

Into the Reverie: Exploration of the Dream Market

Theo Carr¹, Jun Zhuang², Dwight Sablan³, Emma LaRue⁴, Yubao Wu⁵, Mohammad Al Hasan², and George Mohler²

¹Department of Mathematics, Northeastern University, Boston, MA

²Department of Computer & Information Science, Indiana University - Purdue University, Indianapolis, IN

³Department of Mathematics and Computer Science, University of Guam, Guam

⁴Department of Mathematics and Statistics, University of Arkansas at Little Rock, AK

⁵Department of Computer Science, Georgia State University, Atlanta, GA

carr.ken@husky.neu.edu, junz@iu.edu, sabland9706@gotritons.uog.edu, exlarue@ualr.edu,
exlarue@ualr.edu, ywu28@gsu.edu, alhasan@iupui.edu, gmohler@iupui.edu

Abstract—Since the emergence of the Silk Road market in the early 2010s, dark web ‘cryptomarkets’ have proliferated and offered people an online platform to buy and sell illicit drugs, relying on cryptocurrencies such as Bitcoin for anonymous transactions. However, recent studies have highlighted the potential for de-anonymization of bitcoin transactions, bringing into question the level of anonymity afforded by cryptomarkets. We examine a set of over 100,000 product reviews from several cryptomarkets collected in 2018 and 2019 and conduct a comprehensive analysis of the markets, including an examination of the distribution of drug sales and revenue among vendors, and a comparison of incidences of opioid sales to overdose deaths in a US city. We explore the potential for de-anonymization of vendors by implementing a Naïve-Bayes classifier to predict the vendor from a given product review, and attempt to link vendors’ sales to specific Bitcoin transactions. On the buyer side, we evaluate the efficacy of hierarchical agglomerative clustering for grouping together transactions corresponding to the same buyer. We find that the high degree of specialization among the small subset of high-revenue vendors may render these vendors susceptible to de-anonymization. Further research is necessary to confirm these findings, which are restricted by the scarcity of ground-truth data for validation.

Index Terms—Dark Web, Cryptomarket, Illicit Drug, Opioid

I. INTRODUCTION

Cryptomarkets are internet-based markets where consumers can purchase a variety of goods and services similar to the popular e-commerce marketplace, Ebay. However, unlike Ebay, cryptomarkets exist on the dark web, where websites cannot be accessed with traditional browsers. Users instead access the dark web using special software such as Tor Browser, which conceals the location and identity of a user browsing the internet [1]. The anonymity afforded by these dark web markets makes them hotspots for the exchange of illegal, stolen, or counterfeit goods, and for other illegal activities such as terrorism [2].

The first largely successful dark web marketplace, Silk Road, came into existence in 2011 [3]. When it was shut down by law enforcement in 2013, more marketplaces rose to take its place [4]. The proliferation of cryptomarkets was a call-to-action for world leaders and prompted “Operation

Onymous” in 2014, a worldwide action taken by law enforcement and judicial agencies aimed to put a kibosh on these illicit behaviors [5]. Law enforcement interventions such as Onymous, along with exit scams and hacks, have successfully shut down numerous cryptomarkets, including AlphaBay, Silk Road, Dream, and more recently, Wall Street [6]. Despite these interruptions, new markets have continued to proliferate. The authors of [7] note that there appears to be a consistent daily demand of about \$500,000 for illicit products on the dark web, and that the termination of individual markets has little long-term effect on this demand, as vendors and buyers are simply displaced to a different marketplace. A recent analysis found that the bitcoin equivalent of 700 million USD was sent to darknet markets in 2017 [8].

Past cryptomarket research has been largely limited to descriptive analyses, as opposed to predictive modeling, due to the challenge of validating results obtained using an anonymized dataset. In this work, we go beyond descriptive analysis by exploring predictive modeling and methods of de-anonymizing vendors and buyers. In light of the opioid crisis, we also compare the frequency of opioid-specific transactions on several cryptomarkets to the frequency of opioid-related deaths in a US city.

Our contributions are as follows: Our team scraped the data from seven cryptomarkets and performed a comprehensive analysis of the data, including an analysis of revenue and vendor behavior. Additionally, we implement a probabilistic model to predict the vendor, study drug co-occurrence, cluster transactions by buyer, and discuss the potential for vendor de-anonymization.

II. RELATED WORK

Soska and Christin conducted a comprehensive analysis of several dark web marketplaces from July 2013 to July 2015 [7], which followed Christin’s original analysis of the Silk Road marketplace [9]. The authors of [7] found that a small set of vendors generated a disproportionate amount of revenue, and that there was a consistent daily demand for drugs; that is, market disruptions such as exit scams and law enforcement interventions only temporarily reduce sales on the dark web.

Other academic researchers, governments, and private companies have conducted similar studies, conducting exploratory analyses on various illicit marketplaces [8], [10], [11]. Past research has explored how cryptomarkets differ from traditional illicit markets [12]–[14]. Several authors have explored the importance of trust and vendor ratings, and how these factors can lead to a market network that is divided into many sub-networks [15]–[17]. Others have attempted to determine whether most sales represent wholesale or retail purchases [18], analyzed the effect of law enforcement interventions on vendor activity [19], and sought to find a correlation between availability of opioids on the dark web with overdose deaths in Ohio [20].

Several researchers, most notably [21], have observed that a large proportion of crimes are often committed by a small proportion of the population. Related to this, the law of crime concentration was proposed in 2014 to describe the occurrences of physical crimes in a geographic context [22], [23]. Notably, these studies do not include cryptomarkets in their analysis.

In the cryptocurrency sphere, several researchers have studied de-anonymization of bitcoin transactions [24]–[26].

Our work contributes to the still-developing literature on dark web marketplaces by characterizing the structure of a thriving dark web marketplace over the course of 10 months in 2018 and 2019, and making connections to the law of crime concentration [22]. In addition to descriptive analyses, we implement a probabilistic model to predict the vendor involved in a given transaction, analyze the value of marijuana over time, compare co-purchase of drugs on the Dream Market to opioid overdose deaths in Indianapolis, explore clustering techniques to group together transactions by the same buyer, and discuss the potential for vendor de-anonymization.

III. DREAM MARKET OVERVIEW

The majority of our data (about 75%) originates from the Dream market, a dark web market predominantly focused on the sale of drugs. The market formed in late 2013 and continued operating until its shutdown on April 30th, 2019. It became one of the most popular markets with over 100,000 product listings available [27]. While the Dream market focused on the sale of drugs, it also offered digital goods such as e-books, credit cards, bank accounts, hacked accounts, software, and entertainment accounts.

A. Privacy and Security

The Dream market, similar to other other dark web markets, shielded users' identities through use of the Tor Browser and Pretty Good Privacy (PGP) keys. Dream also sought to protect its users from hackers and scams by implementing two-factor authentication, withdrawal pin codes, escrow service, and a rating system [27].

1) *Privacy*: Dark web markets can only be accessed using anonymous browsers such as Tor or the Invisible Internet Project (I2P). Tor uses onion routing to encrypt transmitted

information and globally reroute the IP address of the computer accessing the deep web [14]. In the case of peer-to-peer messaging, PGP encrypts messages and creates a pair of files called keys to access and decrypt those messages. One key is intended to be public and the other to be private. The owner of the keys shares public key in a public place, and anyone that wants to message the owner will use this public key to encrypt their message. The owner will receive the encrypted file and use their private key to decrypt and read the message.

2) *Security*: The Dream market implemented two-factor authentication using the standard password and PGP private keys. An encrypted text is presented which must be decrypted with the user's private key before access to the account is granted. All accounts have the option to set-up a withdrawal PIN. For accounts with such a PIN, funds can't be withdrawn by password alone, providing more security from theft. Escrow is a payment method in which a vendor does not receive the buyer's money for a product until the buyer confirms its arrival. The Dream market allows the option of escrow to prevent buyers from getting scammed, as can happen when finalizing early. Finalizing early means the vendor receives the buyer's payment before the product has arrived. This method allows for easy scamming. Many members advise not to finalize early unless the vendor is well trusted. To know if a vendor is reliable, the Dream market has a rating and feedback system. Buyers can give a vendor a star rating from one to five and leave a comment. This system separates the top-quality vendors from the ones of poor quality.

B. Shutdown

On April 30th, 2019, the Dream market announced it was shutting down and moving to a new onion address. Some speculate law enforcement have taken over the site. According to an article on ZDNet, a moderator posted on a dark web social network site that the market is moving due to many recent DDoS attacks and a \$400,000 ransom request [28]. DDoS, or distributed denial-of-service, floods a network or service with traffic causing an overflow and preventing access to regular traffic. The moderator asks readers to be patient for the next market that will open soon after Dream's closure.

IV. DATA

Our team developed unique scrapers in Python for 7 cryptomarkets shown in Table I. After scraping, the data is stored in a MySQL database [29]. The scraper primarily targets publicly accessible web pages, such as product rating pages. These pages provide a short description of the product along with ratings and comments from buyers. The information is stored in the database as a product rating table. The table contains several features, such as transaction ID, vendor ID, vendor name, product description, bitcoin or US dollar amount, feedback, product rating, and date. The data was collected from the 7 marketplaces over the course of 10 months, from April 2018 to January 2019. This study mainly focuses on the 92,388 product reviews from the Dream market, which account for 77% of our total data. However, data from the other markets

TABLE I
SEVEN CRYPTOMARKETS

| Cryptomarkets | #Listings | Percentage |
|----------------------|-----------|------------|
| <i>Dream</i> | 92,388 | 77.03% |
| <i>Wall Street</i> | 8,902 | 7.42% |
| <i>Empire</i> | 9,947 | 8.29% |
| <i>Silk Road 3.1</i> | 718 | 0.60% |
| <i>Berlusconi</i> | 6,007 | 5.01% |
| <i>TochkaPoint</i> | 1,442 | 1.20% |
| <i>Valhalla</i> | 530 | 0.44% |

is used to analyze buyer behavior (e.g. buyer clustering and co-purchase of drugs), because buyer names are fully obscured in the Dream market.

V. METHODOLOGY

A. Product Categorization

In order to analyze the type of products sold on the market, we first categorize the reviews by the type of drug. We utilize sklearn’s tokenizer function to split the product title for each review into tokens. We then compare these tokens with words and phrases in a custom ‘dictionary’ which contains a set of synonyms, slang words, and associated phrases for each drug type. For example, the category ‘marijuana’ includes the word “weed” but also many specific strains of marijuana such as “Girl Scout Cookies.” The benzodiazepine category contains many specific prescription pills such as Valium and Xanax. Some drugs that would not make up a large percentage by themselves are grouped in a broader category. This is the case with codeine, morphine, and hydrocodone, which are all classified as “opioid”. However, oxycodone and heroin are placed in their own categories because each constitutes a relatively large number of listings. We also created a new category, ‘not drugs’, to represent products not related to drugs.

Several of the product titles were ambiguous, or proffered a second drug as a ‘bonus.’ For ambiguous product titles, we remove the corresponding reviews from the data set; and for product titles offering a ‘bonus’ or ‘sample’ of a second drug type, we label the review using the primary drug type. The remaining uncategorized listings (<1% of all listings) are removed from the data set.

B. Purity/Quantity

Using the tokenized titles, we also extract the purity and quantity of drug for each listing. Our motivation is to better understand drug rates and vendor behavior. We use the quantity and purity information to study the change in price of marijuana over time (described below). Notably, many product titles do not report quantity, and many more do not report purity (purity is most relevant for specific drug types, such as cocaine); therefore, the data on quantity and purity of drugs sold is relatively sparse.

C. Filtering

Next, we filter out several outlier listings. First, we remove listings with timestamps from 1969-1970 (2,402 listings), 2017

(25 listings), and January 2018 – March 2018 (62 listings). Thus, the interval of time for our filtered data set is April 2018 – January 2019 (89,842 listings).

VI. ANALYSIS

Dark web markets do not publicly post records of transactions, complicating analysis of vendor and buyer activity. However, many of the markets strongly encourage consumers to post product reviews in order to build trust between buyers and reputable vendors. Therefore, we use product reviews as proxies for transactions (i.e. each product review is counted as one transaction), similar to previous analyses of dark web markets [9]. However, while reviews are encouraged on the Dream market, they are not mandatory, and there is risk of under-counting the number of transactions.

Temporal analysis of sales on the markets is further complicated because of the uncertainty associated with scraping on the dark web. Even if reviews are a perfect proxy for transactions, observing and assigning a timestamp to reviews is often challenging. The number of reviews in the dataset obtained by scraping increases exponentially toward the end of 2018, which likely does not reflect the actual market activity in the time period: the Dream market was founded in late 2013, and there is no indication outside of the scraped dataset that activity furiously increased in late 2018. To mitigate the effect of potential temporal biases in the data, we focus on the proportion of transactions (between vendors, or drugs, for example) rather than the absolute quantity. However, we do offer estimates for absolute revenue and number of transactions to illustrate the scale of the markets; these estimates should be viewed as a lower bound on the actual figures for the reasons stated above.

A. Goods and Services

In our analyses of the goods being sold in the Dream market, we explore what drugs are available in the market as well as the proportions of each drug sold over time. Based on transaction volume, the most popular drugs on the Dream market are marijuana, cocaine, and benzodiazepines. These three categories account for slightly less than half of all transactions (Figure 1). There are a wide variety of drugs, including methamphetamine, heroin, performance-enhancers, and prescription drugs. Notably, the proportion of transactions involving opioids (e.g. opioid, oxycodone, and fentanyl categories) is relatively small.

Figure 2 illustrates the proportion of drug sales over time. The proportion of sales is relatively constant over time (unlike the absolute sales numbers), providing evidence that our data accurately reflects the proportion of drugs for actual transactions.

We also examine the quantities at which these drugs are sold to determine if purchases are primarily wholesale or retail. Researchers have speculated that certain drugs may be more often purchased in larger quantities [18], so we create boxplots showing the mass of purchase for different drug types (Figure 3). There are several drugs, such as benzodiazepines

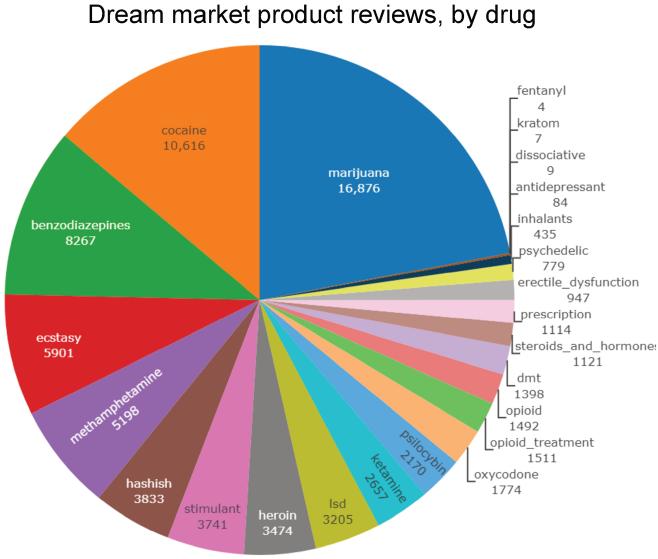


Fig. 1. Total drug distribution in Dream market.

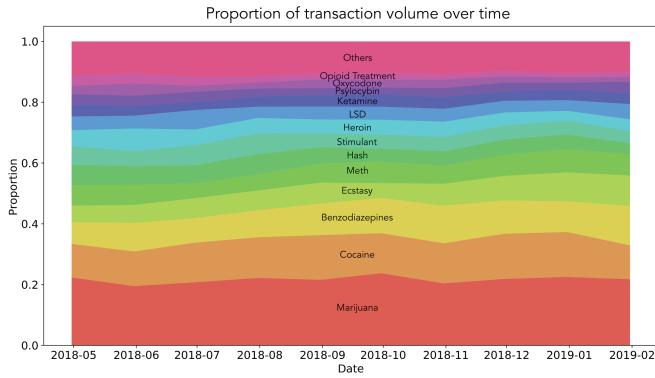


Fig. 2. Proportion of drugs sold over time.

and oxycodone, which are generally sold in smaller quantities. However, for most of the drugs (especially marijuana, abbreviated as ‘weed’), the spread is quite large (note the logarithmic y-axis scale), indicating that there is a mix of small transactions and larger bulk purchases for almost all drug types.

B. Vendor Behavior

To determine if certain combinations of drugs are commonly sold by a given vendor, and to assess the degree of specialization among vendors, we created a graph showing the drugs sold by high-volume vendors (Figure 4). An edge between a vendor (blue dot) and a drug (red dot) means the vendor has conducted at least 50 transactions involving the drug. Most of the vendor nodes have only one or two edges, indicating that most high-volume vendors tend to specialize in one or two types of drugs.

We also attempt to quantify the amount of time for which vendors stayed active in the dark web markets. We obtained the Kaplan-Meier estimate for vendor survivability using the

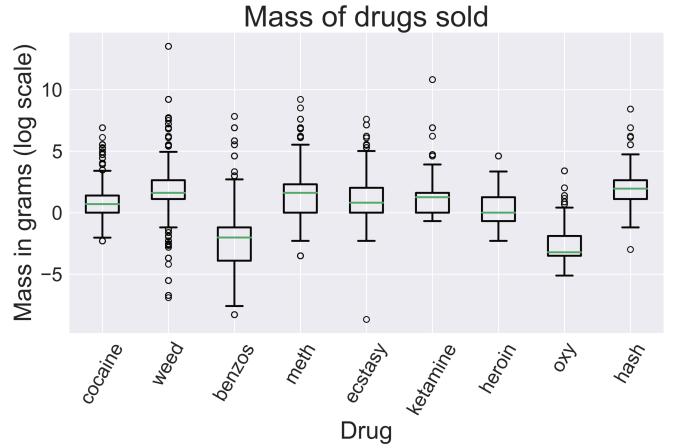


Fig. 3. Mass of products drugs sold. Note the log scale for the y-axis.

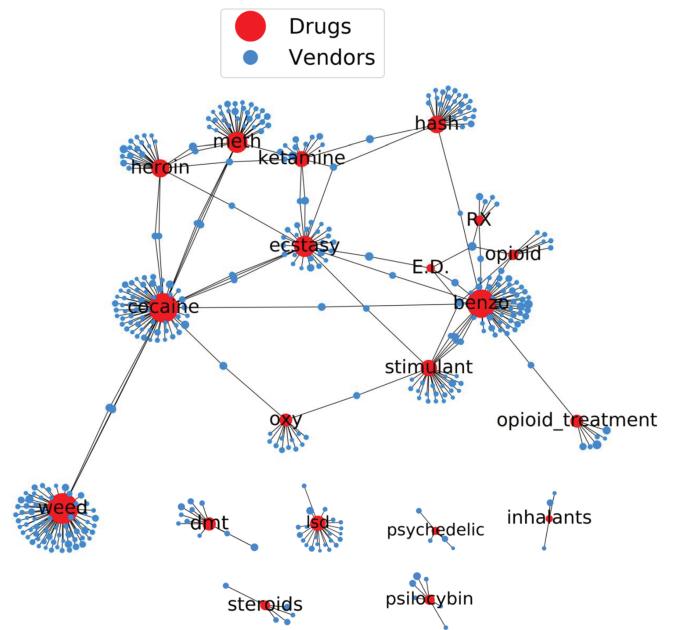


Fig. 4. A connection between a vendor and a drug indicates the vendor having at least 50 sales of the drug.

‘Lifelines’ Python package. The resulting curve in Figure 5 shows the probability that a vendor will continue to stay active as a function of time. This result echoes Christin’s findings that most vendors on the Silk Road market did not remain active for long [9]. According to our data, there is only about a 50 percent chance that a vendor will continue to remain active on the market after 100 days from his or her first transaction. This means that many vendors in the Dream market tend to sell their goods and exit the market shortly thereafter.

C. Revenue Analysis

We also explore the distribution of revenue among vendors. Figure 6 shows a plot of the cumulative revenue across all vendors. The plot illustrates that revenue is heavily concentrated among a small subset of vendors, rather than being evenly

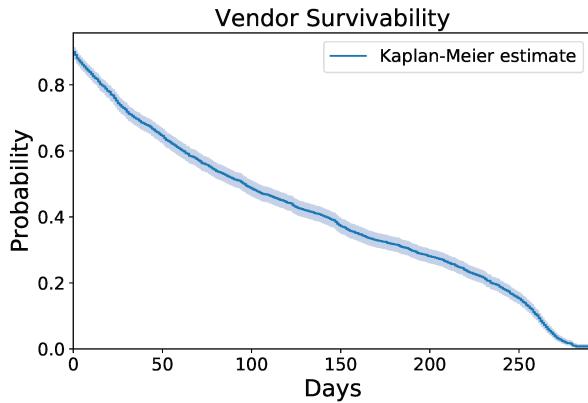


Fig. 5. Probability that a vendor will continue to stay active in market as a function of time.

distributed. This concentration is even more pronounced when considering a specific type of drug (Figure 7), which makes sense given that most vendors specialize in a particular drug, as depicted in Figure 4. While Figure 5 shows that the rate of vendor turnover is high, Figures 6 and 7 make it apparent that the market is dominated by a few high-revenue vendors.

The observation that few vendors account for the majority of revenue invites comparison to Wolfgang's landmark 1972 study of Philadelphia youths, which found that 18% of juvenile offenders were responsible for over 50% of the crimes committed by those in the study [21]. More recently, David Weisburd proposed a law of crime concentration to describe the geographic distribution of crimes. The law states that *for a defined measure of crime at a specific microgeographic unit, the concentration of crime will fall within a narrow bandwidth of percentages for a defined cumulative proportion of crime* [22]. Notably, the law of crime concentration proposed by Weisburd addresses the distribution of crime in a geographical context, not amongst a set of people; therefore, the Wolfgang study [21] may have more relevance in this context. However, Weisburd's proposition that *the concentration of crime will fall within a narrow bandwidth of percentages for a defined cumulative proportion of crime* may also apply to vendor sales on dark web markets.

While it is straightforward to compute the Lorenz curve and Gini index for revenue distributed among vendors, it is more challenging to quantitatively assess the 'concentration' of the market, especially when differentiating between different drugs. For example, suppose there are 2 vendors who have at least one sale of Drug A; the more prolific vendor earns \$90, and the other vendor earns \$10. There is only one vendor of Drug B, and the vendor earns \$35 in revenue. Then is revenue more concentrated for Drug A, or for Drug B? There is only one vendor who earns revenue for Drug B, so it would appear that sales for Drug B are more concentrated (i.e., the vendor of Drug B earns 100% of revenue for Drug B). However, the prolific vendor of Drug A earns more than twice the vendor of Drug B, and accounts for 90% of total revenue for the

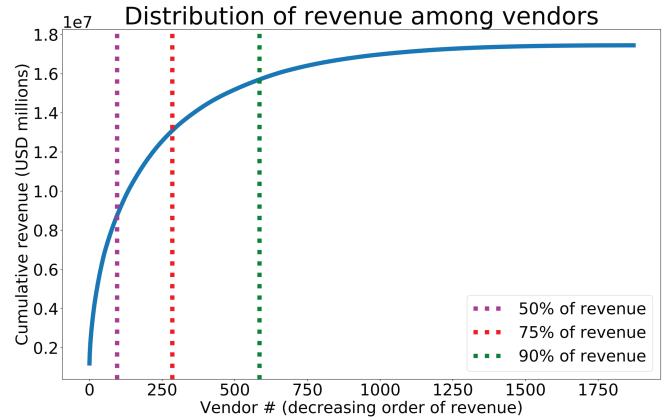


Fig. 6. Just 5% of vendors account for 50% of total revenue. The Lorenz curve has a Gini coefficient of 0.778.

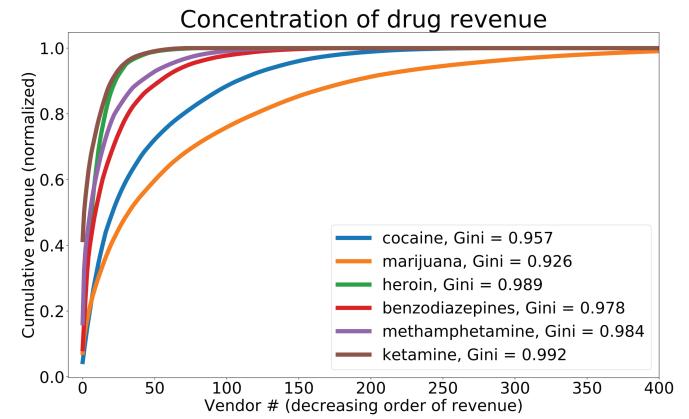


Fig. 7. Revenue is even more concentrated for each drug. In the case of ketamine, the top 2 sellers account for 50% of revenue.

drug. This example highlights a potential weakness of the Gini index in evaluating relative concentrations of revenue among vendors.

The authors of [23] suggest a standardized method for measuring crime concentration. The method that they suggest is not directly applicable to distribution of drug revenue among vendors, because they study the discrete number of crimes distributed among a set of fixed geographic units, while in our case revenue is continuous and distributed among a variable number of vendors. However, we believe that when the Gini index is taken into consideration along with overall revenue, it provides a useful representation of the relative concentrations of different drugs.

Figure 8 displays the revenue for vendors that generated more than \$25,000 during the 10 month period. While 18 vendors grossed more than \$100,000, the revenues appear to be less than observed in other works. For example, Soska and Christin found 35 vendors in various dark web markets selling over \$1 million over a two year period in various markets. The reason for the discrepancy may be that in this study we only analyzed one market, and do not track sales for vendors across

Revenue per vendor (USD)

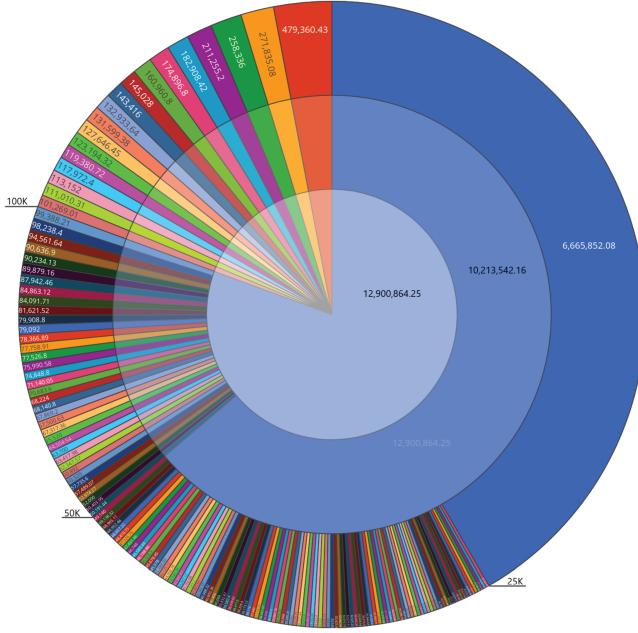


Fig. 8. Revenue for the vendors in \$100,000, \$50,000, and \$25,000 brackets

several different markets.

The distribution of revenue among drugs, shown in Figure 9 does not exactly mirror the distribution of transactions among drugs. While marijuana accounts for the majority of transactions, the revenue is the second-highest among all drugs, falling behind cocaine. This difference among transaction count and revenue could result from a difference in price; another explanation might be that certain drugs are more often purchased wholesale. The authors of [18] speculate that drugs with more complicated global distribution networks, such as cocaine and heroin, would be more often purchased wholesale than drugs such as marijuana, which can be more easily produced on a smaller scale and by a greater number of producers.

VII. DISCUSSION

A. Modeling

We seek to determine whether the vendor of a particular transaction can be inferred based on several basic features: the drug sold and the price. Because the market was dominated by a relatively small set of vendors, we do not include transactions corresponding to low-volume vendors, defined to be vendors with fewer than 100 transactions.

As a baseline, we use conditional probability to predict the vendor from one of the two possible features, drug and price (in order to use a continuous variable as a feature in our probabilistic models, we binned the price of a transaction into a given price interval. Both \$20 and \$1 bins are considered). The test prediction is based on the relative frequency of observations in the training set. For each test sample, the

Revenue per drug (USD)

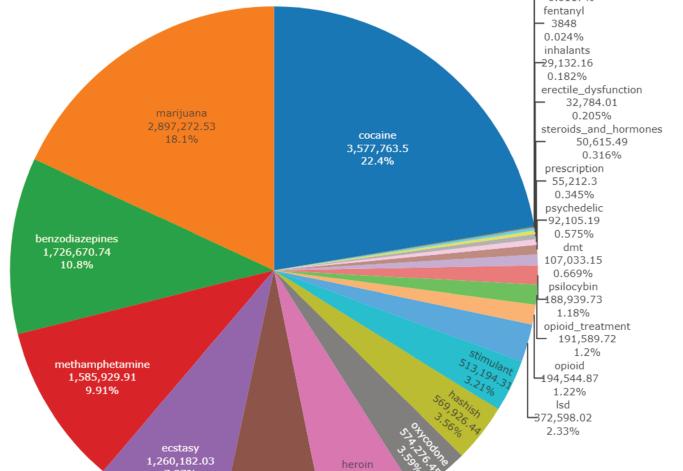


Fig. 9. Percentage of revenue for each drug category.

TABLE II
VENDOR PREDICTION ACCURACY USING PROBABILISTIC MODELS. IN THE Feature ROW, D REPRESENTS DRUG, P_{\$N} REPRESENTS PRICE WITH \$N BINS.

| Model | Conditional Probability | | | Naïve-Bayes | |
|---------|-------------------------|-------------------|------------------|-----------------------|----------------------|
| | D | P _{\$20} | P _{\$1} | D & P _{\$20} | D & P _{\$1} |
| Feature | | | | | |
| Top 1 | .121 | .072 | .240 | .222 | .488 |
| Top 5 | .430 | .192 | .514 | .602 | .660 |
| Top 10 | .652 | .284 | .583 | .807 | .767 |

probability for each vendor is computed, and the vendor with highest probability is selected as the prediction.

We then use a Naïve-Bayes classifier to predict a given vendor from both drug and price. The Naïve-Bayes classifier assumes that features are conditionally independent, given the class. This may be problematic for our problem, because certain drugs may be associated with higher costs.

The prediction accuracy of each model is shown in Table II. Despite the potential violation of conditional independence, the vendor associated with a particular listing can be inferred with almost 50% accuracy using the Naïve-Bayes model with drug and \$1 price bin as predictors.

The time of day at which a product review was posted may also serve as a useful predictor for vendor. However, the timestamps for the product reviews collected in our data set have variable granularity, ranging from hours to months. The variability results from the Dream market not indicating the exact time that a product review was posted, but instead an estimate of how long ago the review was posted (for example, ‘8 days ago’ or ‘2 months ago’). This issue of varying granularity among transactions was addressed by Soska and Christin [7]. To simulate time’s suitability as a predictor, we

artificially generate data assuming that a Poisson process can be used to describe the posting of reviews for each vendor. Similar to the price feature, we bin the timestamps into intervals experimenting with several interval sizes. Applying the conditional probability model on the artificially generated data yields no benefit over random guessing. We also apply a Hidden Markov Model (HMM) on the artificially generated data, with the vendor as the hidden state and drug as the observed state. This model also does not perform better than random guessing.

B. Drug Co-Occurrence

We also explore the ‘co-purchase’ of drugs by buyers on the dark web. Co-purchase refers to the same buyer purchasing two different drugs, and is important to examine because of the increased potential for overdose when certain drug types are mixed.

This analysis is complicated by the partial or full censoring of buyer names by many of the markets. To simplify the task, we assume that in markets with partially-censored buyer names, buyers with matching first and last characters are the same (we further explore the issue of grouping transactions by buyer in Section VII D, *Clustering Product Reviews*). We exclude the Dream market from this analysis, because buyer names in Dream are fully censored. Applying this assumption, we analyze the reviews from 4,400 buyers.

We compare our findings with a data set of drug overdose deaths in Indianapolis to determine if there is a correlation between buyer behaviors and drug-related deaths. The results are shown in Table III. The most common pairings of drugs purchased among dark web buyers are compared to the drug pairings which are the most common cause of drug overdose deaths in Indianapolis. The results indicate that dark web users more commonly purchase recreational drugs, as opposed to the highly addictive opioids which cause the majority of overdose deaths.

TABLE III
COMPARISON OF DRUG CO-PURCHASING ON DARK WEB TO
CO-OCCURRENCE IN OVERDOSE DEATHS

| Dark web pairs | Count | Percent | Overdose pairs | Count | Percent |
|--------------------|-------|---------|------------------|-------|---------|
| cocaine, marijuana | 256 | 3.49% | morphine, opioid | 723 | 2.09% |
| ecstasy, marijuana | 221 | 3.01% | benzos, opioid | 606 | 1.75% |
| cocaine, ecstasy | 197 | 2.69% | 6-mam, opioid | 565 | 1.63% |
| marijuana, meth | 166 | 2.27% | 6-mam, morphine | 562 | 1.62% |
| hashish, marijuana | 140 | 1.91% | codeine, opioid | 533 | 1.54% |

There are several pairs found in both datasets (Dream market and Indianapolis overdoses). However, none of these pairs account for more than 1% of either dataset, supporting the conclusion that cryptomarkets are not a primary source for users of dangerous drug combinations.

C. Price of Marijuana

In light of the decriminalization of marijuana across several states and countries in recent years, we explore if increased

accessibility to marijuana has influenced the value of the drug on the dark web. We include data across all seven markets (totaling 16,284 listings) and analyze how the price of marijuana products evolved over the 10-month period.

We sift through each unique listing and manually extract information about the product, including the quantity, mass, THC amount, and whether it was a vape product or edible. We then use the mass and quantity to compute the price per gram for each product and find that there were about 1,400 kg of marijuana products sold, generating approximately 2.8 million U.S. dollars over the course of the 10 months.

To better understand how the prices of these products are changing, we filter out listings that did not have a duration longer than 200 days. In Figure 10, we see that the average price for marijuana decreased from about \$22 to \$15 per gram from May 2018 to February 2019, though it did peak at \$23 per gram in August 2018.

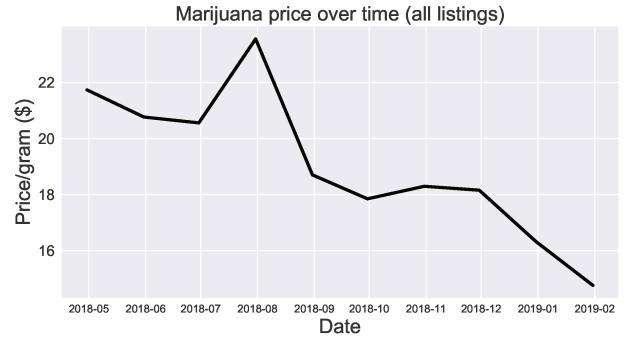


Fig. 10. Monthly average price per gram of all marijuana listings.

Next, we look at several ‘long-duration’ marijuana listings (listings which were posted on a given market for longer than 200 days) to determine if the observed price decrease over time reflected a change in the value of marijuana, as opposed to an increase in quantity of low-priced products (or decrease in quantity of high-value products). Several of the listings do not change in price over time; however, those that do change follow a similar trend to the overall price of marijuana. The price over time for four of these listings are shown in Figure 11. Similar to the average price of all listings, shown in Figure 10, prices increased from March to April and significantly declined toward the end of the year.

The overall decline in the price of marijuana products could result from marijuana being legalized and more accessible for consumers, or the trends for supply and demand of marijuana products online and offline.

D. Clustering product reviews

For many dark web marketplaces, the names of buyers posting a review on the site are obscured. In some cases, the first and last letter are given (but not the length of the name), while in other markets, such as the Dream market, no information about the name is given. One such instance of partially obscured names is shown in Figure 12.

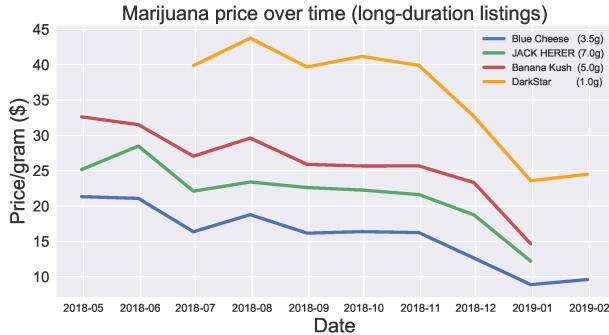


Fig. 11. Monthly average price for several ‘long-duration’ listings, each of which was posted for at least 200 days.

| vendor_name | buyer_name | date | USD | product_title |
|---------------|------------|----------------|-------|--|
| GrimReefer420 | s***9 | 20181231080300 | 110.0 | 1 Ounce - 110 USD - JAGER (PURPLE HINDU KUSH) |
| DeadHeadFred | s***9 | 20190619065900 | 170.0 | 1 Piece - 170 USD - LSD 50 Hits 125µg GDF W... |
| DrFrosty | s***9 | 20181219065100 | 80.0 | 1 Piece - 80 USD - 3.5 grams of shard |

Fig. 12. In several of the markets, all buyer names were replaced with the first letter, three asterisks, and last letter.

To extend our analysis of markets to the demand side, we apply hierarchical agglomerative clustering to estimate which product reviews were posted by the same buyer. One of the markets we examine includes the full buyer names, so we use this market to evaluate our clustering technique. The data subset consists of 593 reviews from 400 unique buyers after reviews for non-drug products are removed. The possible features to cluster on include drug, vendor name, price, date, price, product title, and buyer comment (several of these features are shown in Figure 12).

We convert the date from a categorical variable to a numeric variable, and also extract two new features to represent the day

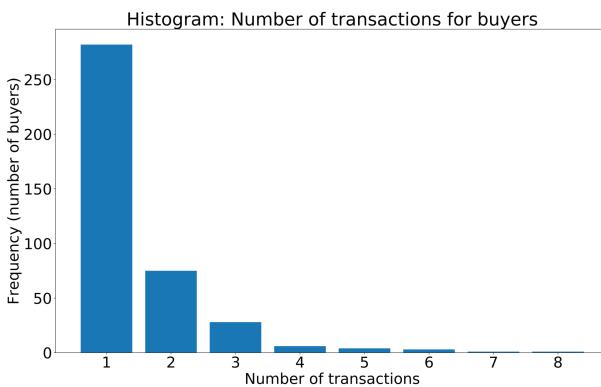


Fig. 13. The number of transactions is not normally distributed among buyers.

TABLE IV
NUMBER OF UNIQUE BUYERS FOR EACH FIRST/LAST CHARACTER PAIR

| Unique buyers in set | 1 | 2 | 3 |
|----------------------|------|------|------|
| Proportion of sets | .842 | .140 | .018 |

of the week (using sine and cosine to represent the cyclical nature of this feature; e.g. Sunday should be closer to Saturday and Monday than it is to Wednesday). The purpose of this feature is to capture purchasing patterns; for example, a certain buyer may purchase a product at a similar time each week.

In agglomerative clustering, each sample is placed in a singleton cluster to start. Then, a distance metric is selected, and the distance between each pair of samples is computed. Depending on a predefined distance threshold, cluster pairs with distances lower than the given threshold are merged together (alternatively, the number of desired clusters is specified, and clusters are merged based on distance until the number is satisfied) [30]. Several of the features are categorical (specifically, drug and vendor name), making it challenging to compute the distance between pairs of reviews using traditional distance metrics such as Euclidean. Instead, we use the Gower similarity coefficient, which allows for the comparison of samples with mixed numeric and categorical features [31], adapting a Python implementation by Marcelo Beckmann [32].

First, the clustering technique is applied *without* any assumed knowledge of buyer name (Method A in Table V), to provide a baseline. Three features are used to perform the clustering: vendor name, drug, and timestamp; this combination of features was found to give the best result, as evaluated by purity, adjusted mutual information score (AMI), and Rand index.

Next, clustering is performed *only* using the first and last letter of the buyer name (Method B in Table V), to provide a second baseline. Product reviews are grouped together by the first and last letter of the buyers’ usernames. All clustering metrics improved, with AMI and Rand most notably increasing.

The preceding methods are then combined: reviews are grouped by first and last letter of buyer name, then agglomerative clustering is performed on each subset of reviews using vendor, drug, and date as features (Method C in Table V). This combined approach slightly improves the AMI and Rand measures, while increasing the purity from about 90% to 98%.

To further increase the performance of the clustering, we perform a frequency analysis on a database of known internet usernames. A similar technique was employed by Décarie-Hétu and Doré [17] to analyze buyer loyalty in another dark web market data set. To incorporate frequency analysis, we compute the frequency of a given first/last letter pair from the database of 10 million usernames, and assign a weight to each unique first/last pair (there are 4,000 unique first/last letter pairs in the database). We then partition reviews into sets based on first/last letter pairs, as in Method C, and use the weight for the first/last letter pair to adjust the distance threshold for merging clusters within the set. We experiment with two different weightings: first, the weight is the inverse of the frequency of the username in the database (Method D in Table V); second, the weight is computed as in [17] (Method E in Table V).

The clustering methods are summarized below:

A Cluster using vendor, drug, and timestamp features

TABLE V
COMPARISON OF DIFFERENT CLUSTERING METHODS.

| Metric | Clustering method | | | | |
|----------|-------------------|------|-------------|------|-------------|
| | A | B | C | D | E |
| Purity | .833 | .898 | .981 | .974 | .961 |
| AMI | .391 | .857 | .858 | .883 | .891 |
| Rand | .317 | .840 | .860 | .869 | .885 |
| Clusters | 420 | 335 | 420 | 404 | 389 |

- B** Group by first/last letter of buyer name
- C** Group by first/last letter, then cluster using method **A**
- D** Apply method **C** with weights of $1/f_n$ for agglomerative clustering, where f_n is frequency of given first/last letter pair
- E** Apply method **C** with weights of $1 - \frac{f_n}{f_{max}}$ for agglomerative clustering, where f_n is frequency of given first/last letter pair, and f_{max} is maximum frequency of first/last letter pair in username database

As shown in Table V, Methods C, D, and E all outperform Method B. Method C, which takes into account the first and last letter as well as three features (drug, vendor, and date) achieves the highest purity, while Methods D and E, which take into account username frequency, achieve purities comparable to C as well as higher AMI/Rand scores.

A high purity indicates that within each identified cluster, there is little variation among the actual buyer name (in the trivial case, assigning each sample to a singleton cluster yields a purity of 1). The high purity can be thought of as analogous to high precision: for each cluster identified, we can be fairly confident that most of the reviews are from the same buyer. The high purity may be desirable for analyses of drug co-purchasing among buyers; in other words, knowing that each cluster is likely to contain reviews from a single buyer allows us to more confidently identify which drugs are commonly purchased together.

Notably, the performance characteristics of our clustering method would likely change when it were applied to a larger data set, due to the increased probability for multiple buyers for each first/last character combination. If there are 76 possible characters, there are 5,776 different possible first/last character combinations [17]. In the labeled data set selected for evaluating the clustering method, there were 400 buyers and 335 first/last letter combinations. As shown in Table IV, 84% of first/last letter groupings represented a single buyer. This percentage should decrease in data sets with a larger number of buyers, as increasing numbers of buyer names are binned into one of the 5,776 possible first/last character combinations.

E. De-anonymization

The pseudo-anonymous characteristics of bitcoin have been explored by various researchers [24]–[26]. Bitcoin de-anonymization research efforts have focused on clustering addresses of major exchanges and entities, and have resulted in sites such as WalletExplorer and the company Chainalysis. However, less work exists on the de-anonymization of individual dark web users.

As part of an exploratory analysis, we gathered Bitcoin blockchain data from January 1, 2019 to January 7, 2019. Using the product cost (in BTC) listed in the product reviews in our Dream Market data, we searched for the amount in the blockchain. Specifically we identified 17 product reviews for the vendor ‘DDUK-NDD’ in the first week of January, each of which had the same product and price, 0.01265 BTC. Searching for this amount as an output in the blockchain, we identified 31 matching transactions. While this may at first appear promising, this amount may not be the amount that the market or vendor actually receives (because of market commission, blockchain fees, or mixing fees). Therefore, we conducted another search for transactions with outputs in the interval [0.01264, 0.01266]; this search yielded 1,215 matching transactions.

This simple attempt at de-anonymization is complicated in several other ways. As explained previously, the granularity of the timestamps in our data are variable. However, because we used reviews from the most recent month collected, the reviews have a granularity of at least a day. While the product reviews’ dates are likely accurate to within a day, each of the 17 product reviews may have occurred several days after the corresponding transaction.

Furthermore, the purchase of a product on the Dream market is more complicated than a buyer sending bitcoin to a vendor. Buyers deposit money into a bitcoin wallet on the Dream market. Upon making a purchase, money is moved into escrow which is released to the vendor upon confirmation of the buyer that she has satisfactorily received the product she ordered. It is likely that at some point in the process of traveling from buyer to seller, the payment also travels through a bitcoin mixer.

Despite these challenges, de-anonymization attempts could be improved by taking into account transaction patterns. As discussed earlier, sales on the Dream market were not evenly distributed among vendors; instead, a few vendors dominated each drug category, earning orders of magnitude more in revenue than most others. Because these vendors have a high volume of sales, there is potentially an increased possibility of identifying patterns in the public blockchain ledger. For example, a given vendor may have a routine of transferring bitcoins through a mixer and to a currency exchange. Searching for a rough price match narrows the search space, and makes the challenge of identifying patterns less daunting.

VIII. CONCLUSIONS

In this work we conducted a comprehensive analysis of the dark web Dream marketplace from April 2018 to January 2019. Our research builds on existing work studying dark web markets by investigating the unit price of drugs over time, comparing opioid overdose deaths in an American city to corresponding sales on the dark web, and by exploring new methods of clustering anonymous transactions.

The Dream market was dominated by a small subset of vendors who generally specialized in one type of drug. Despite a high turnover rate for vendors overall, the cohort of high-volume vendors stabilized the market, which was shown to

have a consistent distribution of drug sales over time. This finding supports [7], in which the authors found a small group of vendors accounting for a large proportion of sales.

Our analysis of the unit price of marijuana reveals a decrease over time in the price of products sold by high-volume vendors and could possibly reflect the increasing availability of marijuana due to relaxation of state laws.

Surprisingly, the Dream market dataset we analyzed contained relatively few opioid transactions. While opioids were still available for purchase on the Dream market, the highest grossing opiates, heroin and oxycodone, ranked 7th and 8th respectively in overall revenue among all drugs. Among co-purchases (i.e. the same buyer purchasing two different drugs), relatively few appeared to match the fatal combinations catalogued in a data set of overdose deaths in Indianapolis. This discrepancy might be partially explained by the banning of fentanyl sales on the Dream market.

Our work explores the potential to identify multiple purchases by the same buyer by clustering partially and fully anonymous transactions, and suggests a new method to use username frequency to improve clustering performance.

While our study provides a thorough characterization of several dark web markets, we do not quantify the privacy risk for vendors (e.g. of having their identities revealed). Future work could attempt to link dark web transactions to transaction flows in the Bitcoin blockchain, as discussed in the *De-anonymization* section.

ACKNOWLEDGMENTS

This research was supported by National Science Foundation Dr. Mohler and Dr. Hasan's NSF grant, numbered REU-1659488.

REFERENCES

- [1] J. Frankenfield, "Darknet market (cryptomarket)," February 2018.
- [2] G. Weimann, "Terrorist migration to the dark web," *Perspectives on Terrorism*, vol. 10, p. 40, 2016.
- [3] J. Bearman and T. Hanuka, "The untold story of silk road, part 1," *Wired Magazine*, 2015.
- [4] A. van Wirdum, "Following wave of shutdowns, remaining darknet markets fill the void (again)," *Bitcoin Magazine*, June 2019.
- [5] Europol, "Operation onymous," November 2014.
- [6] A. Greenberg, "Feds dismantled the dark-web drug trade—but it's already rebuilding," 2019.
- [7] K. Soska and N. Christin, "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem," in *Proceedings of the 24th USENIX Security Symposium*, pp. 33–48, 2015.
- [8] Chainalysis, "Crypto crime report: Decoding increasingly sophisticated hacks, darknetmarkets, and scams," tech. rep., Chainanalysis, 2019.
- [9] N. Christin, "Traveling the silk road: A measurement analysis of a large anonymous online marketplace," in *Proceedings of the 22nd international conference on World Wide Web*, ACM, pp. 213–224, 2013.
- [10] J. Broséus *et al.*, "Studying illicit drug trafficking on darknet markets: Structure and organisation from a canadian perspective.," *Forensic Science International*, vol. 264, pp. 7–14, July 2016.
- [11] J. van Buskirk *et al.*, "Drugs and the internet, issue 4," tech. rep., Sydney: National Drug and Alcohol Research Centre, March 2015.
- [12] J. Martin, "Lost on the Silk Road: online drug distribution and the online 'cryptomarket,'" *Criminology & Criminal Justice*, vol. 14(3), pp. 351–367, 2014.
- [13] J. Aldridge and D. Décaray-Hétu, "Not an 'ebay for drugs': The cryptomarket 'silk road' as a paradigm shifting criminal innovation," May 2014. Available at SSRN 2436643.
- [14] J. Aldridge and D. Décaray-Hétu, "Cryptomarkets and the future of illicit drug markets (the internet and drug markets)," Tech. Rep. Insights: 21, European Monitoring Center for Drugs and Drug Addiction, 2016.
- [15] S. W. Duxbury and D. L. Haynie, "The network structure of opioid distribution on a darknet cryptomarket," *Journal of Quantitative Criminology*, vol. 34, pp. 921–941, 2018.
- [16] L. Norbutas, "Offline constraints in online drug marketplaces: An exploratory analysis of a cryptomarket trade network," *International Journal of Drug Policy*, vol. 56, pp. 92–100, 2018.
- [17] D. Decary-Hétu and O. Quessy-Dore, "Are repeat buyers in cryptomarkets loyal customers? repeat business between dyads of cryptomarket vendors and users," *American Behavioral Scientist*, vol. 61, pp. 1341–1357, 2017.
- [18] J. Aldridge and D. Décaray-Hétu, "Hidden wholesale: The drug diffusing capacity of online drug cryptomarkets," *International Journal of Drug Policy*, vol. 35, pp. 7–15, September 2016.
- [19] R. van Wegberg and T. Verburgh, "Lost in the dream? measuring the effects of operation bayonet on vendors migrating to dream market," in *Proceedings of the Evolution of the Darknet Workshop at the Web Science Conference (WebSci 18)*, pp. 1–5, 2018.
- [20] U. Lokalo *et al.*, "Global trends, local harms: availability of fentanyl-type drugs on the dark web and accidental overdoses in ohio," *Computational and Mathematical Organization Theory*, vol. 25, no. 1, pp. 48–59, 2019.
- [21] M. Wolfgang, R. Figlio, and T. Sellin, *Delinquency in a Birth Cohort*. University of Chicago Press, 1972.
- [22] D. Weisburd, "The law of crime concentration and the criminology of place," *Criminology*, vol. 53, no. 2, pp. 133–157, 2015.
- [23] W. Bernasco and W. Steenbeek, "More places than crimes: Implications for evaluating the law of crime concentration at place," *Journal of Quantitative Criminology*, vol. 33, no. 3, pp. 451–467, 2017.
- [24] S. Meiklejohn *et al.*, "A fistful of bitcoins: characterizing payments among men with no names," in *Proceedings of the Internet measurement conference*, pp. 127–140, ACM, 2013.
- [25] A. Narayanan *et al.*, *Bitcoin and cryptocurrency technologies: a comprehensive introduction*. Princeton University Press, 2016.
- [26] J. Nick, "Data-driven de-anonymization in bitcoin," Master's thesis, ETH Zürich, 2015.
- [27] "Dream market – the biggest darkweb marketplace 2018,"
- [28] C. Cimpanu, "Top dark web marketplace will shut down next month," March 2019.
- [29] Y. Wu, F. Zhao, X. Chen, P. Skums, E. L. Sevigny, D. Maimon, M. Ouellet, M. H. Swahn, S. M. Strasser, M. J. Feizollahi, *et al.*, "Python scrapers for scraping cryptomarkets on tor," in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*, pp. 244–260, Springer, 2019.
- [30] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008.
- [31] J. Gower, "A general coefficient of similarity and some of its properties," *Biometrics*, vol. 27, pp. 857–871, Dec 1971.
- [32] M. Beckmann, "Gower distance-4python," Jun 2018.