jianwei Zhuge, Yunqian Fan, Kun Du, and Shuqiang Lu. 2020. A Market in Dream: the Rapid Development of Anonymous Cybercrime. *Mobile Networks and Applications* 25, 1 (2 2020), 259–270. https://doi.org/10.1007/s11036-019-01440-2