Quantitative study on the Database marketplace: What does it sell? How much are they?

Cuong Nguyen

25/11/2022

Contents

Glossary

BIN Bank Identity Number. 19

CS Credit Score. 14, 19, 23

CVV Card Verification Value. 19

DHT Distributed Hash Table. 6

DL Driver's License. 14, 19, 23

DNS Domain Name System. 6

DOB Date of Birth. 9, 14, 19, 23

EU European Union. 15

MITM Man-in-the-middle. 6

MMN Mother Maiden Name. 14, 19

SSH Secure Shell. 4

SSN Social Security Number. 9, 14, 19, 20, 23

US United States. 15

1 Abstract

In this report, I aim to study the detail numerical statistic of Database marketplace, which are price range, mean, median price of each type of products and all items in the whole dataset. Given a JSON dataset, scrapping and crawling the whole marketplace website by Juha Nurmi and others, I developed and extracted more key fields which are very informative in grouping product types and generating statistical results. Based on the analytical results, I observe that the average price of products traded in Database marketplace is not very high, affordable for most of buyers. In addition, personal information and banking/credit data are widely sold on the dark marketplace. Those sensitive data is highly possible leaked from major data breaches. Without raising the awareness of securing personal data, more people will suffer increasingly financial and mental damage.

2 Background

In this section, I will describe theoretical concepts: Tor network, Onion service, Cryptocurrency, Dark marketplace.

2.1 Tor network

Tor is the second generation of onion routing system that addresses shortcomings in the original design by adding perfect forward secrecy, congestion control, directory servers, integrity checking, configurable exit policies, and a practical design for location-hidden services through rendzvous points [3]. Tor is a low-latency network system, which means that the period of delaying is negligible for most users [5]. This advantage makes Tor as a suitable design for interactive tasks like web browsing or SSH connections [12]. In addtion to Tor's fast performance, Tor communication service also provides the high degree of anonymity and privacy in transferring network packets. Hence, Tor browser, a Firefox-based web browser under Tor network, is widely used by activists and journalists for untraceable communication [12]. Moreover, ordinary citizens who have concerned their privacy during Internet surfing select Tor-based applications as safe guards in the ungovernable Internet environment.

The Tor users are partly protected against the common method of Internet surveillance, called traffic analysis. The source and destination of Internet communication can be exposed by traffic analysis. This information reveals users' behaviour and interests [15]. To reduce the risk of traffic analysis, Tor distribute the connection between the client and the final point over the chain of 3 Tor nodes, called *relays*. The circuit of encrypted connection over relays is described by the following steps.

1. The Tor user obtains a list of Tor nodes from a directory server. Then public keys of selected relays are used to encrypt and encapsulate the network packet.

2. The Tor client connects to the guard relay, the first relay in the circuit of three Tor nodes. The encapsulated message is delivered to the guard relay.

3. The guard relay decrypts the packet, using its private key, and forward to the second node, the middle relay.

4. The second node continues decrypting, using its private key, and then forwarding the message to the third node, the exit relay.

5. The exit relay decrypts, using its private key, and forward the packet to the final destination, e.g, a web server.

As shown in the aforementioned steps, a specific Tor relay only knows about the previous sender and the next recipient. For example, the guard relay only learns about the Tor client and the second hop, but not the final recipient.

Thus, a complete path from the original sender to the final destination is not exposed.

At each hop, the packet is decapsulated and decrypted by using private key of each relay, "peeling off" each encryption layer. It is comparable to layers of onion and that is also the reason why it is called *onion routing*.

Although the traffic correlation attack is unavoidable in any low-latency anonymity network, it is challenging to perform global traffic analysis due to the huge amount of Tor users [17, 3]. In addition to traffic and time analysis, there are other techniques that reveal real identity of Tor users [13].

There are several types of routers, or relays, in the Tor network and soem of them are aforementioned, including guard, middle, and exit. Since the exit relay communicates publicly, outside the Tor network, to the destination, the IP address of it is exposed. Moreover, the sender can also know the IP address of the first relay in the circuit of three Tor routers. By blocklisting the IP addresses of theses publicly known relays, many governments block connections from their citizens to the Tor network. To address this problem, the bridge node which is not listed in the public Tor directory is introduced. Thus, it is difficult to block an IP address if it is not publicly known. However, several countries have found means to detect and block connections to Tor bridges. Pluggable transport, a special kind of bridge, addresses this problem by adding an extra layer of obfuscation [16]. According to `https://metrics.torproject.org` in 2021, there are around 7000 relays and more than 2000 bridges running in the Tor network.

Tor enables TCP-based applications to acquire online anonymity without modification [3]. One of popular Tor application is Tor browser. It is a modified version of Mozilla Firefox Extended Support Release (ESR) with the most-strict security settings, such as NoScript and HTTPS Everywhere. This browser routes all traffic through the Tor network and removes all possible fingerprinting methods, including forging information about the operating system and hardwares [12]. Furthermore, Tor Browser does not save sensitive data, including browsing history, cache and cookies [12].

2.2 Onion service

Onion services, Tor hidden services, are an overlay network on top of TCP/IP; thus, IP addresses are not used in the protocol [14]. The location and real IP address of Onion Service are protected, hidden from the user. These hidden services are only available through the Tor network [3]. Instead of IP address, onion service is indentified by the *identity public key*. This onion address which is in the form of *x.onion* whose the first part x is the hash of public key of onion service, e.g, **database6e2t4yvdsrbw3qq6votzyfzspaso7sjga2tchx6tov23nsid.onion** [12].

The onion service hides and protects itself by only allowing direct connections from three pre-selected Tor relays, *introduction points*. Every Tor client and intermediary points have to be introduced by the three-hop circuit to reach the onion service. To be noticeable to clients, instead of using Domain Name System (DNS), Distributed Hash Table (DHT) is utilized to match a onion service to its corresponding introduction points.

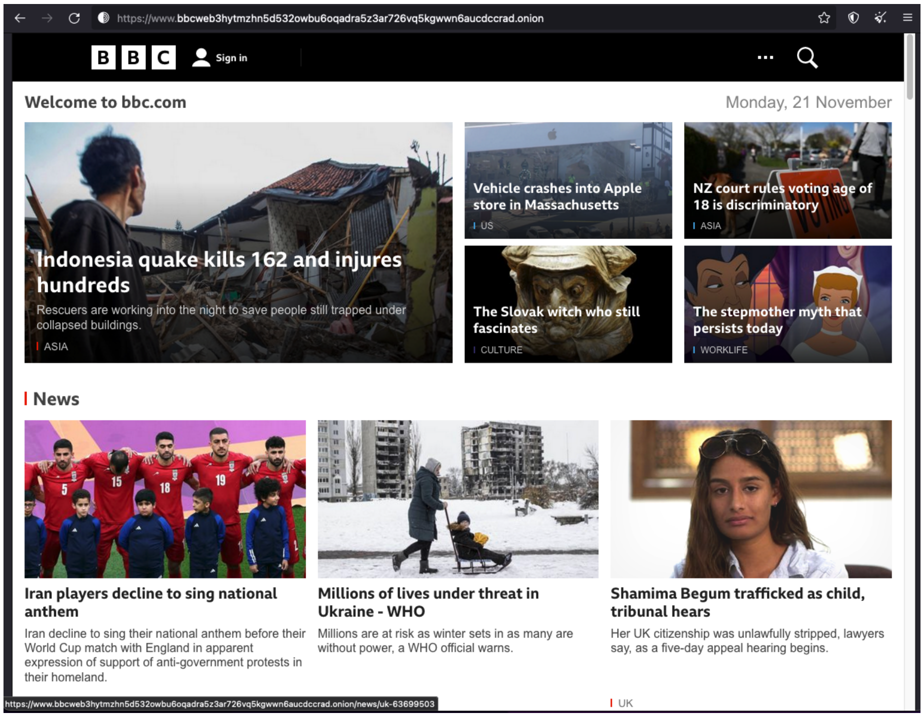
The onion service protocol is summarized in the following steps [3].

1. The onion service selects three Tor relays acting as its introduction points to build an anonymized circuit.
2. The onion service generates a *descriptor*, containing a list of introduction points, and signs this descriptor with the *identity private key*. The key used to signing is the private part of the identity public key which is encoded in the onion service address.
3. Given the onion address, the client requests the signed descriptor from from the DHT. Next, the client verify the signature of returned descriptor with the public key that is encoded in the onion address.
4. The client randomly selects a Tor relay acting as a rendezvous point, build anonymized, three-hop circuit to it and send it an one-time secret.
5. The generated one-time secret and the address of the rendezvous point are encrypted with the public key of the onion service and delivered to one of the introduction points through a anonymized circuit.
6. The onion service decrypts and retrieve the rendezvous point and one-time secret. It connects and sends the one-time secret to the rendezvous point via a three-hop circuit.
7. A complete circuit between the client and the onion service is built. It contains total five Tor relays, two from the client to the rendezvous point and three from the rendezvous point to the onion service.

The onion service protocol provides both end-to-end encryption and authentication [14]. Hence, it is impossible to censor onion services and Man-in-the-middle (MITM) attack is prevented [12]. As onion service can publish contents anonymously, some of them are the sources of adverse content, including illegal traded products, images of child abusing, etc [12]. However, many onion services share content supporting human rights, meaningful for journalism and publishing content that is censored by oppressive governments [12].

2.3 Cryptocurrency

Onion service provides location hiding capacity and anonymity in content publishing for illegal markeplaces. However, they also need a payment system that is untraceable and secure, which is impossible to know exactly involving parties. Hence, the creation of pseudonymous unregulated electronic money leads to the increasing number of illegal marketplaces is formed [12].


https://www.bbcweb3hytmzhn5d532owbu6oqadra5z3ar726vq5kgwwn6aucdccrad.onion
BBC
Sign in

Welcome to bbc.com

Monday, 21 November

Indonesia quake kills 162 and injures hundreds

Rescuers are working into the night to save people still trapped under collapsed buildings.

Vehicle crashes into Apple store in Massachusetts

NZ court rules voting age of 18 is discriminatory

The Slovak witch who still fascinates

The stepmother myth that persists today

News

Iran players decline to sing national anthem

Iran decline to sing their national anthem before their World Cup match with England in apparent expression of support of anti-government protests in their homeland.

Millions of lives under threat in Ukraine - WHO

Millions are at risk as winter sets in as many are without power, a WHO official warns.

Shamima Begum trafficked as child, tribunal hears

Her UK citizenship was unlawfully stripped, lawyers say, as a five-day appeal hearing begins.

https://www.bbcweb3hytmzhn5d532owbu6oqadra5z3ar726vq5kgwwn6aucdccrad.onion/news/uk-63699503
UK

Figure 1: Onion service mirroring BBC news website on Tor network.

In 2009, someone who has pseudo name Satoshi Nakamoto created the peer-to-peer digital cash system [10]. This electronic money, widely known as Bitcoin, is the first distributed digital currency that works without a trusted third party [12, 10]. Bitcoin operates based on cryptographic proof instead of trust, trading directly between users [10]. Several security mechanism, including hash checksums and proof-of-work, are implemented to protect users from transaction spoofing [12].

The amount of bitcoins is limited, 21 million bitcoins; thus, its price depends on the continuos demand to exchange this currency among its netowrk or to other currencies [10, 12, 6].

Although Bitcoin system does not provide high degree of technical anonymity, it is widely used as the main method of payment in illegal commerce website hosted in the Tor network [12]. In addition to Bitcoin, there are several cryptocurrencies, including Monero, Zcash, Litecoin, etc., that are utilized for trading in Tor-based marketplaces.

2.4 Dark marketplace

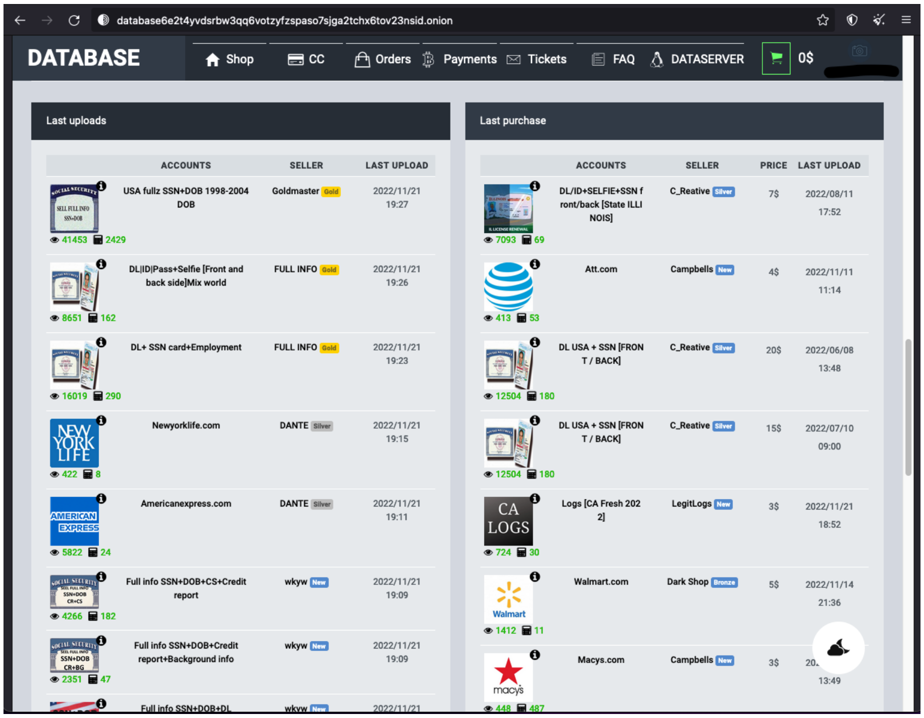
Silk Road, a dark marketplace, is an ideal case to study the impact of anonymous online communication on the transformation of crime, from offline to online environment [1]. With the support of identity-hidden services, illegal trade is significantly expanded. There are two essential components of a dark marketplace [12]:

1. A network that provides high degree of anonymity whose the market website is hosted.
2. A payment system that is not only secure but also exhaustively to trace involving parties.

Silk Road was the first combination of these components [12]. In this online market, users — buyers and sellers — have their own Bitcoin wallers [18]. An *escrow system* that charges a commission fee to the market site and locks the payment between a buyer and a store [18]. In addition, a reputation system is introduced to motivate the vendors sell products matching their advertisements. When a buyer receive a product, if the buyer is happy with the product, the payment is transferred to the vendor. Otherwise, if the buyer does not satisfy, the buyer can send a complaint ticket and the market site resolves the conflict between the buyer and the seller [18]. The buyer can give feedbacks that is publicly visible for other buyers, making the reputation system is important for both buyers and sellers [7].

Several academic papers are published researching different sides of Silk Road [1, 18, 17]. It is studied that Silk Road was a profitable marketplace which trades millions of USD values of illegal drugs per month [4].

These kind of black marketplaces operate based on the financial motivation and



databaseбe2t4yvdsrbw3qq6votzyfzspaso7sjga2tchx6tov23nsid.onion
DATABASE
Shop
CC
Orders
Payments
Tickets
FAQ
DATASERVER
0$
[redacted]

Last uploads

	ACCOUNTS	SELLER	LAST UPLOAD
👁 41453 🖩 2429	USA fullz SSN+DOB 1998-2004 DOB	Goldmaster Gold	2022/11/21 19:27
👁 8651 🖩 162	DL|ID|Pass+Selfie [Front and back side]Mix world	FULL INFO Gold	2022/11/21 19:26
👁 16019 🖩 290	DL+ SSN card+Employment	FULL INFO Gold	2022/11/21 19:23
👁 422 🖩 8	Newyorklife.com	DANTE Silver	2022/11/21 19:15
👁 5822 🖩 24	Americanexpress.com	DANTE Silver	2022/11/21 19:11
👁 4266 🖩 182	Full info SSN+DOB+CS+Credit report	wkyw New	2022/11/21 19:09
👁 2351 🖩 47	Full info SSN+DOB+Credit report+Background info	wkyw New	2022/11/21 19:09
	Full info SSN+DOB+DL	wkyw New	2022/11/21

Last purchase

	ACCOUNTS	SELLER	PRICE	LAST UPLOAD
👁 7093 🖩 69	DL/ID+SELFIE+SSN front/back [State ILLINOIS]	C_Reative Silver	7$	2022/08/11 17:52
👁 413 🖩 53	Att.com	Campbells New	4$	2022/11/11 11:14
👁 12504 🖩 180	DL USA + SSN [FRONT / BACK]	C_Reative Silver	20$	2022/06/08 13:48
👁 12504 🖩 180	DL USA + SSN [FRONT / BACK]	C_Reative Silver	15$	2022/07/10 09:00
👁 724 🖩 30	Logs [CA Fresh 2022]	LegitLogs New	3$	2022/11/21 18:52
👁 1412 🖩 11	Walmart.com	Dark Shop Bronze	5$	2022/11/14 21:36
👁 448 🖩 487	Macys.com	Campbells New	3$	20 13:49

Figure 2: Homepage of Database marketplace onion service hosted on Tor network.

the increasing demand of customers. In addition to Silk Road, was shut down in 2013, there are around 20 different dark marketplaces which are observed huge traffic [12]. These marketplace can be considered as an illegal version of Ebay with additional cryptocurrency payment systems, escrow systems, and reputation systems [12].

In this study, Database marketplace, also a black marketplace, is examined. Unlike Silk Road or its variants, whose main products are illegal drugs, Database market mainly provides personal data — full name, DOB, SSN, etc. — and banking information serving for illicit activities.

3 Research questions

Database marketplace sells access to online compromised accounts and other credential information which can be used for deleterious purposes. For example, buyers can use sold credit cards to purchase their debts or online items. Another scenario is that a criminal can utilize personal information to borrow money from banks, or to make fraudulent health insurance claims[11]. Hence, personal data is an attractive item on dark webs, and the breach of it has the significant impact on society in the financial perspective. In this study, I will explore the wide range of sensitive data sold on Database marketplace, revealing a detail picture of stolen personal data sold on the hidden Tor network through the following research questions:

1. What types of stolen information are traded?
2. What are the prices for different product types?
3. Where the stolen information is from?
4. What we can learn about the sellers?

4 Methodology

4.1 Research design

In order to answer research questions, I explored more key fields from the initial database. These fields includes price, date, product type, and shop information. By applying appropriate statistic on key fields, I produced conclusions based on numerical results. In addtion, manually discovering the Database market, I captured a certain of screenshots that support my observation.

In order to replicate my result, please do the following steps:

1. Access the original dataset provided by Juha Nurmi at `https://mega.nz/folder/aJwVyIYJ#9SWh-Z3-TpPfjHZeFxbeew`.
2. Download *database.tar.gz* (7 MB) and follow the instructions in *README.md*.
3. Check *sha256sum*: **5a5f2cb4feb7fee597b0a26b1dc2fb33b1f9cae639e995a89663198bcfa76f1a**.
4. Extract the compressed dataset (tar -xf database.tar.gz).
5. 33,896 JSON files are generated in the destination folder.
6. Clone the GitHub repo `https://github.com/ancuongnguyen07/Database_Market.git` and follow the comprehensive *README.md* document.

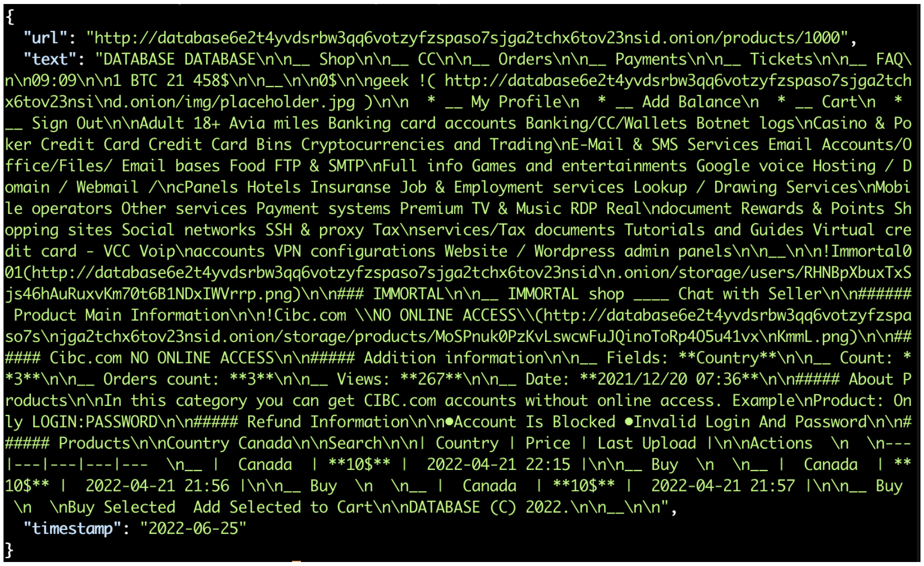
4.2 Samples

The Database Market that sells personal data is accessible through `http://database6e2t4yvdsrbw3qq6votzyfzspaso7sjga2tchx6tov23nsid.onion/` inside the anonymous Tor network. The initial dataset, provided by teacher Juha Nurmi, can be downloaded from `https://mega.nz/folder/aJwVyIYJ#9SWh-Z3-TpPfjHZeFxbeew`. The provided dataset is an archive of 33,896 JSON files, 400 MB of raw size. Each JSON file is represented as the following fields as shown in Figure 3:

- *url*: URL address of webpage.
- *text*: Text of the webpage.
- *timestamp*: Data collection date.

In order to achieve more insights of credentials for sale, I produced an additional JSON file ,*ProductPages*,that reveals more key features of products sold on a targeted marketplace. Each entry in the customized JSON file represents a webpage of product containing many sorts of items for sale. Figure 4 is an example of an entry, where I assigned the following fields.

- *id*: Product ID.
- *time-stamp*: Data collection date.



```
{
  "url": "http://database6e2t4yvdsrbw3qq6votzyfzspaso7sjga2tchx6tov23nsid.onion/products/1000",
  "text": "DATABASE DATABASE\n\n__ Shop\n\n__ CC\n\n__ Orders\n\n__ Payments\n\n__ Tickets\n\n__ FAQ\
n\n09:09\n\n1 BTC 21 458$\n\n__\n\n0$\n\ngeek !( http://database6e2t4yvdsrbw3qq6votzyfzspaso7sjga2tch
x6tov23nsi\nd.onion/img/placeholder.jpg )\n\n  * __ My Profile\n  * __ Add Balance\n  * __ Cart\n  *
__ Sign Out\n\nAdult 18+ Avia miles Banking card accounts Banking/CC/Wallets Botnet logs\nCasino & Po
ker Credit Card Credit Card Bins Cryptocurrencies and Trading\nE-Mail & SMS Services Email Accounts/O
ffice/Files/ Email bases Food FTP & SMTP\nFull info Games and entertainments Google voice Hosting / D
omain / Webmail /\ncPanels Hotels Insuranse Job & Employment services Lookup / Drawing Services\nMobi
le operators Other services Payment systems Premium TV & Music RDP Real\ndocument Rewards & Points Sh
opping sites Social networks SSH & proxy Tax\nservices/Tax documents Tutorials and Guides Virtual cre
dit card - VCC Voip\naccounts VPN configurations Website / Wordpress admin panels\n\n__\n\n!Immortal0
01(http://database6e2t4yvdsrbw3qq6votzyfzspaso7sjga2tchx6tov23nsid\n.onion/storage/users/RHNBpXbuxTxS
js46hAuRuxvKm70t6B1NDxIWVrrp.png)\n\n### IMMORTAL\n\n__ IMMORTAL shop ____ Chat with Seller\n\n######
 Product Main Information\n\n!Cibc.com \\NO ONLINE ACCESS\\(http://database6e2t4yvdsrbw3qq6votzyfzspa
so7s\njga2tchx6tov23nsid.onion/storage/products/MoSPnuk0PzKvLswcwFuJQinoToRp405u41vx\nKmmL.png)\n\n##
#### Cibc.com NO ONLINE ACCESS\n\n##### Addition information\n\n__ Fields: **Country**\n\n__ Count: *
*3**\n\n__ Orders count: **3**\n\n__ Views: **267**\n\n__ Date: **2021/12/20 07:36**\n\n##### About P
roducts\n\nIn this category you can get CIBC.com accounts without online access. Example\nProduct: On
ly LOGIN:PASSWORD\n\n##### Refund Information\n\n●Account Is Blocked ●Invalid Login And Password\n\n#
##### Products\n\nCountry Canada\n\nSearch\n\n| Country | Price | Last Upload |\n\nActions  \n  \n---
|---|---|---|---  \n__ |  Canada  | **10$** |  2022-04-21 22:15 |\n\n__ Buy  \n  \n__ |  Canada  | **
10$** |  2022-04-21 21:56 |\n\n__ Buy  \n  \n__ |  Canada  | **10$** |  2022-04-21 21:57 |\n\n__ Buy
 \n  \nBuy Selected  Add Selected to Cart\n\nDATABASE (C) 2022.\n\n__\n\n",
  "timestamp": "2022-06-25"
}
```

Figure 3: An example JSON file from the provided dataset by Juha.

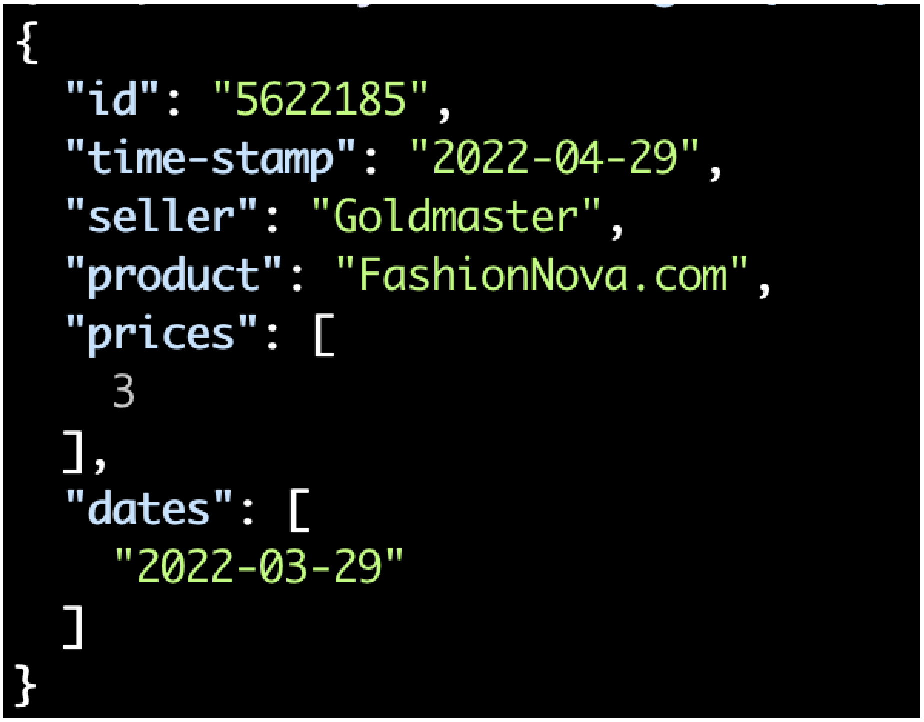
- *category*: Type of product.
- *seller*: Username of seller.
- *product*: Name of product.
- *prices*: Prices of items.
- *dates*: Item uploaded date.

The datasets are under CC BY 4.0 license: You are free to copy, share, redistribute, remix, transform, and build upon the material for any purpose, even commercially. You must give attribution and appropriate credit. Follow the terms: `https://creativecommons.org/licenses/by/4.0/`.

4.3 Data collection

The initial dataset captured web pages of stolen credentials sold in the Database market from November 2021 to June 2022. Based on that dataset, I extracted a certain of key fields, including *product, prices, dates* and *seller* for the *text* field. In addition, I manually took screenshots of pages showing information of products and stores in Database Market. In order to access a marketplace hosted in Tor network, it is common that you need an invitation code to register an account on the marketplace. However, in case of Database Market, I can create an account without an invitation code.

Filtering interesting fields from the text content in the webpage of product requiring some marking letters. For example, I noticed that the title of item sold



```
{
  "id": "5622185",
  "time-stamp": "2022-04-29",
  "seller": "Goldmaster",
  "product": "FashionNova.com",
  "prices": [
    3
  ],
  "dates": [
    "2022-03-29"
  ]
}
```

Figure 4: An example JSON entry from my customized dataset

Total number of items	53815
Total number of stores	30
Maximum price	2500.00 USD
Minimum price	0.20 USD
Average price	15.87 USD
Median price	5.00 USD

Table 1: Statistical result of the *ProductPages* dataset.

in Database marketplace is enclosed by "######", and the store name is covered by "###". However, it is more complicated in the cases of *prices* and *dates*. As criminals can unconstrainedly format the description of their products, I am unable to conduct a common pattern to capture all prices and dates of uploaded items. Thus, in addition to the quite effective pattern, returned correct fields in most of pages, I added some criteria for specific cases. In terms of category, I filtered top common keywords in the title of products, and grouped items that contain the shared keyword into a category. For example, any titles containing "info", "ssn", "dob", and "dl" belong to the *Personal data* category.

4.4 Data analysis

Applying statistic to the whole dataset, *ProductPages*, I obtained the following numerical results shown in Table 1.

The maximum 2500 USD, or other prices greater than 1000 USD, appears with a low frequency, an outlier. In this case, 2500 USD is a price of a set of 1000 personal data entries. On the other hand, the average price of personal data is 7.73 USD shown in Table 4. The price of personal data item depends on the amount of data pieces included. For example, in the product ID 978, 100 USA SSN+DOB is sold at 80 USD, while a buyer has to purchase 250 USD to retrieve 500 USA SSN+DOB. In addition, the quality of identity information, also influences the price. For example, in the product ID 5507764, a set of SSN+DOB+DL+MMN is sold at 5 USD, while a batch of SSN+DOB+CS in ID 5434366 is assigned at 7 USD. The quality of identity information here is interpreted as the level of financial impact and the simplicity to expose other credential data if this information is compromised.

There are a certain of product types are traded in Database marketplace. However, two main categories that most of traded items belong to are *Personal data* and *Online account*, as shown in Figure 5. *Online account* category contains compromised online accounts from popular services, including *Netflix*, *Amazon*, *Venmo*, etc. A full list of attacked websites is accessible via `https://github.com/ancuongnguyen07/Database_Market/blob/main/analysis_result/leaked_websites.txt`.

There are in total 30 stores, or sellers, in the database. *Goldmaster*, the most

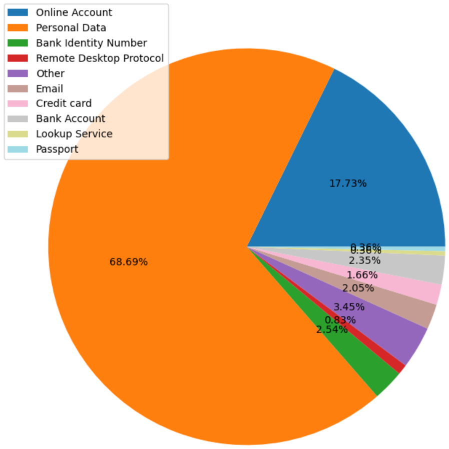


Figure 5: Product types allocation.

active seller, has sold the largest amount of products during 11/2021–06/2022, 11882 items. However, *bussman shop*, a store that has the greatest average price (430 USD), only sold 9 items. Furthermore, *bussman shop*, provides service of customizing ID documents for EU/US; thus, the prices of it's products are significant higher than those of other shops, ranging from 80 USD to 1200 USD. A table of full numerical results from analyzing stores is accessible in `https://github.com/ancuongnguyen07/Database_Market/blob/main/analysis_result/seller_stat.csv`.

4.5 Limitations

The initial database recorded product information in 8 months, from November 2021 to June 2022, covering 33896 product pages. First, the time of writting this study, November 2022, is 5-month after the last captured product in the database; therefore, observations and propositions instroduced in this study do not provide the most updated status of the Database marketplace. In comparison to product entries in the original database, the marketplace now provides more items; and many products recorded in the database is now either withdrawed or sold out, i.e product ID 5609797 (**??**).

In Database marketplace, shops can unconstrainedly format their product description; thus, I cannot filter exactly interesting fields in all cases, resulting

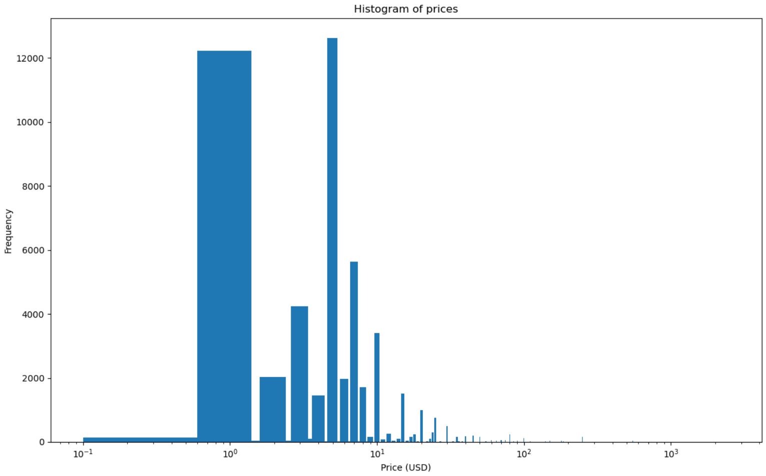


Figure 6: Price distribution.

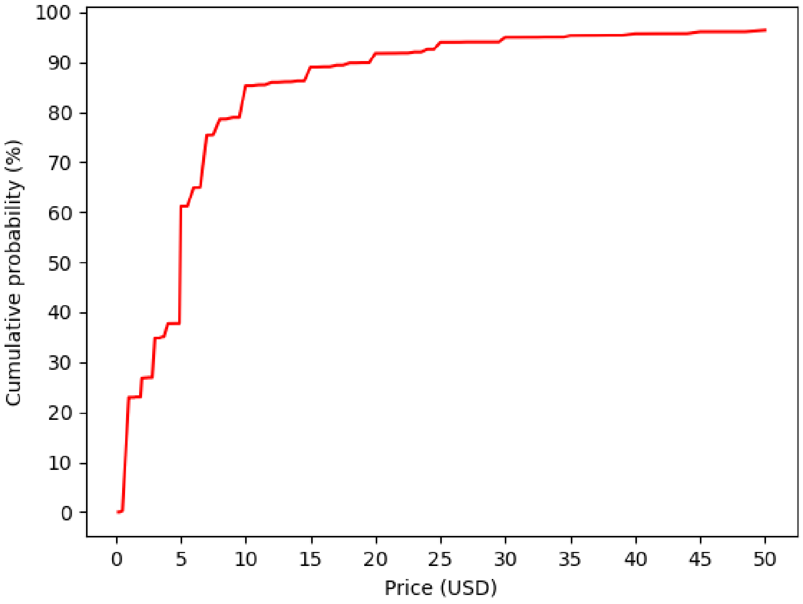
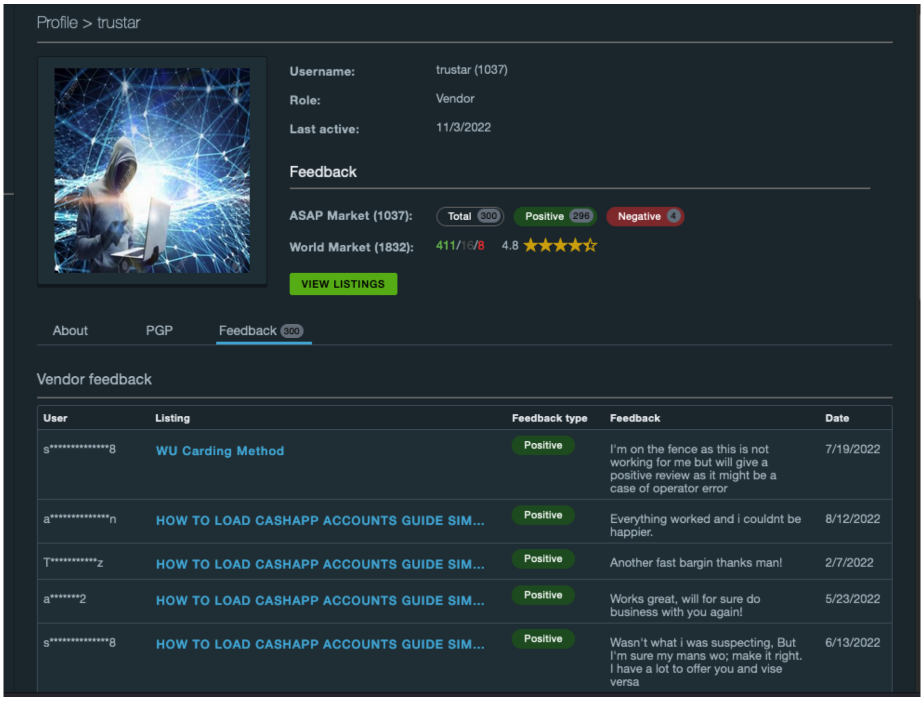


Figure 7: Cumulative probability of price range 0–50 USD.



Profile > trustar

Username: trustar (1037)

Role: Vendor

Last active: 11/3/2022

Feedback

ASAP Market (1037): Total 300 Positive 296 Negative 4

World Market (1832): 411/16/8 4.8 ★★★★☆

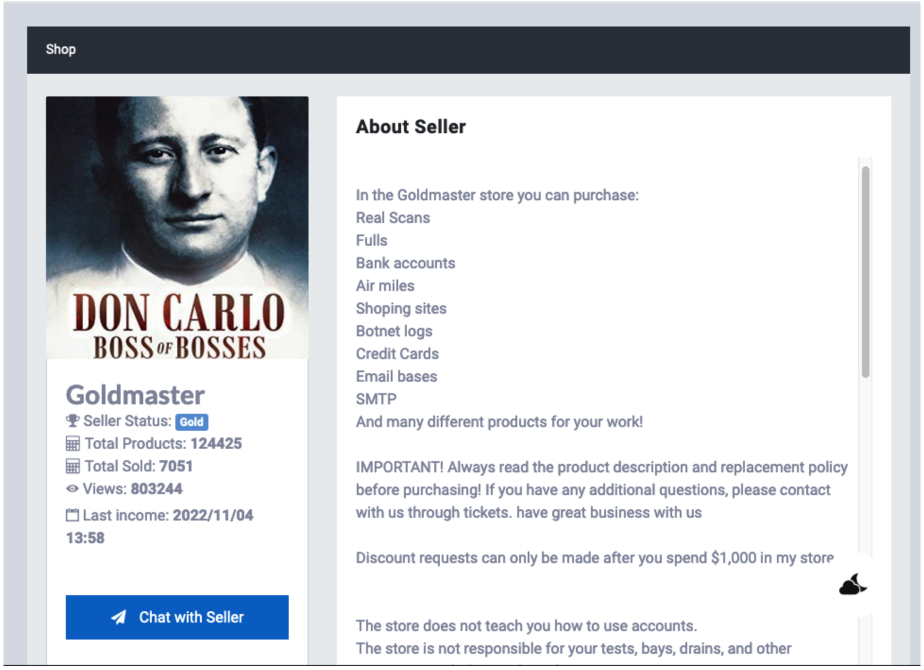
Vendor feedback

User	Listing	Feedback type	Feedback	Date
s**************8	WU Carding Method	Positive	I'm on the fence as this is not working for me but will give a positive review as it might be a case of operator error	7/19/2022
a**************n	HOW TO LOAD CASHAPP ACCOUNTS GUIDE SIM...	Positive	Everything worked and i couldnt be happier.	8/12/2022
T**********z	HOW TO LOAD CASHAPP ACCOUNTS GUIDE SIM...	Positive	Another fast bargin thanks man!	2/7/2022
a*******2	HOW TO LOAD CASHAPP ACCOUNTS GUIDE SIM...	Positive	Works great, will for sure do business with you again!	5/23/2022
s**************8	HOW TO LOAD CASHAPP ACCOUNTS GUIDE SIM...	Positive	Wasn't what i was suspecting, But I'm sure my mans wo; make it right. I have a lot to offer you and vise versa	6/13/2022

Figure 8: A shop page in ASAP marketplace.

minor noise in the examination. Moreover, product category is not provided in the inital database, not even in the product page. In order to attain product type data, I grouped products that have the common keywords in a category as presented in Table 4. This text-based clustering method is not reliable in all cases.

I examined that there is a shortage of feedback information about stores, or sellers, in the Database market. As shown in Figure 8, in the ASAP marketplace, feedbacks from customers are classified as negative and positive, that is helpful to detect an unreliable shop. On the other hand, as presented in Figure 9, Database market does not display feedbacks a specific store, but buyers can send comments via the internal ticket which is not shown in the store dashboad.



Goldmaster

Seller Status: Gold

Total Products: **124425**

Total Sold: **7051**

Views: **803244**

Last income: **2022/11/04 13:58**

About Seller

In the Goldmaster store you can purchase:
Real Scans
Fulls
Bank accounts
Air miles
Shoping sites
Botnet logs
Credit Cards
Email bases
SMTP
And many different products for your work!

IMPORTANT! Always read the product description and replacement policy before purchasing! If you have any additional questions, please contact with us through tickets. have great business with us

Discount requests can only be made after you spend $1,000 in my store

The store does not teach you how to use accounts.
The store is not responsible for your tests, bays, drains, and other

Figure 9: A shop page in Database marketplace.

Product type	Description
Personal Data	Full name, date of birth, social security number, home address, zip code, driver license, employer info, etc.
Online Account	Online account of services such as Netflix, Amazon, Venmo, Booking.com, etc.
Bank Account	Username/password of online bank accounts, and credential information linked to bank accounts.
Credit Card	Card holder name, CVV, full name, address, bank name, card type.
Passport	Real photos and scan of passports.
Email	Emails leaked through data breaches.
Lookup Service	SSN/DOB/MMN/DL /CS lookup service.
Bank Identity Number	BIN is the first 6 digits of the credit card (with the help of this information you will understand what credit cards you need to work with 3-D Secure VBV) 3-D Secure VBV is a protocol that is used as an additional layer of security for online credit and debit cards, two-factor user authentication.
Remote Desktop Protocol	The protocol provides user graphical interface to connect to other computers over the network connection [19].
Other	Tutorials, digital document templates, document forms, domains, etc.

Table 2: Description of product types.

5 Results

In this section, I will provide logical explaination and statistical results, illustrated as figures or tables, to address research questions mentioned in section 3.

5.1 RQ1: What types of stolen information are traded?

As shown in Figure 5, I captured in total 10 product types based on my own standard. Among all categories, *Personal Data* is the most active field, contributing 68, 69% of all traded items. *Online Account* represents the second larges portion 17, 73%. The rest of product types ranges from 0, 36% to 3, 45%. Detailed description of each product type is in Table 2.

Name of service	Year	Number of victims
Capital One	2019	100 million Americans and 6 million Canadians
Yahoo	2013, 2014	3 billion, 500 million
Marriott International	2018	500 million
First American Financial	2019	885 million
Facebook	2019	540 million

Table 3: List of major data breaches [8, 2, 11].

5.2 RQ2: What are the prices for different product types?

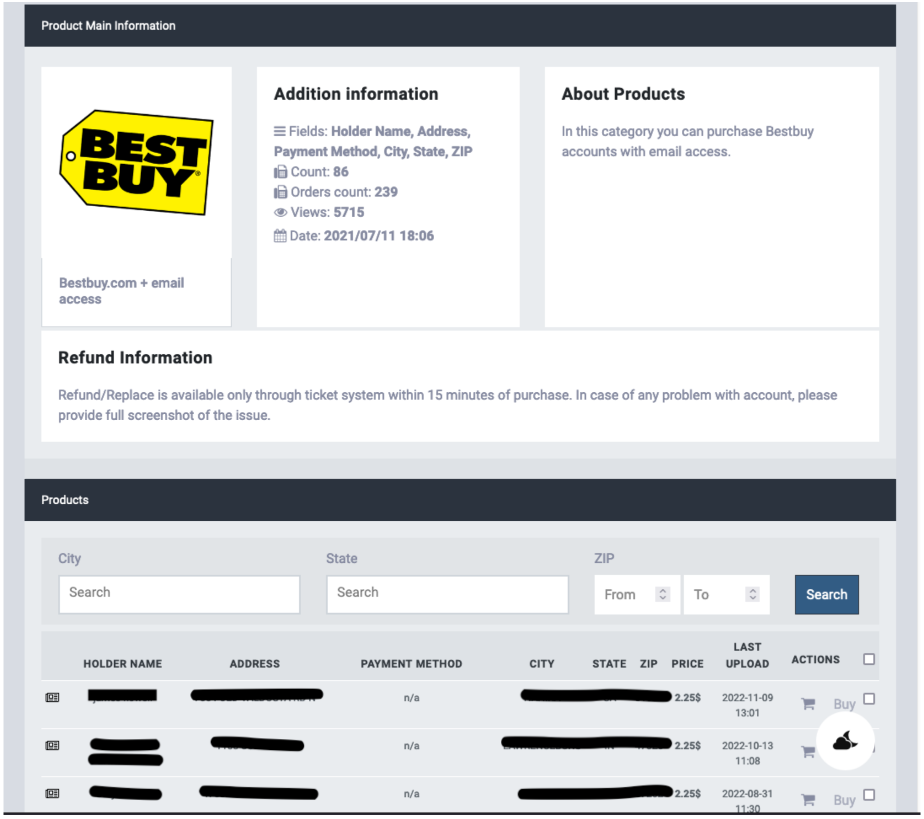
There are 6/10 categories, providing more than 90% of all traded items, recording the arithmetic mean price less than 20 USD. Furthermore, it supports the numerical result of 15,87 USD (Table 1), the average price of the whole dataset. The average price of compromised emails is 158,5 USD, the highest number over all product types. About 10% of them are email leads, being recorded 939,19 USD of average price. That explains why the average price of email type is significantly higher than the overall mean price. In constrast, products in *Personal Data* field is observed the lowest average price — 8,71 USD. It can be infered that regular Internet users are able to afford others' indentity information at unexpected-low cost.

5.3 RQ3: Where the stolen information is from?

The traded information is mainly from data breaches at high-traffic web services. In Table 3, I collected a list of major data breaches. A potential dataset from the breach is traded in Database marketplace, as shown in Figure 10. Information leaked through data breaches can be email address, password, or more sensitive data, including banking account, SSN, and identity information. Moreover, exposed data can be utilized in attacking campaigns that either trick victims to provide more their own personal data or unauthorizedly access others account owned by the same user.

5.4 RQ4: What we can learn about the sellers?

In general, there are two types of sellers, a store sells a large amount of low-cost items and the one provides a small amount of expensive products (Figure 11). In anonymous marketplaces, personal contacts of the vendors are unavailable. To become a seller, a visitor has to provide a list of his/her available products. Moreover, feedbacks from other anonymous forums and markeplaces need to be sent to administrators of Database market to prove that the applicant is a valid merchant.



Product Main Information

Bestbuy.com + email access

Addition information

☰ Fields: **Holder Name, Address, Payment Method, City, State, ZIP**
Count: **86**
Orders count: **239**
Views: **5715**
Date: **2021/07/11 18:06**

About Products

In this category you can purchase Bestbuy accounts with email access.

Refund Information

Refund/Replace is available only through ticket system within 15 minutes of purchase. In case of any problem with account, please provide full screenshot of the issue.

Products

City: Search
State: Search
ZIP: From To
Search

	HOLDER NAME	ADDRESS	PAYMENT METHOD	CITY	STATE	ZIP	PRICE	LAST UPLOAD	ACTIONS	□
	[redacted]	[redacted]	n/a	[redacted]	[redacted]	[redacted]	2.25$	2022-11-09 13:01	Buy	□
	[redacted]	[redacted]	n/a	[redacted]	[redacted]	[redacted]	2.25$	2022-10-13 11:08		
	[redacted]	[redacted]	n/a	[redacted]	[redacted]	[redacted]	2.25$	2022-08-31 11:30	Buy	□

Figure 10: Leaked information of Bestbuy's customer.

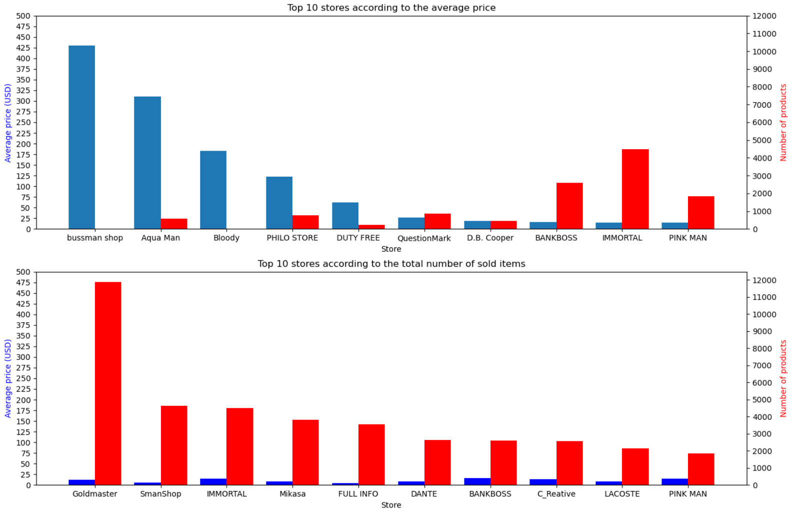


Figure 11: Top sellers sorted based on the average price and total number of traded products.

6 Conclusions

It is recorded that there are more than 53815 items was traded in the period 11/2021–06/2022, and more than half of them — 68,69% — are personal data , including full name, Date of Birth, Social Security Number, Driver's License, Credit Score, etc. The second most traded product type is username and password pair for accessing online services of banking, entertaining, and shopping. Both aforementioned product categories are traded at unexpectedly low cost, less than 15 USD for an entry or a set of data. The cheapest item are sold at only 0.2 USD for an entry of personal data (*full name + date of birth + social security number*). Those personal information is highly possible that are stolen from major data breaches. Through those data breaches, an attacker gains a set of intial dataset of credential data that is utilized to trick people trustfully provide more their sensitive data or access to their online accounts of essential services such as banking or social security. The consequences of identity theft have been documented thoroughly. For example, identity thieves can use stolen SSNs to apply for more credit and then do not pay the bills, damaging credit of compromised identities [9]. As there is a increasing number of data breaches over time, resulting the personal data is traded at low-cost and highly available, the negative consequences can impact a wider range of people.

Besides stolen sensitive data, Database marketplace also provides illicit services such as faking passports, and other identity documents. Ranging from 20 to 1200 USD, you can order a customized identity document. In addition to an affordable price range, those services are coveniently traded through the Tor-based marketplace, Database in this study. The marketplace supports functions, including chatting with sellers, searching keywords, and refunding orders, as the normal marketplace, e.g, Ebay.

In conclusion, Database marketplace provides a wide range of affordable products which can be utilized to do illegal activities. Moreover, these products are highly available and sold at reasonable prices; thus, an increasing number of criminal acts related to identity theft, bank withdrawing, or credit abusing has been observed. Those acts have rocketing impact on victims, not only financial but also mental health. The result of this study may raise the awareness of Internet users and other organizations for protecting sensitive data.

References

[1] MONICA J. BARRATT. “SILK ROAD: EBAY FOR DRUGS”. In: *Addiction* 107.3 (2012), pp. 683–683. DOI: https://doi.org/10.1111/j.1360-0443.2011.03709.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1360-0443.2011.03709.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1360-0443.2011.03709.x.

[2] Clifford Colby. *Capital One data breach: What you can do now following bank hack.* 2019. URL: https://www.cnet.com/tech/computing/capital-one-data-breach-what-you-can-do-now-following-bank-hack/ (visited on 11/16/2022).

[3] Roger Dingledine, Nick Mathewson, and Paul Syverson. “Tor: The Second-Generation Onion Router”. In: *Proceedings of the 13th Conference on USENIX Security Symposium - Volume 13.* SSYM’04. San Diego, CA: USENIX Association, 2004, p. 21.

[4] Diana S Dolliver. “Evaluating drug trafficking on the Tor Network: Silk Road 2, the sequel”. en. In: *Int. J. Drug Policy* 26.11 (Nov. 2015), pp. 1113–1123.

[5] Caliskan Emin, Minarik Tomasm, and Osula Anna-Maria. *Technical and Legal Overview of the Tor Anonymity Network.* Tech. rep. NATO Cooperative Cyber Defence Centre of Excellence, 2015.

[6] Sean Foley, Jonathan R. Karlsen, and Tālis J. Putniņš. “Sex, drugs, and Bitcoin: How much illegal activity is financed through cryptocurrencies?” English. In: *Review of Financial Studies* 32.5 (May 2019), pp. 1798–1853. ISSN: 0893-9454. DOI: 10.1093/rfs/hhz015.

[7] ROBERT AUGUSTUS HARDY and JULIA R. NORGAARD. “Reputation in the Internet black market: an empirical and theoretical analysis of the Deep Web”. In: *Journal of Institutional Economics* 12.3 (2016), pp. 515–539. DOI: 10.1017/S1744137415000454.

[8] Sean Hollister. *Best Buy hit by [24]7.ai data breach, too.* 2018. URL: https://www.cnet.com/news/privacy/best-buy-data-breach-24-7-ai/ (visited on 11/09/2022).

[9] *Identity Theft and Your Social Security Number.* URL: https://www.ssa.gov/pubs/EN-05-10064.pdf (visited on 11/16/2022).

[10] Satoshi Nakamoto. *Bitcoin: A Peer-to-Peer Electronic Cash System.* 2008.

[11] Niles Nerd. *Is your information being sold on dark web.* 2019. URL: https://www.nerdsonsite.com/blog/is-your-information-being-sold-on-dark-web/ (visited on 10/23/2022).

[12] Juha Nurmi. “Understanding the Usage of Anonymous Onion Services: Empirical Experiments to Study Criminal Activities in the Tor Network”. PhD thesis. Tampere University, 2019.

[13] Juha Nurmi and Mikko S. Niemelä. “Tor De-anonymisation Techniques”. In: *Network and System Security.* Ed. by Zheng Yan et al. Cham: Springer International Publishing, 2017, pp. 657–671. ISBN: 978-3-319-64701-2.

[14] Tor project. *How Do Onion Services Work?* URL: `https://community.torproject.org/onion-services/overview/` (visited on 11/17/2022).

[15] Tor project. *Tor:Overview.* URL: `https://2019.www.torproject.org/about/overview.html.en#whyweneedtor` (visited on 11/17/2022).

[16] Tor project. *Types of relays.* URL: `https://community.torproject.org/relay/types-of-relays/` (visited on 11/17/2022).

[17] Andrei Serjantov. "On the anonymity of anonymity systems". PhD thesis. University of Cambridge, 2004.

[18] Marie Claire Van Hout and Tim Bingham. "'Surfing the Silk Road': a study of users' experiences". en. In: *Int. J. Drug Policy* 24.6 (Nov. 2013), pp. 524–529.

[19] Wikipedia. *Remote Desktop Protocol.* URL: `https://en.wikipedia.org/wiki/Remote_Desktop_Protocol` (visited on 11/14/2022).

A Tables of statistic results of stores and item categories

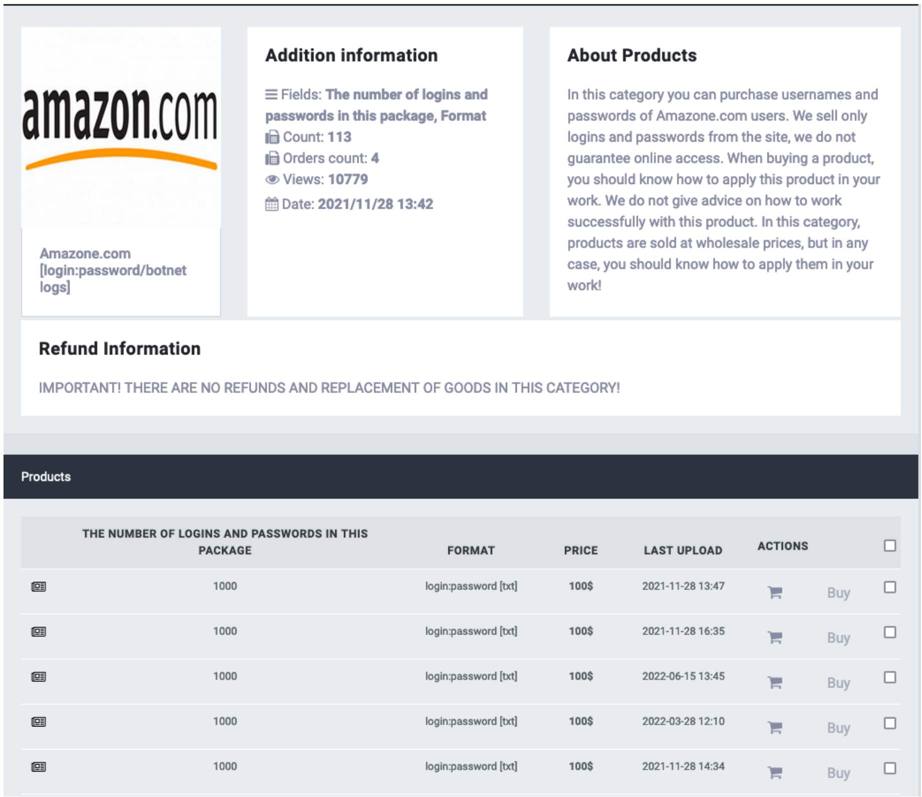
Category	Amount of products	Average price
Personal Data	36963	8.71
Online Account	9539	13.88
Other	1855	68.54
Bank Identity Number	1368	9.24
Bank Account	1266	48.49
Email	1102	158.50
Credit card	892	11.00
Remote Desktop Protocol	446	10.85
Lookup Service	192	33.30
Passport	192	14.94

Table 4: Amount of products and average price of product categories. The currency of price is USD.

Store	Amount of products	Range of types	Min price	Max price	Average pirce	Median price
Goldmaster	11882	4	1.0	2500.0	12.24	3.0
SmanShop	4614	1	1.0	35.0	5.24	5.0
JOKER	1700	5	1.0	500.0	13.72	6.0
Goodnik7	1505	2	1.0	15.0	5.06	7.0
Mikasa	3811	5	1.0	200.0	8.22	5.0
PINK MAN	1850	4	1.0	1000.0	14.35	5.0
DANTE	2644	3	0.5	490.0	8.38	3.0
Dark Shop	1444	6	1.0	350.0	6.5	1.0
LACOSTE	2124	4	1.0	500.0	8.23	8.0
IMMORTAL	4484	7	0.5	1600.0	14.48	7.0
Radikula Store	776	4	0.5	350.0	13.41	6.0
C_Reative	2569	5	1.0	2000.0	13.13	10.0
Hurricane	1254	6	0.5	700.0	7.83	4.0
FULL INFO	3550	2	1.0	80.0	3.74	4.0
D.B. Cooper	441	3	1.0	300.0	19.04	5.0
QuestionMark	857	3	1.0	650.0	26.45	5.0
NERO	1229	4	1.0	1000.0	12.81	1.0
Carsh	92	2	3.5	75.0	10.14	3.5
makataO	833	2	1.0	1000.0	10.51	1.0
The Best Banks	641	5	0.5	80.0	11.12	10.0
bussman shop	9	1	80.0	1200.0	430.0	150.0
DUTY FREE	229	5	2.0	250.0	62.47	10.0
PHILO STORE	770	7	1.0	550.0	122.22	18.0
Ninja Secret's	421	4	7.0	50.0	8.8	7.0
BANKBOSS	2583	6	1.0	670.0	16.64	10.0
Aqua Man	590	5	0.2	1900.0	310.68	15.0
PROMETHEUS	315	3	1.0	60.0	6.62	1.0
BestLink	588	4	3.0	250.0	10.21	6.0
Spamking	7	1	7.0	7.0	7.0	7.0
Bloody	3	1	150.0	200.0	183.33	200.0

Table 5: Statistical results of stores. The currency of price is USD.

B Screenshots of Database marketplace



Amazone.com [login:password/botnet logs]

Addition information

☰ Fields: **The number of logins and passwords in this package, Format**
Count: **113**
Orders count: **4**
Views: **10779**
Date: **2021/11/28 13:42**

About Products

In this category you can purchase usernames and passwords of Amazone.com users. We sell only logins and passwords from the site, we do not guarantee online access. When buying a product, you should know how to apply this product in your work. We do not give advice on how to work successfully with this product. In this category, products are sold at wholesale prices, but in any case, you should know how to apply them in your work!

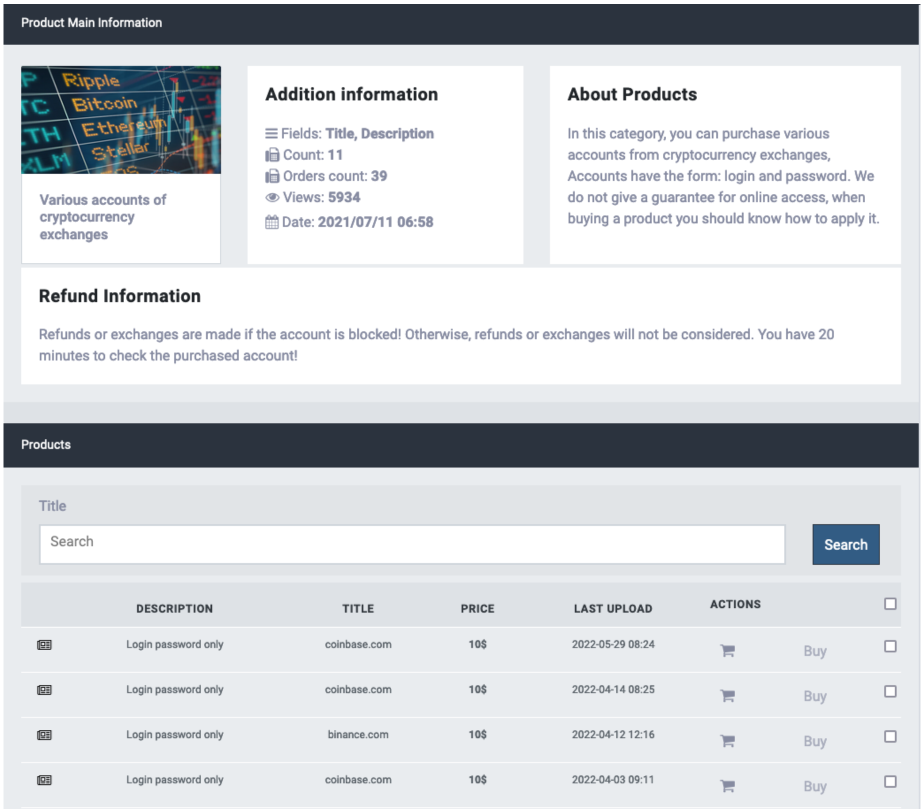
Refund Information

IMPORTANT! THERE ARE NO REFUNDS AND REPLACEMENT OF GOODS IN THIS CATEGORY!

Products

	THE NUMBER OF LOGINS AND PASSWORDS IN THIS PACKAGE	FORMAT	PRICE	LAST UPLOAD	ACTIONS		□
	1000	login:password [txt]	100$	2021-11-28 13:47	🛒	Buy	□
	1000	login:password [txt]	100$	2021-11-28 16:35	🛒	Buy	□
	1000	login:password [txt]	100$	2022-06-15 13:45	🛒	Buy	□
	1000	login:password [txt]	100$	2022-03-28 12:10	🛒	Buy	□
	1000	login:password [txt]	100$	2021-11-28 14:34	🛒	Buy	□

Figure 12: Leaked Amazon accounts



Product Main Information

Various accounts of cryptocurrency exchanges

Addition information

☰ Fields: **Title, Description**
Count: **11**
Orders count: **39**
Views: **5934**
Date: **2021/07/11 06:58**

About Products

In this category, you can purchase various accounts from cryptocurrency exchanges, Accounts have the form: login and password. We do not give a guarantee for online access, when buying a product you should know how to apply it.

Refund Information

Refunds or exchanges are made if the account is blocked! Otherwise, refunds or exchanges will not be considered. You have 20 minutes to check the purchased account!

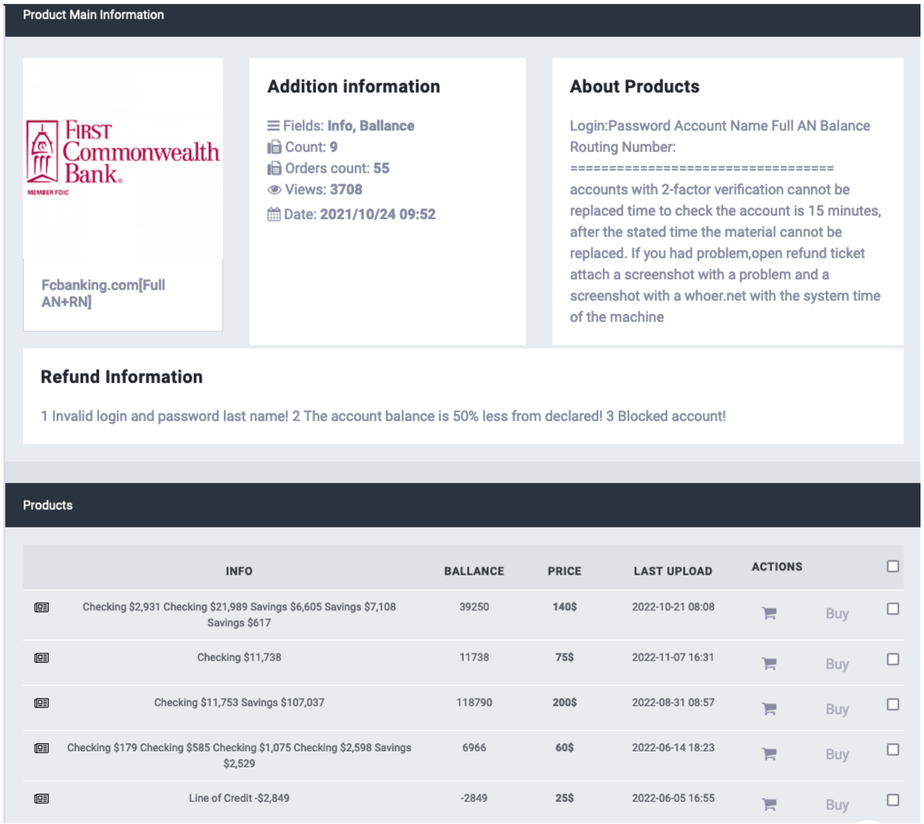
Products

Title

Search | Search

	DESCRIPTION	TITLE	PRICE	LAST UPLOAD	ACTIONS		□
	Login password only	coinbase.com	10$	2022-05-29 08:24	🛒	Buy	□
	Login password only	coinbase.com	10$	2022-04-14 08:25	🛒	Buy	□
	Login password only	binance.com	10$	2022-04-12 12:16	🛒	Buy	□
	Login password only	coinbase.com	10$	2022-04-03 09:11	🛒	Buy	□

Figure 13: Leaked Cryptocurrency exchange accounts



Product Main Information

Fcbanking.com[Full AN+RN]

Addition information

☰ Fields: **Info, Ballance**
Count: **9**
Orders count: **55**
Views: **3708**
Date: **2021/10/24 09:52**

About Products

Login:Password Account Name Full AN Balance Routing Number:
====================================
accounts with 2-factor verification cannot be replaced time to check the account is 15 minutes, after the stated time the material cannot be replaced. If you had problem,open refund ticket attach a screenshot with a problem and a screenshot with a whoer.net with the system time of the machine

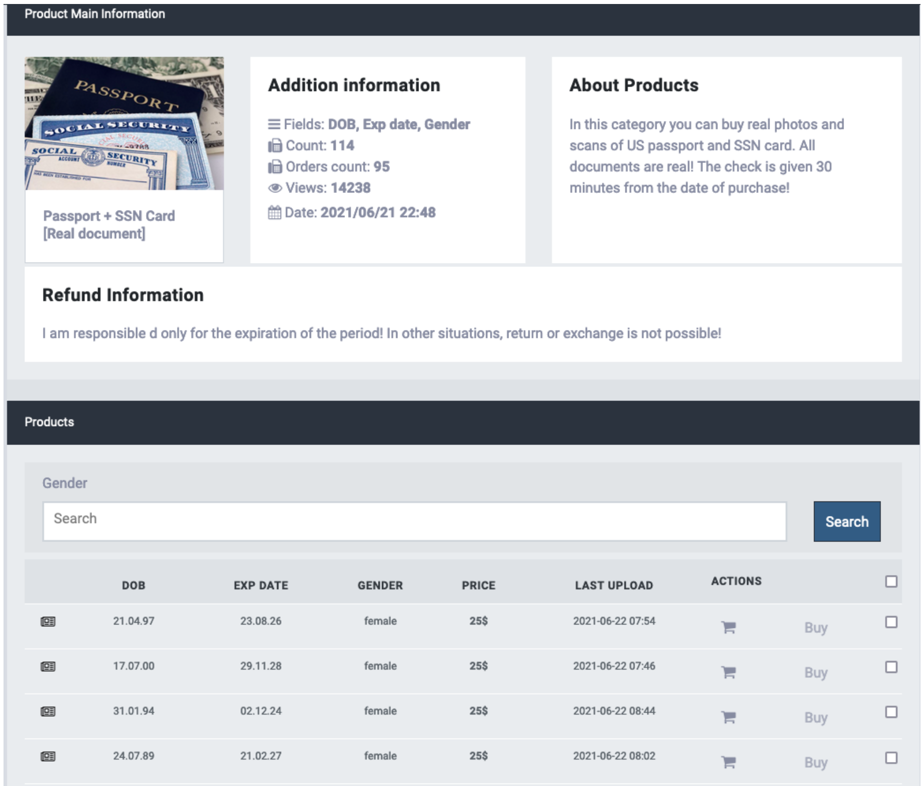
Refund Information

1 Invalid login and password last name! 2 The account balance is 50% less from declared! 3 Blocked account!

Products

	INFO	BALLANCE	PRICE	LAST UPLOAD	ACTIONS		☐
	Checking $2,931 Checking $21,989 Savings $6,605 Savings $7,108 Savings $617	39250	**140$**	2022-10-21 08:08		Buy	☐
	Checking $11,738	11738	**75$**	2022-11-07 16:31		Buy	☐
	Checking $11,753 Savings $107,037	118790	**200$**	2022-08-31 08:57		Buy	☐
	Checking $179 Checking $585 Checking $1,075 Checking $2,598 Savings $2,529	6966	**60$**	2022-06-14 18:23		Buy	☐
	Line of Credit -$2,849	-2849	**25$**	2022-06-05 16:55		Buy	☐

Figure 14: Leaked bank accounts



Product Main Information

Passport + SSN Card
[Real document]

Addition information

☰ Fields: **DOB, Exp date, Gender**
Count: **114**
Orders count: **95**
Views: **14238**
Date: **2021/06/21 22:48**

About Products

In this category you can buy real photos and scans of US passport and SSN card. All documents are real! The check is given 30 minutes from the date of purchase!

Refund Information

I am responsible d only for the expiration of the period! In other situations, return or exchange is not possible!

Products

Gender

Search

	DOB	EXP DATE	GENDER	PRICE	LAST UPLOAD	ACTIONS		□
🪪	21.04.97	23.08.26	female	**25$**	2021-06-22 07:54	🛒	Buy	□
🪪	17.07.00	29.11.28	female	**25$**	2021-06-22 07:46	🛒	Buy	□
🪪	31.01.94	02.12.24	female	**25$**	2021-06-22 08:44	🛒	Buy	□
🪪	24.07.89	21.02.27	female	**25$**	2021-06-22 08:02	🛒	Buy	□

Figure 15: Leaked passports