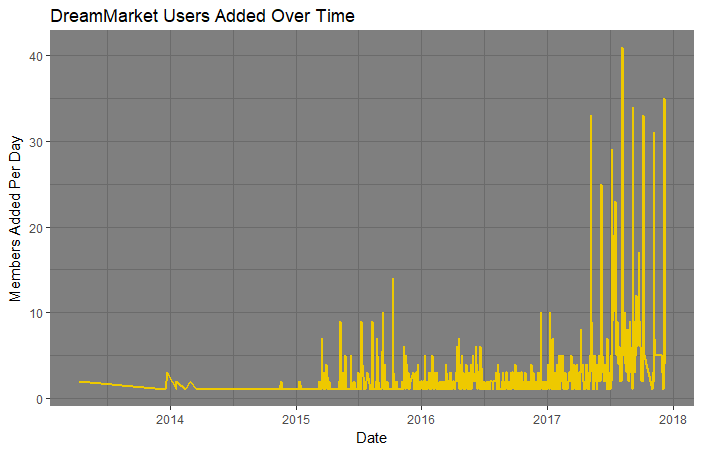
[[1]](#footnote-1)

Black Market Product Sales

Ryan Bass, Brandon Hill, Ann Nelson*Abstract*— An analysis of the [The Dark Net Market SQL dataset](http://www.azsecure-data.org/dark-net-markets.html), available at [www.azsecure-data.org](http://www.azsecure-data.org), developed by the University of Arizona’s Artificial Intelligence Lab was used to investigate what types of goods are being sold, common encryption programs used, and public keys used by sellers of digital goods sellers. Analysis of the encryption programs and public keys could provide insight to how secure the transferring of digital goods is in Dream Market. This dataset also includes data about the items each user is selling, allowing insight into which users are key players in this dark net market. The Python analyses include data cleaning to ensure a solidly structured dataset, exploratory data analysis to understand trends, correlations and other patterns/relationships in the data. Key Python libraries used in this analysis included numpy, pandas, and matplotlib.

*Index Terms*—Network Security, Information Security, Encryption Protocols

O

# INTRODUCTION

NLINE cybercriminal websites are used for advertising, buying, and selling of illegal products. These websites are termed ‘Dark Net Markets’ (DNMs). DNMs are relatively new compared to other merchandizing platforms such as Amazon, Ebay, or CraigsList. National security agencies are particularly interested in monitoring the use of these systems in order to halt or diminish the online sale of illegal products. These investigations are often complicated by international sales of these products.

In July 2017, after several months of joint investigation between Thailand, the Netherlands, Lithuania, Canada, the United Kingdom, and France, two dark web marketing sites, AlphaBay and Hansa Market, were shutdown. Shortly after these takedowns, Dream Market (the second largest DNMs before the collapse of AlphaBay) membershi exploded (Figure 1)t became the largest Dark Net Market.

Figure 1: Dream Market members added per day. Username data was associated with the date the user made an account. These time stamps show a marked increase in memberships during the shutdown of AlphaBay and Hansa dark net markets in July 2017

# Data Exploration

The Dark Net Market SQL dataset at [www.AZsecure-data.org](http://www.AZsecure-data.org), was developed by the University of Arizona’s Artificial Intelligence Lab. The data was collected from Dream Market using a python-based webcrawler and parsed via [Beautiful Soup](https://www.crummy.com/software/BeautifulSoup/), a python library for modifying, searching and parsing web pages. The data set contains a product table, and a sellers table. The product table contains Dream Market’s product data from 2013-2017 including product name, category, description, shipping options, shipping departure, shipping destination, price, and payment information for 39,473 products. The sellers table provides data about Dream Market’s 2,092 sellers who participated on Dream Market from 2013-2017. Attributes in the sellers table include the seller’s name, how long they’ve been a member, pgp\_key, seller’s description, and feedback rating.

With 91,463 products being offered by 2,092 different usernames on Dream Market, data mining and statistical analytics weere used to decipher the contents of the database generated by AZSecure-data. After mining and cleaning of the database, the products being sold on Dream Market were classfied into three major categories; drugs, digital goods, and services.

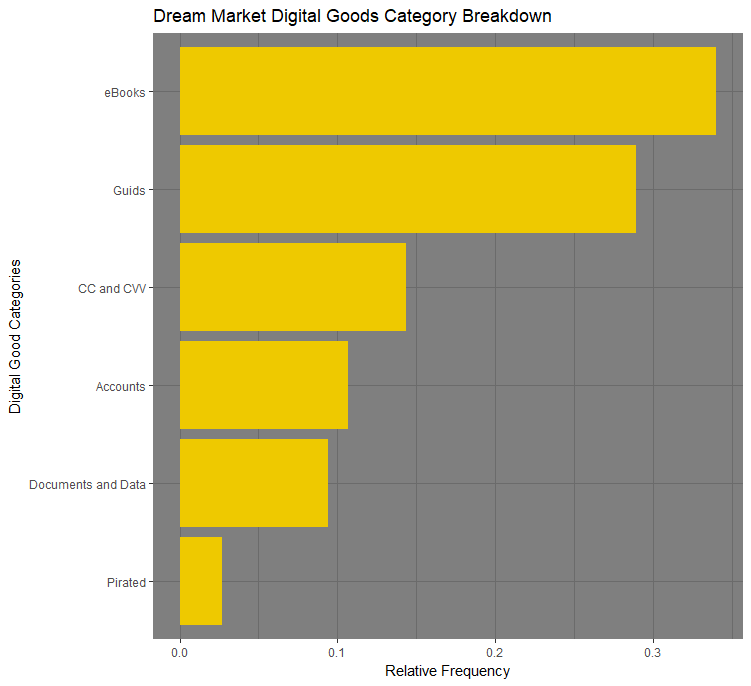
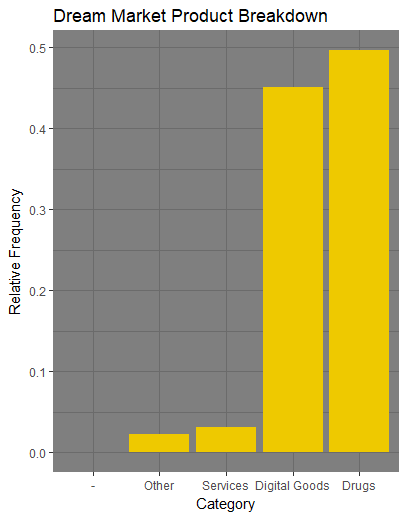
After the closing of the previous dark web market mainstay AlphaBay, Dream Market saw its registered users skyrocket with illegal goods and services being exchanged everyday Its user base relied on their information being kept secret over the dark web using anonymized connections utilizing Tor and virtual private networks (VPN) to access the Dream Market site.

Figure 2: Relative Frequency of DreamMarket Product Categories. Mining of the database created by University of Arizona Artificial Intelligence Lab yielded three major categories of products; drugs (49.69%), digital goods (45.02%), services (3.10%), and others (2.19%).

Drug sale on Dream Market makes up 49.69% of the overall goods provided. Of all the advertised drugs on Dream Market, marijuana makes up the largest percentage at 21.25%. This is to be expected as it is a plant whch can be grown indoors, so on a global site, like Dream Market, users from all over the world in different climates could potentially supply themselves. Cocaine was the second most sold drug at 13.60% of all the drugs on the site. With users being anonymized on the dark web, illegal drugs will always be on sale on websites such as this.

The sale of digital goods on Dream Market makes up 45.02% of the overall goods provided. A large majority of this are eBooks (at 37.68% of the digital goods available, or nearly 17% of the total amount of goods for sale), these can easily be bought and once the security features are disabled can be distributed cheaply to provide a profit for the user, This should be highly concerning to all eBook distributors as their encryption was put in place to keep from their books being readily distributed

# Encryption Technologies

Since digital goods represent 45% of the total products available on DreamMarket we wanted to analyze how these goods were being transferred from seller to buyers, therefore

Figure 3: Dream Market Digital Goods Categories. A majority of the digital goods available on Dream Market are eBooks and guides. Sensitive personal information items include credit card information, accounts on websites, and documents/data.

all further analysis refers to only digital goods sellers.

The dataset scraped by AZSecure-data.org includes information on the encryption program used by sellers as well as the public keys they’re inputting into their algorithm. We wanted to see if there were dominant programs used for encryption, and if there is a potential security lapse in this transfer of information that would lead to anonymity of the sale to be compromised. Out of the 578 sellers that sell digital goods and render services, 570 of them had their public key encryption methods listed. 742 sellers did not have their encryption means listed on their seller’s profile but did provide their public keys. This shows that they are still likely encrypting the data they transfer to buyers, just not overtly listing their encryption methodology. When analyzing the encryption programs used by these users, we grouped encryption programs versions together to provide a more comprehensive view of the programs used rather than the specific version (ex. GnuPGP v1.4.9, v1.4.10, etc were grouped as GnuPG v1). Results in Figure 5 show a high percentage users using the free cryptography program GnuPG. The only other encryption program with a significant userbase is BCPG (Bouncy Castle Public Key).

GnuPG (GNU Privacy Guard, or GPG) is a free hybrid cryptosystem based on Symantec’s PGP (Pretty Good Privacy), an encryption program made in 1991 by Phil Zimmerman to encrypt and decrypt files, e-mails, texts, directories, and even whole disk partitions. A hybrid cryptosystem combines public-key and symmetric-key cryptosystems whose general layout can be seen in Figure 6. OpenPGP was then derived from PGP to serve as a free open source encryption standard set by Internet Engineering Task Force (IETF) in RFC 4880. GnuPG is based on these OpenPGP standards and was made by Werner Koch in 1999.

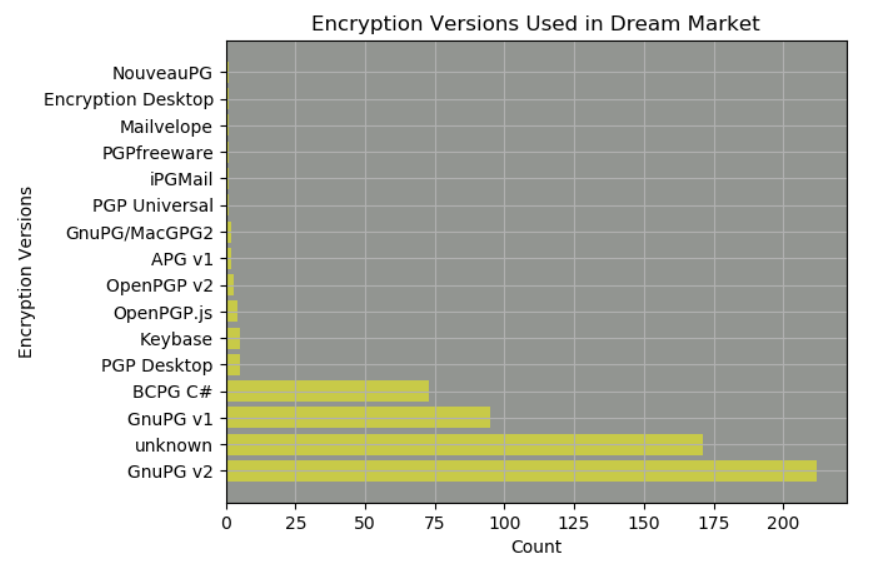
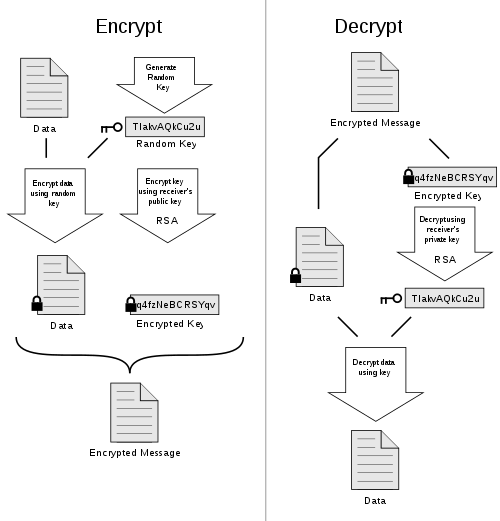


Figure 5: Of the reported encryption programs used by DreamMarket sellers (versions of GnuPG made up approximately 80% of the encryption programs used on the site. Only BCPG, with approximately 15% usage rate, was the only other program of note being used.



.

Figure 6: PGP encryption design layout.

1.) Symmetric key generated

2.) Symmetric key encrypts plaintext/data

3.) Symmetric key is encrypted using public key

4.) Both encrypted data and key sent to recipient

5.) Decrypt key using recipient’s private key

6.) Plaintext/data decrypted using symmetric key. [1]

GnuPG can serve as a “one stop shop” for algorithms needed at each step of cryptography. The list of functions and algorithms supported by GnuPG is listed in Table 1. GnuPG has been a popular way to encrypt digital communications for many years. It was made even more popular after 2013, expanding it’s userbase to non-dark web users, when a popular online video “GPG for Journalists” was published to help journalists keep their sources safe when receiving whistleblower information by encrypting their emails[2]. This video has garnered hundreds of thousands of views after Edward Snowden later confirmed he used GnuPG to help communicate with the journalist he was in touch with to leak NSA secrets. Because of GnuPG is free to use and popular it’s widespread use in the dark web market is understandable.

|  |  |
| --- | --- |
| GnuPG Functions | GnuPG Supported Algorithms |
| Public Key | DSA, EIGamal, RSA |
| Cipher | 3DES, AES-128, AES-192, AES-256, Blowfish, Camellia-128, Camellia-192, Camellia-256, CAST5, IDEA, Twofish |
| Hash | MD5, RIPEMD-160, SHA-1, SHA-224, SHA-256, SHA-384, SHA-512 |
| Compression | BZIP2, Uncompressed, ZIP, ZLIB |

Table 1: Lists of functions and their corresponding algorithms supported by GnuPG as of versions v1.4.18 and v2.0.26.

# Key Lengths

After looking at encryption techniques, the next data security information that was analyzed was the key length of the sellers. All the sellers had their public keys listed on their profile.We were able calculate the key lenghts and use them to judge the level of security of the encryptions used for communications/file transfers. Figure 7 shows a filtered histogram of the key lengths in the dataset of digital goods and services sellers, and a count of the most relevant groupings of key lengths.

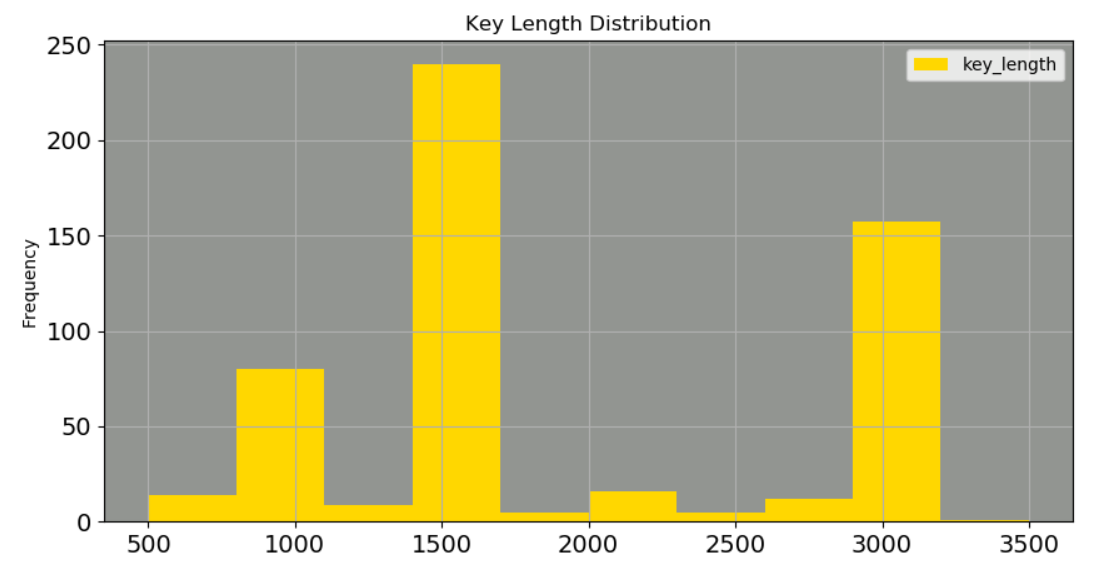


Figure 7: Distribution of key lengths between 500 and 3500 bit lengths for all users in the dataset. The graph was limited to more appropriately display the most common key lengths.

When the distribition is limited to key lengths between 500

and 3500 (94.6% of the data) the trends can be better seen. There are approximately 200 key lengths that are approximately 768 bits, the minimum key size set by GnuPG [3] for an ElGamal key. The other two spikes in key length counts are between about 1500-1750, and 3000. The 3000 key length spike is likely due to RSA or DSA 3072 bit keys, which RSA Security claims their 3072-bit keys have the strength of a 128-bit symmetric key [4]. As of 2016 even the NSA’s Commercial National Security Algorithm Suite includes RSA 3072-bit keys as the minimum encryption setting for key establishment [5]. This level of key encryption is estimated to withstand brute force attacks even from quantum computers for the foreseeable future.

# Encryption & Key Length

When comparing the key length distribution to each program being used by sellers on DreamMarket (Table 2) other trends can be found.

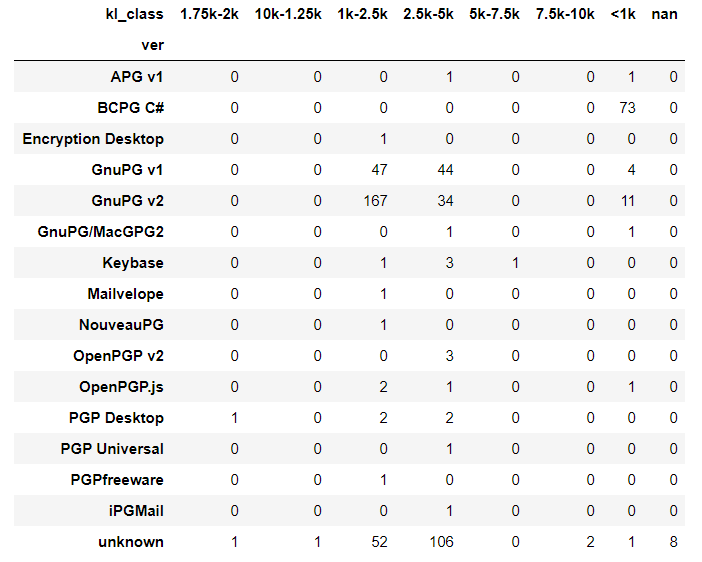


Table 2: Two-way table of key lengths and encryption programs used by DreamMarket sellers.

GnuPG v2 and BCPG are the only commonly used versions/programs that use key lengths below 1000-bits. As noted before with GnuPG, this is likely due to the use ot the minimum key length option in ElGamal, 768-bits in length. It seems likely that users simply opting to use the minimum suggested key length.

Key lengths of 128-bits are typically sufficient enough to ensure security. Larger key sizes exist to satify US military regulations to achieve specific security levels. Larger key sizes would require increaed resources to generate as well as considerable more resource to withstand an attack Because of the increased resources required to generate larger keys, there is some reasonable consideration for a user to use a key that is only as large as needed to ensure security. A brute-force attack on any k-bit key block cipher would only take O(2k/2)O(2k/2) steps, so a 256-bit key would still give 128-bit security, while a 128-bit key could be cracked in 2^64 operations, which is achievable.

# Conclusion

Analysis of the product and seller data for the Dream Market dark net e-commerce site provides interesting insights into the information security methods commonly used with this large data. The predominant encryption method, GnuPG, an open-source, hybrid encryption tool has proven to be secure with 4+ years of usage without breaking. Analysis of the key lengths documented in the data set suggests that about 10% of the sellers maybe at increased risk, because they had defaulted to the lowest key length used the the ElGamal Algorithm to generate their public keys, while about 25% of the sellers are using NSA’s recommended key length and are secure from brute force attacks.

References

1. “(Pretty Good Pivacy”, n.d.) Available:

“<https://upload.wikimedia.org/wikipedia/commons/4/4d/PGP_diagram.svg>)

1. anon108, “GPG for Journalists - Windows edition | Encryption for Journalists | Anonymous 2013” <https://vimeo.com/56881481>
2. g10 Code GmbH ,“The GNU Privacy Handbook”, July 21, 2018, <https://www.gnupg.org/gph/en/manual/c14.html>
3. B. Kaliski. ” TWIRL and RSA Key Size”, May 6, 2003, <https://web.archive.org/web/20170417095741/https://www.emc.com/emc-plus/rsa-labs/historical/twirl-and-rsa-key-size.htm>
4. NSA, CSS, “Information Assurance Directive”, January 2016, <https://cryptome.org/2016/01/CNSA-Suite-and-Quantum-Computing-FAQ.pdf>
5. Mohammadreza Ebrahimi, Po-Yi Du “Dark Net Markets Data Collection: Dream Market Dataset”. Publisher: University of Arizona Artificial Intelligence Lab, AZSecure-data, Director Hsinchun Chen. <https://www.azsecure-data.org/dark-net-markets.html>

1. Submittal Date: 7/22/2018. SMU MSDS-7349-401

   Professor Mohamed Meky.

   Ryan Bass, (email: [crbass@mail.smu.edu](mailto:crbass@mail.smu.edu))

   Brandon Hill (email: [bdhill@smu.edu](mailto:bdhill@smu.edu)­)

   Ann Nelson (email: [alnelson@smu.edu](mailto:alnelson@smu.edu)). [↑](#footnote-ref-1)