# Ranking potentially harmful Tor hidden services: Illicit drugs perspective

Mohd Faizan, Raees Ahmad Khan and Alka Agrawal
*Department of Information Technology, Babasaheb Bhimrao Ambedkar University, Lucknow, India*

Received 24 December 2019
Revised 3 February 2020
Accepted 21 February 2020

## Abstract

Cryptomarkets on the dark web have emerged as a hub for the sale of illicit drugs. They have made it easier for the customers to get access to illicit drugs online while ensuring their anonymity. The easy availability of potentially harmful drugs has resulted in a significant impact on public health. Consequently, law enforcement agencies put a lot of effort and resources into shutting down online markets on the dark web. A lot of research work has also been conducted to understand the working of customers and vendors involved in the cryptomarkets that may help the law enforcement agencies. In this research, we present a ranking methodology to identify and rank top markets dealing in harmful illicit drugs. Using named entity recognition, a harm score of a drug market is calculated to indicate the degree of threat followed by the ranking of drug markets. The top-ranked markets are the ones selling the most harmful drugs. The rankings thus obtained can be helpful to law enforcement agencies by locating specific markets selling harmful illicit drugs and their further monitoring.

**Keywords** Dark web, Illicit drugs, Ranking, Tor hidden services

**Paper type** Original Article

## 1. Introduction

The proliferation of the Internet and communications technology has paved the way for a variety of services to the general public. This has also made it easier to carry out a range of illicit activities. Dark web, an anonymous platform is one such place where illegitimate activities are prevalent like sale of illegal drugs, distribution of child abuse content, violence and hate content [1–3]. The Onion Router (Tor) is the most common method to access the dark web, although other methods like Freenet and I2P do exist but are used by fewer users [4]. A website on the Tor dark web is known as hidden service (HS). Among all the goods and services present on these online platforms, illicit drugs are the most popular product [5]. These online platforms are generally referred to as cryptomarkets or darknet markets (DNM). The easy availability of these illicit drugs may significantly harm consumers due to drug

Publishers note: The publisher wishes to inform readers that the article “Ranking potentially harmful Tor hidden services: Illicit drugs perspective” was originally published by the previous publisher of Applied Computing and Informatics and the pagination of this article has been subsequently changed. There has been no change to the content of the article. This change was necessary for the journal to transition from the previous publisher to the new one. The publisher sincerely apologises for any inconvenience caused. To access and cite this article, please use Faizan, M., Khan, R.A., Agrawal, A. (2020), “Ranking potentially harmful Tor hidden services: Illicit drugs perspective”, New England Journal of Entrepreneurship. Vol. 18 No. 3/4, pp. 267-278. https://10.1016/j.aci.2020.02.003. The original publication date for this paper was 24/02/2020

*Declaration of Competing Interest*: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Applied Computing and Informatics
Vol. 18 No. 3/4, 2022
pp. 267-278
Emerald Publishing Limited
e-ISSN: 2210-8327
p-ISSN: 2634-1964
DOI 10.1016/j.aci.2020.02.003

overdoses [6] which require continuous surveillance of DNM by the law enforcement agencies.

Much research was done to aid the law enforcement agencies in getting the modus-operandi of DNM [7–9]. However, it could not provide methods to quantify the overall threat or harm incurred by online illicit drug platforms in terms of health concerns. There exist key players in the online ecosystem in terms of health impact and their activity requires immediate investigation from the law enforcement agencies. They are characterized in the dark web by the broad range of illicit drugs they sell.

Therefore this study tries to fill this gap by proposing the content-based ranking methodology that can help the law enforcement agencies in identifying the most harmful dark markets dealing with drugs. Consequently, the concerned agencies can put more significant efforts in order to shut down those markets. However, this does not imply that the other DNM should be left off instead the law enforcement agencies may prioritize their efforts towards DNM that requires immediate investigation and action. The proposed methodology is evaluated with standard metrics by conducting experiments on the Tor dark web dataset. The term HS shall refer to the DNM and websites on Tor dark web that deal with illicit drugs.

The rest of the paper is organized as follows: Section 2 describes the related work. Section 3 elaborates on the proposed ranking methodology. Section 4 provides the experiment settings followed by results and discussion. Finally, Section 5 draws the conclusion and possible future work.

## 2. Related work

The dark web markets have witnessed a surge in their size and significance after the launch of Silk Road, with the estimated annual drug trade of USD 170 million [10]. Several studies have managed to reveal the properties, working and impacts of DNM [11,12]. One of the first studies on cryptomarkets was focused on the Silk Road market, where the author collected and analyzed the data from the live market on a regular basis. The author found that drug-related products were the most popular [9]. The qualitative analysis of the discussion forums of the DNM has uncovered the user experiences regarding the purity, effects, potency of drugs and the methods to evaluate the quality of drugs [13].

The quantitative method generally uncovers the geographical reach of vendors and their lifespan, customer retention methods, trade statistics and market challenges. A descriptive study on DNM has found that the largest number of vendors operate from the United States followed by European countries. However, The Netherlands has the highest number of vendors per 100 thousands of the population [14]. In an attempt to maintain a healthy customer base and to inflate the product reviews on the DNM, the vendors often resort to giving out free samples of their major drug products to the customers [15]. The analysis of the data collected from the AlphaBay cryptomarket indicates a highly competitive ecosystem among the vendors where only a small number of vendors manage to sustain their business through aggressive advertising [16]. A study focused on the sale of psychoactive substances has found that the vendors involved in the business of psychoactive substances have a short lifespan. However, the number of vendors on DNM shows an increasing trend [17].

The crawling mechanism for the surface web is simple given the easy availability of the website addresses on the Internet. Contrary to that there does not exist any system that contains the addresses of the Tor hidden services. This renders the crawling of hidden services a slow and challenging task. Moreover, a complete scan of the entire Tor web would be impractical as there would be $32^{16}$ different addresses to be crawled (the address of the Tor HS is composed of 16 characters) [18]. The researcher has used some of the publicly available listings of Tor HS addresses on the surface web for further exploration of the Tor dark web. The crawling process also gets marred by the requirement of login credentials in the case of

DNM. In order to maintain their privacy, there is also a risk of user account getting blocked by the DNM if they suspect crawling activity. A systematic methodology was proposed for scraping the DNM data where the automated spiders were used to overcome the crawling challenges [19].

A recent study [20] has proposed a link-based ranking technique called ToRank to identify influential Tor HS. As per the authors, the HS that is ranked higher by ToRank is more popular HS among others in the Tor dark web ecosystem. Identification of popular HS may help the law enforcement agencies in getting clues about the working of DNM. However, being a link analysis algorithm, ToRank purely relies on the hyperlinks between the HS and does not take into account the content of the HS while ranking them. Other studies have focused on dark web forums to identify key users [21,22,34].

The User Rank algorithm based on Page Rank was put forth to identify influential users through message content analysis and the level of attentiveness between the users [23].

The existing work explores the general working of the markets by collecting and analyzing customer feedback and comments, product ratings, vendor profiles, etc. The analysis of the DNM data can also be leveraged to quantitatively calculate the negative implications of the DNM resulting from the usage of illicit drugs. A hyperlink based approach may not be efficient in ranking the influential HS as most of the web pages in a Tor HS are difficult to find as they are linked by very few other pages, also the Tor HS has very low outgoing links to the other services [24]. Therefore we propose a content-based approach to rank HS trading in harmful illicit drugs. The proposed approach can be utilized to proactively monitor and investigate such HS. Moreover, the proposed approach also recognizes the *isolated* HS (one with no incoming and outgoing hyperlinks) if they possess the illicit drug, which otherwise is not detected by the conventional link analysis algorithms.

## 3. Proposed method

The proposed algorithm is specifically designed to work in the domain of illicit drug trading in the dark web. The computational time of the algorithm can be reduced if it is fed with the pre-identified dataset of drug-related HS. The existing work that focused on classifying suspicious activities on the dark web using machine learning can be used for this purpose [25]. The proposed algorithm can then be applied to retrieve top dangerous services from the bunch of HS. Figure 1 shows the proposed ranking technique. The design of the proposed ranking technique consists of the four main components: data preprocessing, illicit drugs name extraction, harm score calculation and rankings of HS.

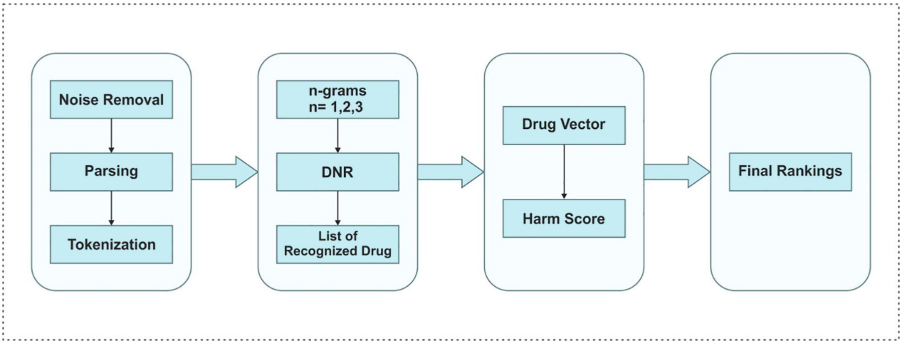

**Figure 1.**
The proposed ranking methodology.

### 3.1 Data preprocessing

The dataset used in this study consist of the HTML file representing each of the HS. A custom made Python script was employed to extract the available product listings from the HTML file for each HS in the dataset. Since a DNM may sell different types of products, the extracted set of listings may contain several products other than drugs and also from different vendors. After extracting all the product listings, the textual content with HTML tags removed is obtained and stored in a plain text file for each of the HS in the dataset. The plain text file is then processed to remove all the irrelevant content like script, hyperlinks, punctuations and white spaces. It is followed by converting all text to lowercase and then removing stop-words and duplicates. A parser using regular expression was employed to identify and remove numbers having either single-digit or more than three digits. The reason for doing this shall be discussed in Subsection 3.2. The obtained data is then put to perform tokenization that breaks long strings of text into smaller pieces called tokens. In our case, tokens shall be single words and numbers of two and three digits only. After the tokenization process, the tokens are stored in the form of a list in a text file for further processing.

### 3.2 Illicit drugs name extraction

The preprocessed data in the form of tokens received from the first step contains other words and product names along with the names of illicit drugs. Drug name recognition (DNR) shall be applied to recognize the name of drug-related products. DNR is a particular type of Named Entity Recognition (NER) task to extract names of drugs from unstructured text. DNR becomes very important and challenging in our task of extracting illicit drug names. To confuse the law enforcement agencies, consumers and vendors trading in illicit drugs use common street names/slang words for illicit drugs with some of the names being common English phrases and words used in day-to-day life. For e.g., *pastas* may refer to the class of drugs called amphetamines. Moreover, they also use the brand name of the drugs instead of their generic names. The drugs may also be referred in numeric form like 77,501 etc. because of this reason specific numeric tokens were kept at the preprocessing stage.

The bigrams and trigrams are generated from the list of tokens obtained after preprocessing. The purpose behind their creation is to identify drug names and their slangs composed of two or three words. A dictionary-based DNR approach is used to identify illicit drug names and their slangs followed by putting them into the appropriate class of drugs. Dictionary-based approaches require a drug dictionary to match against the text document. The dictionary-based approach is utilized in this study due to the context of the problem. As discussed above, drug vendors mostly use slang terms for trading illicit drugs on the DNM. These slang terms bear no relation with the original name of the drugs at all; also, there is no naming convention or nomenclature used for generating slang terms. In this context, a rule-based approach could not be effectively applied; on the other hand, a comprehensive dictionary of common/slang drug terms can be used for exact matching of the text. However, the drug dictionary should be updated regularly given the ever-changing ecosystem of the illicit drug trade. In our work, we have used the dictionary of slang and code words of drugs by the US Drugs Enforcement Administration [26].

In our approach, we shall be using DNR to identify eighteen different types of controlled and prohibited drugs identified in a study that evaluate the harmful effects of such drugs [27]. The study is discussed in Section 3.3. The description of the drugs is given in Table 1.

The final list of tokens containing the single tokens, bigrams and trigrams for each HS in the dataset is obtained. An associative array for each of the HS is created that records the drug type and its frequency in a *key-value* format. The final list of tokens is then matched against the drug dictionary. There may be typos present in the listings which can affect the search mechanism in the drug dictionary. To overcome this, we shall use the Levenshtein

| Drug Class | Drug Type |
|---|---|
| Stimulants | Amphetamines<br>Cocaine<br>Crack Cocaine<br>Khat<br>Methamphetamine<br>Mephedrone |
| Depressants | Benzodiazepines<br>Gamma-Hydroxybutyric Acid (GHB) |
| Hallucinogen | Ecstasy<br>Ketamines<br>Lysergic Acid Diethylamide (LSD)<br>Mushrooms |
| Narcotics (Opioids) | Buprenorphine<br>Heroin<br>Methadone |
| Inhalant | Butane |
| Steroid | Anabolic Steroids |
| Cannabis | Cannabis |

**Table 1.** Illicit drugs commonly available on the dark web.

distance to match the similarity of the token with the drug dictionary in case of typos. The Levenshtein distance measures the number of characters that should be changed to convert a string to the other one. If an exact match is found for a token or the Levenshtein distance of the token is less than 25 percent, then the corresponding drug type of that token from Table 1 is added in the *key* field of the associative array of that particular HS. After that, the matched token is searched in the set of extracted product listings for the HS. Since there may be multiple listings of a single drug type from various vendors in a HS, therefore the number of listings matched with the token is counted and stored in the *value* field of their corresponding drug type in the associative array which shall be the frequency of the drug type. The matched listings are removed from the extracted set of listings. This procedure is repeated for each of the tokens that get matched with the drug dictionary. If a token is matched against the drug type that already exists in the associative array, then the frequency of that drug type is updated accordingly. Once all the tokens are matched, the remaining listings left in the set are non-drug products and these listings are discarded. The associative array containing drug types recognized along with their frequency for each of the HS is passed on to the next stage for the calculation of harm score. An example of the final associative array of a HS offering *cocaine*, *LSD, cannabis* and *ecstasy* is shown in Figure 2.

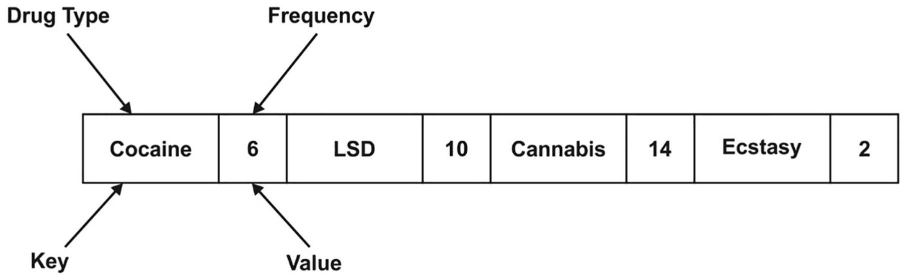

**Figure 2.** Associative array.

### 3.3 Measuring harm score of HS

Drug abuse has become a significant health concern for individuals and society. Therefore the use of such drugs is controlled and prohibited by the policy-making bodies. The abuse of drugs can affect in multiple ways from individual harm to environmental and economic damage. Hence it is required to estimate the harm of each of the controlled drugs in terms of the multitude of ways it affects. An existing study has proposed a method to solve this problem where the authors' aim was to aid policymakers in the field of health and investigation by evaluating the harms caused by drug abuse [27]. A committee of drug experts from the United Kingdom was formed to assess the harm of 20 drugs based on 16 criteria using multi-criteria decision analysis (MCDA). The drugs were given an overall score on a scale from 0 to 100, with 0 indicating the least harmful and 100 being the most harmful drug on all 16 criteria. The name of the illicit drugs and their corresponding score is shown in Table 2. The list of criteria were: Drug-specific mortality, Drug-related mortality, Drug-specific damage, Drug-related damage, Dependence, Drug-specific impairment of mental functioning, Drug-related impairment of mental functioning, Loss of tangibles, Loss of relationships, Injury (to others), Crime, Environmental damage, Family adversities, International damage, Economic cost, and Community. After some criticism of the applied methodology, the authors have come up with a follow-up study to assess the harm of drugs on a broader scope. However, the scorings obtained in the follow up were very similar to the previous one with the correlation of 0.993 [28].

To calculate the overall harm score of an HS, we shall be adopting the scoring of the illicit drugs indicating their cumulative harm from the study discussed above [27]. It should be noted that *alcohol* and *tobacco* were included in the 20 drugs that were assessed in the study; however, our work is in the context of illicit drugs hence we did not consider *alcohol* and *tobacco* as they are not controlled substances.

Let $L_i(i = 1, 2, \ldots M)$ be the associative array of the $i^{th}$ HS $H_i$ in the dataset, a drug vector $V_i$ of length $n(n = 18)$ is created for $H_i$. The elements $x_i^k(k = 1, 2, \ldots 18)$ of $V_i$ are obtained using Eq. (1). $d_k$ and $t(d_k)$ are the drug type and its individual harm score obtained from Table 2.

| S. No. ($k$) | Drug Type ($d_k$) | Individual Harm Score $t(d_k)$ |
|---|---|---|
| 1. | Heroin | 55 |
| 2. | Crack Cocaine | 54 |
| 3. | Methamphetamine | 33 |
| 4. | Cocaine | 27 |
| 5. | Amphetamines | 23 |
| 6. | Cannabis | 20 |
| 7. | Gamma-Hydroxybutyric Acid (GHB) | 19 |
| 8. | Benzodiazepines | 15 |
| 9. | Ketamines | 15 |
| 10. | Methadone | 14 |
| 11. | Mephedrone | 13 |
| 12. | Butane | 11 |
| 13. | Anabolic Steroids | 10 |
| 14. | Khat | 9 |
| 15. | Ecstasy | 9 |
| 16. | Lysergic Acid Diethylamide (LSD) | 7 |
| 17. | Buprenorphine | 7 |
| 18. | Mushrooms | 6 |

**Table 2.**
Illicit drugs and their harm score based on 16 criteria.

$$x_i^k = \begin{cases} t(d_k)*f(d_k), & \text{if } d_k \in L_i \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$f(d_k)$ is the frequency of drug type $d_k$ in the associative array $L_i$.

A harm score $\tau(H_i)$ is assigned to $H_i$ given by Eq. (2).

$$\tau(H_i) = log_{10}\left(1 + \frac{|V_i|}{\upsilon_o}\right) \quad (2)$$

where $|V_i| = \sum_k x_i^k$ and $\upsilon_o = \min[t(d_k)]$

As the HS may sell a number of different products with multiple listings, $|V_i|$ may get a very large numeric value. Therefore, the logarithm function is used to calculate the harm score to conveniently express the large values. One is added in the logarithm function to avoid getting zero in the argument (when $|V_i| = 0$) for which the log function is undefined. An ethical HS that does not deal in illicit drugs shall have $|V| = 0$ and subsequently, a harm score of zero indicating that it does not pose any ill effects on its users.

### *3.4 Ranking*

As there would be some HS that trade in less harmful drugs while others may offer potentially harmful drugs, so in order to identify the most severe HS, we need to rank them. The overall ranking shall depend on the harm score of the HS. The HS are ranked in the descending order of their harm score implying that the HS with the highest harm score would get the top rank. In the case of a tie, the HS with a drug having the highest individual harm score is placed above in the rankings. For e.g: let X and Y be two HS, X contains a single listing of *crack cocaine* and Y contains two listings of *cocaine*. Both X and Y have the same harm score but X shall be placed above Y in ranking because of the presence of *crack cocaine* (with individual harm score 54).

## 4. Experiments

The proposed ranking methodology attempts to uncover potentially harmful HS using content analysis technique. The rankings generated by the proposed technique need to be evaluated on the standard metrics for ranking problems. However, to the best of authors' knowledge, there is no *gold standard* or *ground truth* on the content-based ordering of the HS against which the generated rankings can be evaluated. Hence, to evaluate the accuracy of our proposed ranking technique, the rankings of the HS in the dataset from the three experts are considered as the ground truth.

### *4.1 Dataset*

We have used the DUTA-10K dataset in our work. The DUTA-10 K has been used in an existing study to test the To Rank algorithm [20]. The dataset contains 10,367 labeled samples spread across 28 categories. Each sample represents a hidden service from the Tor dark web and contains the root page and the first level subpages of an HS in a single HTML file. In our case, we were only interested in the HS related to illicit drug trades therefore we have taken 255 HS samples dealing in illicit drugs and they were in English. The DUTA-10K dataset is publicly available for download at [29].

### *4.2 Ground truth generation*

The dataset was presented before the three experts for ranking the HS to obtain the ground truth. One expert was a professional medical doctor, another was a psychologist and the last one was from academia. All three experts were asked to independently rank each of the HS in the dataset based on the availability of illicit drug listings on HS, the severity of available drugs and the frequency of listings. This generates three rankings from each of the independent experts. Since the problem of ranking is very subjective and each of the expert perceptions on the drug harms may vary so the evaluation may be biased if we consider ranking from any of the single experts. To tackle this problem, we shall be using the aggregation method called rank-based aggregation (RBA) [30] to create final rankings from three experts. In this method, the individual rank of each HS from three experts is combined, and then the final rankings are generated. This method ensures that the noise that may creep in shall be compensated by other experts during aggregation. The final list thus generated shall be used as ground truth for evaluating the accuracy of the rankings generated by our proposed method.

### *4.3 Evaluation metrics*

The correctness of the rankings obtained from the proposed ranking methodology is evaluated by the Kendall's tau [31] metric commonly used in the field of information retrieval. Kendall's tau has been chosen given its wide use in the literature [23,35–38] and has been shown to be a more robust and efficient metrics than the others [39]. We have also used rank-biased overlap (RBO) metric that puts more importance to the top of the ranked list similar to the weighted Kendall's tau [40] as our work is focussed on identifying the top ranked HS. Moreover, RBO can efficiently handle the non-conjointness in the rankings as compared to the other metrics [32].

### *4.4 Results*

Python v3.6 [33] is used to implement the proposed ranking algorithm using harm score. The preprocessing of the dataset, including the removal of stop-words, was performed by the NLTK package. The harm score for each of the HS is obtained using Eqs. (1) and (2).

The top ten ranked HS retrieved by the proposed method and from the ground truth are shown in Table 3, respectively. For simplicity, the HS is denoted by alphabets instead of their onion domain. The HS with the highest-ranking obtained by the proposed algorithm is the same as evaluated by the experts. This HS is found to be selling eleven different illicit drugs from Table 1.

Kendall's tau measure is used for comparing the rankings generated by the proposed methodology to the ground truth rankings to assess the correctness of the ranks assigned to each of the HS.

For each pair of rank $(p_{i,}q_i),\ (p_jq_j)\cdots\cdots\cdots..(p_{n,}q_n)$ in the list $P$ and $Q$, $n_c$ and $n_d$ be the number of concordant pairs (if $p_i < p_j$ and $q_i < q_j$ or $p_i > p_j$ and $q_i > q_j$) and discordant pairs

**Table 3.**
Top ten ranked HS obtained from the proposed algorithm and the ground truth.

| HS | Retrieved Rank | Ground Truth | Difference |
|---|---|---|---|
| A | 1 | 1 | 0 |
| B | 2 | 2 | 0 |
| C | 3 | 3 | 0 |
| D | 4 | 4 | 0 |
| E | 5 | 5 | 0 |
| F | 6 | 6 | 0 |
| G | 7 | 7 | 0 |
| H | 8 | 8 | 0 |
| I | 9 | 10 | 1 |
| J | 10 | 9 | 1 |

(if $p_i < p_j$ and $q_i > q_j$ or $p_i > p_j$ and $q_i < q_j$) respectively. The Kendall's tau of the two ranking lists $P$ and $Q$ of size $n$ is given by Eq. (3):

$$R(P,\ Q) = \frac{n_c - n_d}{\frac{1}{2}n(n-1)} \tag{3}$$

Table 4 shows Kendall's tau for the proposed ranking algorithm for different values of $k$, where $k$ represents the number of pairs in the rankings. The maximum value of $k$ is set to $k = 50$ i.e. we examine the closeness between the two rankings up to the top 50 ranked HS. From Table 4, it is evident that the proposed algorithm can accurately predict the top ten harmful HS when compared to the ground truth. The Kendall's tau value close to one in Table 4 indicates that the two rankings are highly related. However, Kendall's tau slightly decreases when $k$ increases. The HS that have been allotted the top ranks indicate their potentially harmful nature and key position in illicit drug trade. The law enforcement agencies may allocate their resources more to these top-ranked HS in busting them down.

To further check the accuracy of our ranking methodology, we have obtained eight random samples of the rankings from the entire list and computed the Kendall's tau of the random samples and the ground truth. Table 5 shows the Kendall's tau of different samples, the size of each of the sample is 25. The high value of Kendall's tau for the random samples of ranks shows that the proposed ranking methodology is close to the ground truth data.

As in our work, the top-ranked nodes are of greater importance to the law enforcement agencies, the accuracy of the proposed method in the high ranks is examined. Rank-biased overlap (RBO) is used to measure the accuracy in high ranks by giving different weights for different ranks and allotting higher weights to high ranks. A higher RBO value indicates greater similarity and correlation between the two rankings. RBO of two ranking lists $P$ and $Q$ is given by Eq. (4):

$$RBO(P,\ Q,\ p) = (1-p)\times\sum_{d=1}^{r} p^{d-1}A(P,\ Q,\ d) \tag{4}$$

| k | Kendall's Tau |
|---|---|
| 10 | 0.9556 |
| 20 | 0.9474 |
| 30 | 0.9218 |
| 40 | 0.9120 |
| 50 | 0.8784 |

**Table 4.** Effectiveness of the proposed algorithm using Kendall's tau.

| Sample # | Kendall's Tau |
|---|---|
| 1 | 0.9232 |
| 2 | 0.8956 |
| 3 | 0.8920 |
| 4 | 0.9347 |
| 5 | 0.9465 |
| 6 | 0.9030 |
| 7 | 0.8845 |
| 8 | 0.9235 |

**Table 5.** Kendall's tau of different samples.

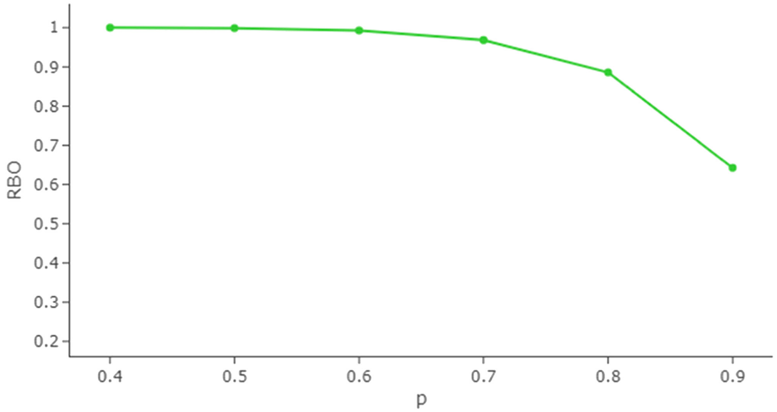

**Figure 3.**
The RBO values between the generated ranking and the ground truth.

where $A(P, Q, p)$ given by Eq. (5) is the overlap value between two rankings $P$ and $Q$ up to rank $d$, $r$ is the number of unique ranks and $p$ is a configurable parameter in (0,1) such that the smaller value of $p$ implies that the metric is more top-weighted.

$$A(P, Q, d) = \frac{|P_{1:d} \cap Q_{1:d}|}{|P_{1:d} \cup Q_{1:d}|} \tag{5}$$

Figure 3 shows the value of RBO between the two rankings at the different values of $P$. The proposed algorithm can be seen producing accurate rankings with the ground truth for the high ranks.

## 5. Conclusion

In this work, a methodology based on content analysis for ranking harmful Tor hidden services dealing with illicit drugs is proposed. A metric is defined to calculate the harm score for the HS based on the different types of illicit drugs present on the HS. The harm score is then used to generate the overall rankings of the HS on the Tor dark web dataset. In order to assess the accuracy and correctness of the proposed ranking methodology, we created the ground truth for the dataset with the help of experts. The standard metrics used for evaluation indicate the good performance of the proposed method in ranking highly harmful HS. The top-ranked HS can then be put under greater monitoring by the law enforcement agencies.

In future works, the proposed methodology can be strengthened by using more sophisticated NLP methodology as introduced in Ref. [41] for identifying illicit drugs code words. Moreover, other factors, like the trustworthiness and usability of HS, can also be quantified to assess the impact of HS.

## References

[1] C. Guitton, A review of the available content on Tor hidden services: the case against further development, Comput. Hum. Behav. 29 (6) (2013) 2805–2815, https://doi.org/10.1016/j.chb.2013.07.031.

[2] M. Faizan, R.A. Khan, Exploring and analyzing the dark web: a new alchemy, First Monday 24 (5) (2019), https://doi.org/10.5210/fm.v24i5.9473.

[3] I. Biryukov, F. Thill Pustogarov, R.-P. Weinmann, Content and popularity analysis of Tor hidden services, in: ICDCSW' 14 Proceedings of the IEEE 34th International Conference on Distributed

Computing Systems Workshops, 2014, pp. 188–193, https://doi.org/10.1109/ICDCSW.2014.20 (accessed 23 April 2019).

[4] R. Graham, B. Pitman, Freedom in the wilderness: a study of a Darknet space, Convergence (2018), https://doi.org/10.1177/1354856518806636/.

[5] D. Moore, T. Rid, Cryptopolitik and the Darknet, Survival 58 (1) (2016) 7–38, https://doi.org/10.1080/00396338.2016.1142085.

[6] J.K. O'Donnell, J. Halpin, C.L. Mattson, B.A. Goldberger, R.M. Gladden, Deaths involving fentanyl, fentanyl analogs, and U-47700—10 states, July-December 2016, MMWR 66 (2017) 1197–1202.

[7] L. Burns, A. Roxburgh, R. Bruno, J. Van Buskirk, Monitoring drug markets in the Internet age and the evolution of drug monitoring systems in Australia, Drug Test. Anal. 6 (7–8) (2014) 840–845.

[8] J. Martin, Lost on the Silk Road: Online drug distribution and the 'cryptomarket', Criminol. Crim. Just. 14 (3) (2014) 351–367.

[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, 2013, pp. 213–224, https://doi.org/10.1145/2488388.2488408 (accessed 13 November 2019).

[10] K. Kruithof, J. Aldridge, D. Décary-Hétu, M. Sim, E. Dujso, S. Hoorens, Internetfacilitated Drugs Trade, in: RAND Corporation, 2016, pp. 21–32.

[11] D. Van der Gouwe, T.M. Brunt, M. van Laar, P. van der Pol, Purity, adulteration and price of drugs bought on-line versus off-line in the Netherlands, Addiction 112 (4) (2017) 640–648.

[12] D. Rhumorbarbe, L. Staehli, J. Broséus, Q. Rossy, P. Esseiva, Buying drugs on a Darknet market: a better deal? Studying the online illicit drug market through the analysis of digital, physical and chemical data, Forensic Sci. Int. 267 (2016) 173–182.

[13] A. Bancroft, P.S. Reid, Concepts of illicit drug quality among darknet market users: purity, embodied experience, craft and chemical knowledge, Int. J. Drug Policy 35 (2016) 42–49.

[14] J. Martin, J. Cunliffe, D. Décary-Hétu, J. Aldridge, The international darknet drugs trade-a regional analysis of cryptomarkets, in: R.G. Smith (Ed.), Organised Crime Research in Australia 2018. Australian Institute of Criminology Research Reports, Australian Institute of Criminology, 2018, pp. 95–103.

[15] I. Ladegaard, Instantly hooked? freebies and samples of opioids, cannabis, MDMA, and other drugs in an illicit E-commerce market, J. Drug Issues 48 (2) (2018) 226–245.

[16] M. Paquet-Clouston, D. Decary-Hetu, C. Morselli, Assessing market competition and vendors' size and scope on AlphaBay, Int. J. Drug Policy 54 (2018) 87–98, https://doi.org/10.1016/j.drugpo.2018.01.003.

[17] E. Wadsworth, C. Drummond, P. Deluca, The dynamic environment of crypto markets: the lifespan of new psychoactive substances (NPS) and vendors selling NPS, Brain Sci. 8 (3) (2018) 46.

[18] Tor Project, Available at: https://2019.www.torproject.org/docs/onion-services/, 2019 (Accessed 23 December 2019).

[19] A. Celestini, G. Me, M. Mignone, Tor marketplaces exploratory data analysis: the drugs case, in: H. Jahankhani (Ed.), Global Security, Safety and Sustainability – The Security Challenges of the Connected World. ICGS3 2017. Communications in Computer and Information Science, Springer, 2016.

[20] M.W. Al-Nabki, E. Fidalgo, E. Alegre, L. Fernandez-Robles, ToRank: identifying the most influential suspicious domains in the Tor network, Expert Syst. Appl. (2019), https://doi.org/10.1016/j.eswa.2019.01.029.

[21] D.B. Skillicorn, Applying interestingness measures to ansar forum texts, in: Proc. of the ISI-KDD, 2010, 2010, pp. 7:1–7:9.

[22] G. L'Huillier, S.A. Ríos, H. Alvarez, F. Aguilera, Topic-based social network analysis for virtual communities of interests in the dark web, in: Proc. of the ISI-KDD, 2010, 2010, pp. 9:1–9:9.

[23] C.C. Yang, X. Tang, B.M. Thuraisingham, An analysis of user influence ranking algorithms on dark web forums, in: Proc. of the ISI-KDD, 2010, 2010, pp. 10:1–10:7.

[24] M. Bernaschi, A. Celestini, S. Guarino, F. Lombardi, Exploring and analyzing the tor hidden services, Graph. ACM Trans. Web 11 (4) (2017), https://doi.org/10.1145/3008662.

[25] M.W. Al-Nabki, E. Fidalgo, E. Alegre, I. de Paz, Classifying illegal activities on tor network based on web textual contents, in: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers. Vol. 1, 2010, pp. 35–43.

[26] DEA, available at: https://www.dea.gov/, 2019 (accessed 15 October 2019).

[27] D.J. Nutt, L.A. King, L.D. Phillips, Drug harms in the UK: a multicriteria decision analysis, Lancet 376 (2010) 1558–1565.

[28] J. van Amsterdam, D. Nutt, L. Phillips, W. van den Brink, European rating of drug harms, J. Psychopharmacol. 29 (6) (2015) 655–660.

[29] DUTA-10K, available at: http://gvis.unileon.es/dataset/duta-darknet-usage-text-addresses- 10k/, 2019 (accessed 23 December 2019).

[30] H.D. Kim, C. Zhai, J. Han, Aggregation of multiple judgments for evaluating ordered lists, in: C. Gurrin (Ed.), Advances in Information Retrieval. ECIR 2010. Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2010.

[31] M.G. Kendall, The treatment of ties in ranking problems, Biometrika 239–251 (1945).

[32] W. Webber, A. Moffat, J. Zobel, A similarity measure for indefinite rankings, ACM Trans. Inf. Syst. 28 (4) (2010) 20.

[33] Python Software Foundation, Python Language Reference, version 3.6. available at: http:// www.python.org/ (accessed 15 September 2019).

[34] G. Me, L. Pesticcio, P. Spagnoletti, Discovering hidden relations between tor marketplaces users, in: 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress, 2017, 2017, pp. 494–501.

[35] Q. Wang, Y. Jin, S. Cheng, T. Yang, ConformRank: a conformity-based rank for finding top-k influential users, Phys. A (2016), https://doi.org/10.1016/j.physa.2016.12.040.

[36] J. Dai et al., Identifying influential nodes in complex networks based on local neighbor contribution, IEEE Access 7 (2019) 131719–131731, https://doi.org/10.1109/ACCESS.2019.2939804.

[37] N. Wang, Q. Sun, Y. Zhou, S. Shen, A study on influential user identification in online social networks, Chin. J. Electron. 25 (3) (2016) 467–473.

[38] C. Mao, W. Xiao, A Comprehensive algorithm for evaluating node influences in social networks based on preference analysis and random walk, Complexity (2018), https://doi.org/10.1155/2018/1528341.

[39] C. Croux, C. Dehon, Influence functions of the Spearman and Kendall correlation measures, Stat. Methods Appl. 19 (2010) 497–515, https://doi.org/10.1007/s10260-010-0142-z.

[40] G.S. Shieh, A weighted Kendall's tau statistic, Statist. Probabil. Lett. 39 (1) (1998) 17–27.

[41] R. Bhalerao, M. Aliapoulios, I. Shumailov, S. Afroz, D. McCoy, K. Levchenko, Mapping the Underground: Supervised Discovery of Cybercrime Supply Chains. arXiv:1812.00381, Proceedings of APWG eCrime, 2019.

**Corresponding author**
Mohd Faizan can be contacted at: imfaizan15@gmail.com