Computer Systems Science & Engineering
DOI:10.32604/csse.2021.017039
Article

Tech Science Press

An Ensemble Approach to Identify Firearm Listing on Tor Hidden-Services

Hashem Alyami[1], Mohd Faizan[2], Wael Alosaimi[3], Abdullah Alharbi[3], Abhishek Kumar Pandey[2], Md Tarique Jamal Ansari[4], Alka Agrawal[2] and Raees Ahmad Khan[2,*]

[1]Department of Computer Science, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia
[2]Department of Information Technology, Babasaheb Bhimrao Ambedkar University, Lucknow, 226025, Uttar Pradesh, India
[3]Department of Information Technology, College of Computers and Information Technology, Taif University, P. O. Box 11099, Taif 21944, Saudi Arabia
[4]Department of Computer Application, Integral University, Lucknow, 226026, Uttar Pradesh, India
*Corresponding Author: Raees Ahmad Khan. Email: khanraees@yahoo.com
Received: 18 January 2021; Accepted: 19 February 2021

Abstract: The ubiquitous nature of the internet has made it easier for criminals to carry out illegal activities online. The sale of illegal firearms and weaponry on dark web cryptomarkets is one such example of it. To aid the law enforcement agencies in curbing the illicit trade of firearms on cryptomarkets, this paper has proposed an automated technique employing ensemble machine learning models to detect the firearms listings on cryptomarkets. In this work, we have used part-of-speech (PoS) tagged features in conjunction with n-gram models to construct the feature set for the ensemble model. We studied the effectiveness of the proposed features in the performance of the classification model and the relative change in the dimensionality of the feature set. The experiments and evaluations are performed on the data belonging to the three popular cryptomarkets on the Tor dark web from a publicly available dataset. The prediction of the classification model can be utilized to identify the key vendors in the ecosystem of the illegal trade of firearms. This information can then be used by law enforcement agencies to bust firearm trafficking on the dark web.

Keywords: Dark web; firearms; pistols; rifles; cryptomarkets; vendors

1 Introduction

The dominance of Internet-based technology is growing at a fast pace which also opens up the opportunity to perform illicit activities online. Cryptomarkets are one such byproduct on the Dark web that provides a safe environment for the number of illegal activities to be performed. From counterfeit currency to forged documents and prohibited drugs to firearms, all under one roof [1]. The Silk Road marketplace was one of the finest examples of these cryptomarkets [2]. The cryptomarkets usually employ The Onion Routing (Tor) [3] technique and cryptocurrency to ensure the anonymity of each transaction thereby disguising the real identity of their customers and vendors.

The weapon trafficking on the dark web came into limelight after several mass shootings and terror attacks that took place in the European countries. The firearms used by the perpetrators in the November 2015 Paris attacks and the 2016 Munich shooting were reportedly bought from the dark web marketplace [4]. Moreover, the same type of firearms called assault rifles were used in the Charlie Hebdo attack followed by the November attacks in Paris in the year 2015 [4,5]. The firearm used in the Munich attack was a Glock 17 pistol that belongs to the category of handguns [6]. Recently, a teenager has also been jailed for ordering a handgun (Glock 17) on the dark web [7]. Studies have found that the pistols and rifles were used in the majority of the mass shootings that happened in the United States [8]. These findings indicate that the executors of the mass shootings prefer pistols and rifles over other firearms.

Despite limited research available regarding the trafficking of illegal firearms on the dark web, two of the studies have identified firearms belonging to the category of pistols and rifles as the most commonly available weapon on the cryptomarkets [9,10]. These findings indicate the popularity of pistols and rifles among the customers on the dark web. The identification of key vendors of these specific types of firearms operating on the single cryptomarket or multiple marketplaces could prove beneficial for the law enforcement agencies. The law enforcement agencies could use the vendor information to tracking them down. In an attempt to aid the law enforcement agencies, we present an automated approach to detect the listings of pistols and rifles on the cryptomarkets for the identification of the key vendors.

The main contributions of our study are as follows:

- We found that the part-of-speech tagged features belonging to the noun category could improve the classification performance when used exclusively in the feature set.
- We investigated the different N-grams (N = 1, 2, 3) model for optimal feature extractions and found that the combination of unigrams and bigrams provide the best results.
- We proposed a stacked ensemble of Naïve Bayes and Random Forest classifiers to predict the firearm listings on the crytomarkets with better performance than the individual base classifiers.
- We carried out the experiments on a publicly available dataset to show the efficacy of the proposed classification model.
- The predicted firearm listings by the ensemble classification model can be mapped to their associated vendor profile. By aggregating all such listings, the key vendors selling illegal firearms across multiple marketplaces are identified.

The rest of the paper is organized as follows: Section 2 describes the related work. Also, section 2 elaborates on the proposed approach. Section 3 provides the experiment settings followed by results. Finally, Section 4 and Section 5 draw the discussion and conclusions of work.

2 Materials and Methods

2.1 Pertinent Related Works

The academic research concentrated on illegal weapon trafficking on the dark web is limited. Despite lacking in research on this issue, there have been a number of cases that indicate the prevalence of weapon trafficking on the Dark web [11]. Moreover, multiple sellers are under prosecution for the charges of illegal weapon trafficking on the dark web cryptomarkets [12–14]. One of the first studies on firearm trafficking on the dark web was conducted by the RAND Corporation [9]. The study has analyzed nearly eight hundred firearm-related listings across twelve cryptomarkets on the Tor dark web. Pistols are the most commonly available firearm followed by the rifles and sub-machine guns. Digital products that include guides and manuals on manufacturing firearms and explosives at home are the second most common category. However, the firearms were reasonably expensive as compared to the retail market. The study has not included the vendor shops operated exclusively by a single seller in the analysis.

Another study, similar to the RAND Corporation, has analyzed six online shops explicitly dealing in the firearms on the Tor dark web (Copeland [10]). Handguns especially pistols, were the most popular product as found in the earlier study (Paoli et al. [9]). Though, only a single shop among the six was accounted for the majority of handguns listings. Handguns were also a popular product among the vendors such that each of the identified vendors was dealing in handguns. Vendors often provide a range of products at premium prices. Rhumorbarbe et al. [15] analyzed the webpages of nine cryptomarkets to get insight into the trafficking of weapons on the dark web. Non-lethal weapons like pepper sprays, tasers, etc. contribute to around 28% of all the 386 weapons listings identified, followed by the firearms with 25%. They have identified 96 different vendor profiles with some vendors that were found operating across several markets. They concluded that the weapons trafficking on the cryptomarkets is much less when compared with other products like illicit drugs. A study has proposed an automated approach to identify the posts on the dark web forum for the procurement of weapons [16]. The datasets from four dark web discussion forums were manually labeled by the experts. The machine learning techniques were used to classify these labeled posts.

The above studies were exploratory in nature that attempts to uncover the size of cryptomarkets in terms of weapon trafficking, characteristics of vendors and the types of weapons on sale. Handguns (pistols) were the most preferred product category both on the cryptomarkets and the vendor shops. Pistols were also used in many of the mass shootings that happened recently. Therefore, identifying the key vendors of pistols and rifles on these online platforms could help the law enforcement agencies in taking down these vendors. To the best of the author's knowledge, no study has leveraged the data of weapon trafficking on cryptomarkets for identifying top vendors. Hence, we present our ensemble machine learning approach to detect firearm of the category of pistols and rifles. Henceforth, the term firearms shall be used to refer to the pistols and rifles together.

2.2 Proposed Approach

The proposed approach for detecting the firearms listings consists of the four steps: i) Pre-processing, ii) Feature Construction, iii) Detection of Firearm Listing and iv) Key Vendor Identification. These steps are explained in detail in the following subsections. Fig. 1 shows the diagrammatic representation of the proposed approach.

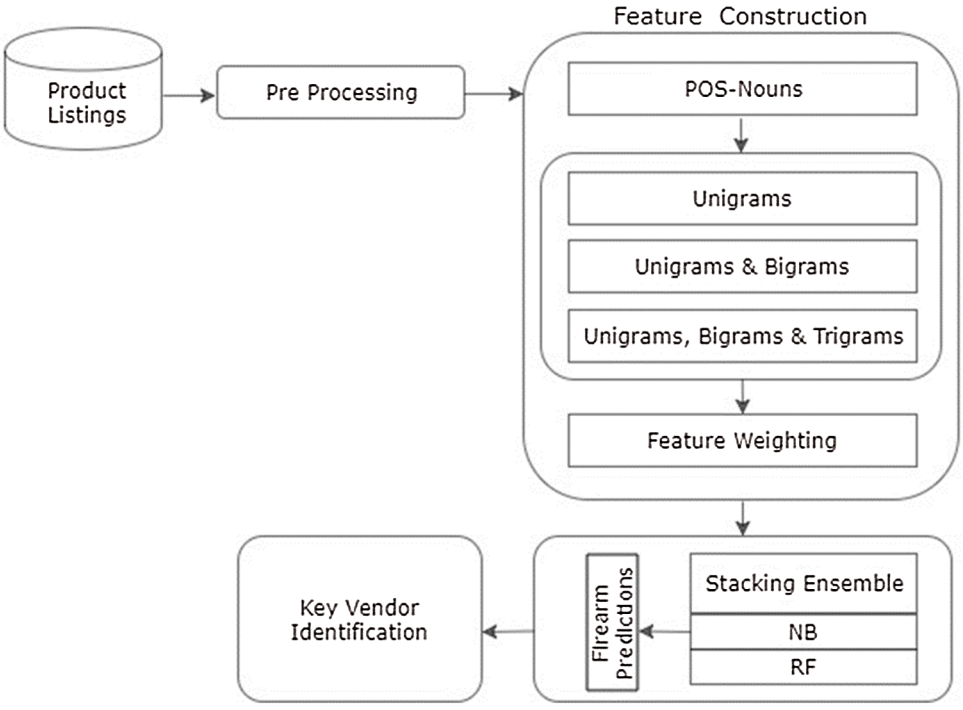

Figure 1: Proposed methodology

2.2.1 Pre-Processing

The aim of the data pre-processing step is to remove all the irrelevant and noisy data before proceeding to the feature extraction and then classification. The following steps were applied to the dataset during pre-processing:

- Tokenization: It breaks the string of text into individual tokens (words or numbers) at every whitespace character encountered.
- Removal of Stop-words: Stop-words are the common words (like a, am, the, is, of, etc.) that are frequently used in text and bear no meaning to the context of the text. Removing stop-words reduce the size of the text and improve the efficiency of the classification model. A list of stop-words is used to identify and remove them from the text.
- Noise Removal: Noisy data like special characters, Uniform Resource Locator (URLs), punctuations are removed as they did not convey any meaning.

2.2.2 Feature Construction

Feature extraction is performed to extract the most appropriate and relevant features to reduce the dimensionality of the feature space, which may help in better classification performance. The feature extraction consists of Part-of-Speech (PoS) tagging and N-grams extraction.

- Part-of-Speech tagging: The Part-of-Speech tagging method has been used to extract key tokens that could affect the classification performance like the verb, noun, adverb, adjective, etc. In our work, we have extracted the tokens that have been tagged as nouns. Much of the information about the weapon listings contains features like the manufacturer name, make, model number, caliber, finesse, action type, etc. which are nouns. Hence, it was decided to extract the noun tagged tokens for further processing and to see its effect on the classification performance.
- Extraction of N-grams: The N-gram techniques are common in classification and have been used in a variety of problems. N-gram is a string of consecutive words where N represents the number of words in the string. The most common N-grams are unigram (N = 1), bigram (N = 2) and trigram (N = 3). The listings of products in the dataset include the name of products, their manufacturers and other identifying information (like semi-automatic) that may consist of two or three words. Therefore in our work, we shall be using the combination of unigram, bigram, and trigram for adequate feature representation.
- Feature Representation and Weighting: The feature vector is represented in the Bag of Words (BoW) model, where the vocabulary of each of the unique features in the dataset is created. Feature occurring in less than two instances are removed from the vocabulary. We have used term frequency-inverse document frequency (TF-IDF) to assign a weight to each of the features in the vocabulary. The TF-IDF assigns high weight to the frequently appearing feature in an individual instance of the dataset. However, if the feature frequently appears in most of the instances of the dataset, then it will be assigned a lower weight.

2.2.3 Detection of Firearm Listing

The identification of firearm listing is achieved with the help of two classification models, namely: Naïve Bayes and Random Forest.

- Naïve Bayes: Naïve Bayes (NB) is a probabilistic supervised classifier based on the Bayes theorem. It makes a naïve assumption that the features are conditionally independent of each other. NB predicts the probability of an event by calculating the joint probability with respect to the occurrence of the other event. NB is among one of the popular classification algorithms with good classification performance. The simplicity along with the easy and fast implementation of the NB classifier motivated us to select it as one of the base classifiers for our ensemble model.

- Random Forest: Random Forest (RF) classifier is based on the decision tree algorithms. It is an ensemble of multiple individual decision trees that form a forest. Multiples samples from the training dataset are chosen to generate each decision tree. A final class value is chosen from the trees based on the majority voting scheme. RF can effectively manage the missing values and proper parameter tuning can prevent it from over-fitting. Given the track record of achieving good performance in previous studies, RF has been selected as another base model for the ensemble.

The ensemble technique leverages the power of multiple base classifiers in order to obtain improved classification performance as compared to a single base classifier. It combines the weak classifier with a strong classifier to get overall better classification output. The base classifier can be combined in several different ways to form the ensemble. To enhance the overall classification performance, we have applied the stacking ensemble technique that combines the two base classifiers in a sequential manner in stacking. In stacking, the base classifiers are placed one over the other resembling stack data structure such that each of the base classifiers passes its prediction to the classifier above it. The classifier at the bottom layer accepts the input data and the top-most classifier makes the final prediction.

2.2.4 Key Vendor Identification

It has been identified that the vendors manage their profile across several cryptomarkets to resist the cost of disruption in markets by law enforcement agencies [17]. It also gives them the opportunity to advertise their products in multiple markets to expand their business [18]. The aggregation of such vendors across several marketplaces will help in identifying the key vendors. To link the vendors operating on several marketplaces, we shall be considering the usernames and the PGP key of the vendors as characteristic features identified by Broseus et al. [18]. However, unlike Broseus et al. [18], we ignore the case of the username while matching them. The method of identification of the key vendor of firearms is described in the next paragraph.

The listings of the firearms predicted by the best classification model are identified. The username and the corresponding PGP key of the vendor associated with the listings are extracted. Two usernames were considered to be belonging to the same vendor profile if the usernames are matched exactly (ignoring the case). In case of mismatch, the Levenshtein distance between the username is calculated. If the Levenshtein distance is less than 25 percent and the PGP key is the same, then the vendor profile is considered to be the same. The common vendor profiles were aggregated based on the number of listings they are associated with. The vendors with the most number of listings might be the key vendors of the firearms.

2.3 Experimental Setup

2.3.1 Dataset

The experiment was conducted on a publicly available dataset collected by an independent researcher Gwern Branwen [19]. The dataset comprises of the downloads from 87 marketplaces and 37 forums related to these marketplace spanning a period of nearly two years. This dataset has been used in the number of previous peer-reviewed studies [20–22]. From the 87 marketplaces, we have selected the data from the following marketplaces: Alphabay, Armory and Dreammarket. These marketplaces have selected because they have been identified as offering the maximum number of weapon-related listings in the previous research [9,15].

The weapon-related listings from the download of each of the marketplace were identified and extracted. These listings were manually labeled into three classes: pistol class, rifle class that contains the listings of the pistols and rifles respectively and the other class contains all other listings including ammunition, other types of guns, digital materials and all other products on cryptomarkets like drugs, etc. The distribution of the listings among the three classes is given in Tab. 1.

Table 1: Dataset description

Class	Number of Listings
Pistols	392
Rifles	378
Others	2230
Total	3000

The listing pages were parsed with the BeautifulSoup to extract all the textual content, including the product and vendor information and stored as textual documents. These text documents shall be used in the classification model. The final model is tested on the 3000 product listings and the key vendors are identified from it.

2.3.2 Evaluation Metrics

The proposed classification models need to be evaluated on certain metrics. Accuracy, precision, recall and F-score can be used to evaluate the effectiveness of the proposed model. Accuracy is the ratio of the correctly classified instances to the total number of instances in the dataset under consideration. However, accuracy may not provide a good measure of the effectiveness of the model when the instances are not distributed equally among the classes. Therefore, in the case of imbalanced classification, precision, recall and F-score can be used for evaluation.

3 Data Interpretation and Results

To detect firearm listings, two classification models are used: NB and RF. Initially, we have used these two models individually to assess their performance on the dataset. The dataset was split into the training and testing set with a k-fold cross-validation scheme. Tab. 2 shows the comparison of the results between the individual classifiers and the stacking ensemble to further improve the prediction performance.

Table 2: Comparison of the individual classifiers and the ensemble

Classifier	Precision	Recall	F-score
NB	78.88	80.49	79.68
RF	85.20	83.73	84.46
Stacking (NB, RF)	87.58	88.37	87.97

In the stacking ensemble, the NB and RF are the base estimators and the logistic regression is the final estimator. Tabs. 3 and 4 shows the effect of taking only PoS tagged noun with N-gram (N = 1, 2, 3) tokens for the feature set on the performance of the classification models.

The size of the features space when using POS tagged features is given in Tab. 5. Further, Tab. 6 shows the top five key vendors identified being involved in the trade of pistols and rifles in the dataset.

All the experiments were conducted on Python v3.6 IDLE on Windows machine with the hardware configuration of 4 GB Random Access Memory and Intel i5 processor.

Table 3: Comparison of the individual classifier with part-of-speech tagged features and n-grams

Classifier		Precision	Recall	F-score
NB	uni	89.42	79.52	84.07
	uni+bi	90.54	79.02	84.39
	uni+bi+tri	91.28	79.02	84.32
RF	uni	88.90	87.26	88.07
	uni+bi	94.83	94.65	94.74
	uni+bi+tri	92.31	91.79	92.05

Table 4: Performance of the stacking ensemble with part-of-speech tagged features and n-grams

Classifier		Precision	Recall	F-score
Stacking (NB, RF)	uni	91.43	92.15	91.79
	uni+bi	97.15	96.84	96.99
	uni+bi+tri	93.46	92.75	93.10

Table 5: Size of the feature space

Type of Feature Set	Size
Full	2003
POS Tagged with Unigrams	1063
POS Tagged with Unigrams and Bigrams	2470
POS Tagged with Unigrams, Bigrams & Trigrams	3874

Table 6: Key firearm vendors

Vendor	Number of Listings	Type of Firearm
V1	14	Pistols & Rifles
V2	10	Pistols
V3	9	Pistols
V4	6	Rifles
V5	5	Pistols

4 Discussion

The assessment of the proposed method is done using precision, recall and the F-score. The stacking ensemble consists of the NB and RF as the base classifiers and the logistic regression as the final estimator. The base classifiers are first evaluated separately and then compared with the performance of the stacking ensemble of NB and RF. While evaluating the base classifier individually, RF performs better than the NB; however, NB was computationally faster than the RF. The ensemble of the two

classifiers produced much better results than the individual models. The output of the stacking managed to get 87.97% for F-score which indicates that the ensemble of NB+RF is more capable of predicting the specific weapon listings. The final estimator of Logistic Regression yielded a better output of integrating NB and RF. The combination of NB and RF has a significant lead over the individual performance of the NB and RF. The only limitation with this ensemble setting is the longer time needed to construct the model which is common in such setup.

Tab. 3 shows the evaluation of the individual classifiers when they are fed with the POS extracted feature set. The result of utilizing the POS tagged features with a different combination of N-grams (N = 1, 2, 3) is presented. The use of POS tagged features instead of the full feature space has significantly increased the F-score both for the NB and the RF classifiers. However, the best performance is achieved when using the combination of unigrams and bigrams of the POS tagged features with a slight increase in the dimensionality of the feature set, as can be seen in Tab. 5. Many of the firearms names are of two words like the Glock pistol series and M1 Garand, M 1841, Ruger M77 rifles. The use of POS tagged features (nouns in our case) with bigrams may act as the discriminative features in the classification and hence produced better performance.

In this work, we have made a comparison between the performance of the stacking ensemble without the original feature set and the stacking ensemble with the POS extracted feature set with unigrams and bigrams. The classification results show that the stacking of the base classifiers NB and RF with the combination of unigram and bigrams POS extracted features significantly improves the performance of the model. The predictions of our proposed model can be applied to the cryptomarket data for identifying the key vendors of firearms.

5 Conclusions

In this paper, we have put forward a methodology to detect the listings of pistols and rifles for the identification of the key vendors across the cryptomarkets. The proposed method aims to identify the features in the dataset having contextual meaning to improve the performance. Therefore, we have extracted the POS tagged nouns, followed by their N-grams, to obtain the most representative feature set. We then use the stacking ensemble of NB and RF for the detection of the specific firearm listings. The experimental results show that the feature set with part-of-speech tagged nouns particularly with the combination of unigrams and bigrams features and ensemble classifiers, produces the best classification performance on the dataset.

Acknowledgement: This research was supported by Taif University Researchers Supporting Project number (TURSP-2020/254), Taif University, Taif, Saudi Arabia.

Funding Statement: Funding for this study is received from the Taif University Research Supporting Projects at Taif University, Kingdom of Saudi Arabia under Grant No. TURSP-2020/254.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] M. Faizan and R. A. Khan, "Exploring and analyzing the dark web: A new alchemy," *First Monday*, vol. 24, no. 5, pp. 1–20, 2019.

[2] N. Christin, "Traveling the silk road: A measurement analysis of a large anonymous online marketplace," in *Proc. 22nd Int. World Wide Web Conf.*, Brazil: Rio de Janeiro, pp. 213–224, 2013.

[3] R. Dingledine, N. Mathewson and P. Syverson, "Tor: The second-generation onion router," in *Proc. the 13th Conf. on USENIX Security Symposium*, San Diego, CA, pp. 21–22, 2004, https://dl.acm.org/doi/10.5555/1251375.1251396.

[4] J. Huggler, "Man arrested in Germany on suspicion of illegal arm dealing in terror crackdown." London, England: Telegraph Media Group Limited, 2015. [Online]. Available at: https://www.telegraph.co.uk/news/worldnews/europe/germany/12020249/Paris-attackers-bought-weapons-from-arms-dealer-in-Germany.html.

[5] C. Woolf, "Where did the Paris attackers get their guns?." Boston, MA: PRX, The World, 2015. [Online]. Available at: https://www.pri.org/stories/2015-01-15/where-did-paris-attackers-get-their-guns/.

[6] Digital Culture, "Munich shooter bought gun online: Here's how the dark networks." Berlin, Germany: Deutsche Welle, 2016. [Online]. Available at: https://www.dw.com/en/munich-shooter-bought-gun-online-heres-how-the-dark-net-works/a-19424920/.

[7] S. Morris, "Teenager obsessed with mass shootings jailed for buying gun online." London, England: The Guardian, 2019. [Online]. Available at: https://www.theguardian.com/world/2019/sep/13/teenager-obsessed-with-mass-shootings-jailed-for-buying-gun-online/.

[8] J. A. Fox and M. J. Delateur, "Mass Shootings in America," *Homicide Studies*, vol. 18, no. 1, pp. 125–145, 2013.

[9] G. Paoli, J. Aldridge, N. Ryan and R. Warnes, "Behind the curtain: The illicit trade of firearms," *Explosives and Ammunition on the Dark Web*, pp. 1–17, 2017. [Online]. Available at: http://www.rand.org/t/RR2091.

[10] C. Copeland, M. Wallin and T. J. Holt, "Assessing the practices and products of darkweb firearm vendors," *Deviant Behavior*, vol. 41, no. 8, pp. 949–968, 2020.

[11] G. Weimann, "Terrorist migration to the dark web," *Perspective on Terrorism*, vol. 10, no. 3, pp. 40–44, 2016.

[12] T. McKay, "Feds bust over 35 suspected dark web vendors seizing everything from drugs to a grenade launcher." CA, United States: Gizmodo, G/O Media Inc., 2018. [Online]. Available at: https://gizmodo.com/feds-bust-over-35-suspected-dark-web-vendors-seizing-e-1827159590.

[13] Justice News, "Montgomery man convicted for illegal gun sales on darknet sites." Montgomery, USA: United States Department of Justice, 2015. [Online]. Available at: https://www.justice.gov/usao-mdal/pr/montgomery-man-convicted-illegal-gun-salesdarknet-Sites.

[14] Justice News, "Kansas man sentenced to 52 months for exporting firearms to overseas purchasers using hidden marketplace website." Washington, DC, USA: United States Department of Justice, 2017. [Online]. Available at: https://www.justice.gov/opa/pr/kansas-man-sentenced-52-months-exporting-firearms-overseas-purchasers-using-hidden.

[15] D. Rhumorbarbe, D. Werner, Q. Gilliéron, L. Staehli, J. Broséus *et al.,* "Characterising the online weapons trafficking on cryptomarkets," *Forensic Science Int.*, vol. 283, no. 6, pp. 16–20, 2018.

[16] J. K. Saini and D. Bansal, "A comparative study and automated detection of illegal weapon procurement over dark web," *Cybernetics and Systems*, vol. 50, no. 5, pp. 405–416, 2019.

[17] K. Soska and N. Christin, "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem," in *Proc. USENIX Security Symposium*, Washington, D.C, pp. 33–48, 2015, [Online]. Available: https://www.usenix.org/system/files/sec15-paper-soska-updated_v2.pdf.

[18] J. Broseus, "Studying illicit drug trafficking on darknet markets: Structure and organisation from a Canadian perspective," *Forensic Science Int.*, vol. 264, no. 3, pp. 7–14, 2016.

[19] G. Branwen, "Dark net market archives," (2011–2015). Ireland, UK: Gwern.net, 2015. [Online]. Available at: https://www.gwern.net/DNM-archives/.

[20] O. Cherqi, G. Mezzour, M. Ghogho and M. Elkoutbi, "Analysis of hacking related trade in the darkweb," in *Proc. IEEE Int. Conf. on Intelligence and Security Informatics (ISI)*, Miami, FL, USA, pp. 79–84, 2018.

[21] A. Berman and C. L. Paul, "Making sense of darknet markets: Automatic inference of semantic classifications from unconventional multimedia datasets," *Lecture Notes in Computer Science*, vol. 11594, pp. 230–248, 2019.

[22] S. Foley, J. R. Karlsen and T. J. Putnins, "Sex, drugs, and bitcoin: how much illegal activity is financed through cryptocurrencies?," *Review of Financial Studies*, vol. 32, no. 5, pp. 1798–1853, 2019.