The Impact of Adverse Events in Darknet Markets: an Anomaly Detection Approach

*

1st Ziauddin Ursani
Department of Computer Science
University of Bristol
Bristol, United Kingdom
zia.ursani@bristol.ac.uk

2nd Claudia Peersman
Department of Computer Science
University of Bristol
Bristol, United Kingdom
claudia.peersman@bristol.ac.uk

3rd Matthew Edwards
Department of Computer Science
University of Bristol
Bristol, United Kingdom
matthew.john.edwards@bristol.ac.uk

4th Chao Chen
Department of Computer Science
University of Bristol
Bristol, United Kingdom
chao.chen@bristol.ac.uk

5th Awais Rashid
Department of Computer Science
University of Bristol
Bristol, United Kingdom
awais.rashid@bristol.ac.uk

***Abstract*—In this paper, the notion of anomaly detection is introduced for the first time in the area of darknet markets (DNMs). Our hypothesis is that like popular social media platforms DNMs also exhibit anomalous behaviour. However, we propose that the meaning of anomalies in DNMs differs from social media anomalies. The social media anomalies are a cause of threat to the real world, while DNM anomalies are caused by threats from the real world. We present an unsupervised learning method developed to detect anomalies. The model is based on a weighted sum of a feature set trained through an evolutionary algorithm. Our approach successfully identifies anomalies in 35 DNMs – both at the community level and at the level of its user types. Our analysis shows that most of the anomalies found align with well-known adverse events—either as a direct consequence or as a cascading effect of the root event. Moreover, the model identified additional anomalies, which we were able to link to other events through post hoc analysis. Furthermore, we show that the adverse event of market shutdown generates a two-pronged impact on the ecosystem, i.e., it not only triggers startups of new markets but it does also inflict anomalies to current markets which may become fatal in some cases. We conclude that this two-pronged impact can be exploited by law enforcement agencies to produce maximum disruption in DNMs.**

***Index Terms*—Darknet Markets, Anomaly Detection, Adverse Events, Unsupervised Learning, Evolutionary Algorithm.**

1. Introduction

Understanding the dark web ecosystem is critically important to cyber security and cyber safety aims, as it typically houses a variety of illegal services such as the distribution of child abuse media, hacking, drug trafficking and weapons proliferation [6, 21]. How one may intervene appropriately and effectively in this ecosystem remains a research area in its infancy. While law enforcement organisations have developed and employed interventions [13], the effectiveness of these interventions remains unknown, with mixed consequences [21]. Positive impacts include decreasing the number of illegal sites and identifying key offenders, which subsequently reduces the benefits of offenders to continue with this type of crime. However, there is a large issue regarding the displacement of users and vendors after site take-downs or seizures — known as the “whack-a-mole” problem — for example, after the seizure of Silk Road (SR), Silk Road 2.0 (SR2.0) was launched within a month [14]. It is not unusual for marketplaces, websites, or even botnets that have been seized or taken down to reappear promptly [9, 21]. For example, the authors of [9] described the Armenian police force arresting the hacker that controlled the Bredolab botnet. However, the effect of this was minimal, with servers reactivated and running two days after the initial seizure. Similarly, work by [8] found that the number of active vendors across various Darknet Markets (DNMs) dropped substantially after Operation Onymous. However, within a month the number of active vendors had almost returned to pre-Onymous levels. This demonstrates the necessity for research into mechanisms that can help understand and model the impact of such actions on DNMs in order to effectively measure such impacts, leading to more impactful interventions. The work presented here is motivated by this necessity.

DNMs exhibit abnormal behaviour when they are disturbed by adverse events. Some examples of adverse events are: the actions taken by law enforcement agencies against DNMs, such as the Silk Road Shutdown [29], including actions taken against its management, such as the arrest of Silk Road founder Ross Ulbricht [31] and its users (e.g., Silk Road Drug dealer Cornelis Jan “Maikel” Slomp) [24], internal fights among DNMs through rumouring, DDOS attacks or hacking (e.g., the compromise

NB: appendices, if any, did not benefit from peer review.
A preprint of this paper has been deposited on ArXiv.

of Silk Road 2 Escrow Accounts [20]) or any negative news about DNMs in the media, such as the Gawker Blog publication about Silk Road [5]. The adverse event directed against one forum can affect the behaviour of other fora too. It is important to understand this effect as this could help law enforcement agencies to design their actions in a way to bring about maximum disruption to darknet markets. Hence, for the purpose of this study, an anomaly represents the effect of an adverse event on the behaviour of DNM users, and its detection is the estimation or measurement of this effect.

In this paper, we introduce our anomaly detection approach not only to identify the adverse events which may or may not be known a priori but also to measure their impact on the activity of the DNM users. The approach is applied on a substantial number of datasets i.e., 35 DNM communities containing over 150,000 users [18]. More specifically, our key contributions are as follows:

- We present an unsupervised learning based anomaly detection approach, which trains a weighted sum model of the selected feature set by minimisation of the standard error of estimate against the data points. Hence, the resulting model is able to identify an anomaly not only for the whole community, but also for its user types.
- Our approach enables an automated analysis of the cascading impact of disruptive events, such as a site shutdown, on the darknet ecosystem. Our results show a two-pronged impact generating not only startups of new markets but also inducing anomalies in the existing markets, which even lead to market shutdown in some cases. Such cascading impacts can be modelled and exploited by law enforcement to bring about maximum disruption to DNMs.

To our knowledge, this is the first study that investigates the feasibility of applying anomaly detection to identify the disruptive or adverse events and model and analyse their impact on DNMs. Such an anomaly detection approach provides a bottom-up focus to detect not only the immediate impact of an adverse event, but also the cascading effects, enabling a better understanding of the consequences of such events both in *space*, i.e., spawning or shutdown of DNMs and *time*, i.e., longer term impacts on the survival of markets and engagement of their users in ongoing activities.

This paper is structured as follows. We provide an overview of the related work in Section 2. The anomaly detection approach is discussed in Section 3. We describe our experiments and results in Section 4. Finally, we conclude this study in Section 5.

2. Background and Related Work

We discuss related work in two domains: anomaly detection (section 2.1) and impact of adverse events in DNMs (section 2.2).

2.1. Anomaly Detection

The term anomaly detection has been used in network intrusion detection and also in social media. Below we describe each of these separately. Later, we will emphasise why anomalies in DNMs are conceptually different from these areas and why they require a different approach.

- The anomalies considered in the area of network and cyber security are Virus, Worm, Trojan, Denial of Service (DOS), Network Attack, Physical Attack, Password Attack, Information Gathering Attack, User to Root (U2R) Attack, Remote to Local Attack (R2L), and Prob [2]. To detect these anomalies three types of systems can be deployed i.e., Misuse based, Anomaly based and Hybrid. However, Anomaly based is most popular nowadays. Network anomalies are not only security based but also performance based. Focusing on security anomalies only, these can be classified as point anomalies, collective anomalies and contextual anomalies. Various methods have been developed to identify these anomalies. These methods can be classified into three broad categories: genetic algorithm GA-based [43], Artificial Neural Network [1] and Artificial Immune System (AIS-based) [7].
- The anomalies considered in the area of Social Media are bullying, terrorist attack planning, dissemination of mis- and disinformation, hoax and rumour spreading, etc. If a single individual is involved in these acts, it is categorised as point anomaly. If several individuals are involved, then it is categorised as group anomaly [50]. The techniques developed to detect these anomalies are broadly categorized as behavior based, structure based and spectral based [26]. The example of behavior based anomaly detection is content based filtering [47]. The structure based anomaly detection consists of link mining [30]. The spectral based techniques explore the spectral graph space by different measures such as eigenvalues [49].

In this study, we extend the notion of anomaly detection to DNMs. There are very few examples where anomaly detection methods are used on darknet data. Those detection methods are aimed at identifying the threats the darknet is posing to legal communities, such as detection of distributed scan attacks [15], identification of hacker threats [40] and exploring hacker assets [39]. However, we take the perspective that the DNMs posing a threat to legal communities is their "normal" behaviour, rather than anomalous behaviour. To our best knowledge, this is the first exploration of anomaly detection with regard to anomalous behaviour of DNM users in response to events considered a threat to DNMs themselves and not the opposite, which is in contrast with previous social media and network intrusion anomaly detection approaches. Table 1 contrasts the anomaly detection in DNMs and other platforms.

Anomalies in DNMs are reactive, i.e., a change in the behaviour of DNM users in reaction to events in the real world that are considered harmful to DNMs by the users. On the contrary, anomalies in other platforms are proactive, i.e., a change in the behaviour of the users aimed at generating events that are harmful to the real world. DNM anomalies are mostly activity based, while anomalies in other platforms are of several kinds, i.e.,

TABLE 1. ANOMALY COMPARISON BETWEEN DNM AND OTHER PLATFORMS.

Criteria	Anomalies in DNMs	Anomalies in other platforms
Type	Reactive	Proactive
Nature	Activity based	Behavioural, performance, structural and spectral based
Exclusive Features	Examples: Use of coding and hacking terms	Examples: login/logoff records, HTTP access records, file access records, and sentiment analysis

behavioural, structural, spectral and performance based. Only behavioural anomalies in social media can be partly (quantitative) activity based as in the case of DNMs. However, again those activities are based on some features that are exclusive to each side, i.e., the use of coding and hacking terms exclusive to DNMs, and login/logoff records, HTTP access records and file access records exclusive to social media. Due to these contrasts, anomaly detection models used in other areas cannot simply be transferred to DNMs.

2.2. Impact of Adverse Events in Darknet Markets

The context of adverse events in cryptomarkets has been studied mostly through the lens of evaluating the effectiveness of law enforcement actions. For example, the authors of [8] studied the long-term impact of one of the largest law enforcement actions taken against cryptomarkets, Operation Onymous, and found that the overall impact of the operation was limited in scope, with the underground marketplaces adapting and recovering from the shutdowns within one to two months. This is a developing area, with some contrasting findings. Work by [4], using a difference-in-difference analysis across three forums, suggested that arrests do have a dampening effect on trade which is not fully accounted for by migration to other markets.

Recent work compared law enforcement action to other forms of adverse event affecting underground markets. In the context of high-risk opioids such as fentanyl, work by [3] studied the effect of law enforcement actions alongside voluntary market closures and exit scams, stating that law enforcement actions can have a greater impact on opioid availability than other closures, primarily by encouraging self-regulation among surviving markets.

In most cases, prior work has focused on this problem by working from *known* adverse events to examine their impact on an outcome measure of choice. This study, however, presents a bottom-up approach by exploring the feasibility of using anomaly detection techniques to identify when adverse events have occurred – both known and unknown – and are causing upheaval that may be worth further study or a response from law enforcement/researchers.

3. Approach

The approach for anomaly detection used in network intrusion detection and social media is that a reference model to represent normal data is developed and then new observations are tested against that model. The new observations are considered "anomalous" if they deviate from the reference model beyond the threshold line. In the case of DNMs, no reference model is available. The models developed for the social media platforms [22] cannot be used here, for the reasons discussed in section 2.1. In this section, we describe our approach to developing a model on a set of DNM data.

3.1. Data

The data consists of darknet datasets and adverse events.

3.1.1. The DNM Dataset. For this analysis, we make use of over 2.5 million posts drawn from over 150,000 users from 35 cybercriminal communities, drawn from the DNM Corpus: a large dataset collected between 2013 and 2015 [18]. All the DNMs have English language as their main medium of communication. In particular, we targeted discussion fora within this collection, which acted as support areas for underground marketplaces dealing in a number of different illicit goods. Table 2 gives a breakdown of the data available for each community. Communities ranged from successfully established markets with thousands of users (though not all were always active posters) to small sites that never moved beyond a handful of initial users.

The raw data provided in [18] captures fora as scraped at several semi-regular intervals by the dataset curators. This leads to heavy redundancy within the data, as threads may be captured at multiple times. However, this redundancy is also useful, as it helps to guard against intermittent faults in the crawling process. Our approach to parsing the data takes a *latest-version-first* view – of all pages captured within the crawling process, we treat as canonical the most recent version, only parsing older pages where they were not captured in later scrapes. We note that capturing pages from older scrapes is an important step in handling this data, as many thousands of threads and user profile pages are not present at all in the most recent scrapes of each forum. Differences could be attributed to crawling failures in later scrapes, incomplete coverage as part of the crawling processes, or to administrator action in taking down or hiding discussion threads over time.

Parsing of the data proceeded in two stages within the scrape history of each community. First, user profile pages were processed to build up a dataset of users and associated information from their profile pages (e.g., PGP public keys, membership status). Next, discussion thread pages were parsed in order to associate posts (including textual content and metadata such as posting time, subforum, etc.) with the user that authored them. Where quotations of other users could be identified within the text of a user's post, these quotations were separated from the authored text, to avoid contamination of profiling analysis. It sometimes occurred that user profile pages were not

TABLE 2. BREAKDOWN OF THE COMMUNITIES TARGETED FOR THIS STUDY.

Community	Posts	Users
Silk Road 2	882,418	26,163
Silk Road	846,077	52,383
Evolution	509,225	33,743
Abraxas	276,300	1,607
Agora	84,914	6,153
Black Market Reloaded	80,467	7,006
Nucleus	65,175	9,478
The Hub	58,642	7,337
Pandora	49,023	8,729
Black Bank	32,817	2,381
The Majestic Garden	26,121	1,858
Utopia	14,458	4,392
Diabolus	11,456	2,151
Kingdom	10,285	856
Project Black Flag	6,131	330
Cannabis Road2	5,842	2,139
Cannabis Road3	4,905	1,903
Bungee54	3,325	1,510
Panacea	2,241	520
Tor Bazaar	2,205	902
The Real Deal	1,049	115
Hydra	937	276
Kiss	933	145
Andromeda	894	1,601
Outlaw Market	689	2,007
Revolver	660	85
Tor Escrow	490	294
Dark Bay	332	484
Doge Road	300	118
Darknet Heroes	190	793
Havana	181	77
Tom	144	4,120
Grey Road	43	24
Tortuga	37	7
Mr Nice Guy	25	6

captured in the scrapes due to sites protecting access to those pages, or where users were observed posting for whom no profile page had been seen (either due to people using guest accounts, or due to incomplete coverage of profile pages in the crawls). In these cases, new user entries were created on the fly during the second stage of parsing, using such metadata as was available about the author account from the post metadata. Finally, metadata about different "user types", such as "Senior Member", "Vendor", "Administrator", "Newbie", etc., was collected for each forum according to the forum registration information that was available. This was done to analyse the forum for anomalies, not only as a whole community, but also at the level of its user types.

We observed two kinds of user types in the darknet forums: (i) *group user types*, i.e., one title assigned to several members and (ii) *individual user types*, i.e., a title exclusive for one member only. A user type can be an individual user type in one market and a group user type in another market. Every member is assigned one title, therefore there is no overlap of members in user types. It should be noted that, in some cases, we found quite atypical attributes of individual user types. Table 3 provides a few examples of the different user types found in the DNM registration information. The division of user types help us to divide anomalies into group and point anomalies. All the anomalies in individual user types are point anomalies. We consider these anomalies as insignificant because they are outcome of individual action rather than the outcome of any adverse events. However, anomalies in group user types are mostly group anomalies and are given serious consideration in our analysis.

3.1.2. Adverse Events. As a baseline for evaluating our approach, we calibrated outcome measures against the impact of the adverse events (E2-E5) in the form of known law enforcement interventions shown in Table 4. Moreover, Table 5 provides a list of additional minor adverse events (E1, E6-E9) identified through a manual Internet search around the dates where unaccounted group anomalies were found. Unaccounted group anomalies refer to those anomalies that were flagged by our approach, but which could not be immediately linked to baseline adverse events in Table 4.

3.2. Feature List

The approach analyses the anomalies based on unit time of one calendar month. Since calendar months are not equal, and in order to maintain equal sample sizes, the feature values are normalised for a 30-day month. We chose a sample period of a calendar month because smaller samples are not consistent in their amount of activity and therefore are bound to generate false positive anomalies. As we explained in Section 3.1, we collected data for each DNM in two files: one file representing activity of users and another file representing their metadata. This helped us to group the data for each user type and at the community level. All features discussed below are countable, so their extraction did not warrant any additional technical difficulties. For each sample time, we extracted the following feature types:

- **Coding Terms:** This is a count of coding terms used in messages. The coding terms considered here are listed in [19]. The use of coding terms reflect the technical ability of users which comprises a very important part of user activity.
- **Hacking Terms:** This is a count of hacking terms found in messages. The hacking terms considered here are listed in [45]. Hacking is a very important service provided by darknets. Use of hacking terms reflects major activity in this area.
- **Attachments:** This is a count of attachments to the messages. This feature is used in DNMs to explain technical things which cannot be described via short messages (see [39]).
- **Quotations:** This is a count of quotations included in users' messages. Quotations are typically used to let other DNM users know the context of a message. This feature is also used in network analysis [37].
- **Number of posts:** This is an account of the number of messages the users post to interact with other users of the community.
- **Number of threads:** A thread represents a conversation covering a group of messages posted to discuss a particular question or statement. This feature is also used for network analysis [51].
- **Number of Active Days:** A forum is considered to have an active day if at least one message is posted to it on that day. The maximum number of

TABLE 3. User types in DNM Forums

Forum	Group User Type	Individual User Type
Hack hound	Advanced, Advanced Member, Banned, Beginner, Expert, Intermediate Member, Member, Newbie, Titleless[1]	[curtailed][2] Intelligence Service, Intermediate, RDG Soft products, Retarded, Suspended, Ub3rnoob
Cracking fire	Active Member, Banned, Cracking Team, Cracking Crew, Ex Staff, Guest, New Member, Titleless[1] Very New Member, Well-Known Member	Android Coder – Retired, CF. Cracker, cracking is my life, Fucker, GFX Expert, Hacking Team Br, Legacy, Moderator, Redyoh approved!, Retired, Risk is My Busines, Super Master, V.I.P Member
The Hub	Brand Spankin' New, Full Contributor, Hero Contributor, Jr. Contributor, Newbie, New Contributor, Senior Contributor, Titleless[1]	Full Member, Hero Member
Evolution	Administrator, Banned, Forum Moderator, Market Moderator, Member, Moderator, Public Relations, Troll, Vendor	Guest, Resident Medical Expert
Nucleus	Administrator, Banned, Member, Moderator	Guest, Scammer
Black Market Reloaded	!!!!!Scammer!!!!!, Administrator, Banned, BMR Vendor, Hero Member, Jr. Member, Member, Moderator, New Member, Sr. Member	Unregistered[3]
Bungie54	Administrator, Junior Member, Member, Newbie, Titleless[1]	Bastard Administrator, Bungee54 Team, Customer Support One, Customer Support Two, Moderator, Senior Member, Your worst nightmare
Abraxas	Administrator, Full Member, Hero Member, Newbie, Sr. Member, Vendor	Jr. Member
Black Bank	Administrator, Banned, Member, Newbie, Vendor	Moderator
Diabolus	Freedom Fighter, Global Moderator, Newbie, Silk Road Vendor, Titleless[1] We rise from the ash	Administrator, Destiny will guide me
Pandora	Administrator, Full Member, Hero Member, Junior Member, Newbie, New Newbie, Pandora Support, Sr. Member, Titleless[1] Vendor, We rise from the ash	Unregistered[3]
Tom	Newbie	Administrator, Global Moderator
Utopia	Administrator, Banned, Member, Moderator, Vendor	SR Moderator
The Darknet Heroes	Daemon, Heroes, Newbie, Root	Member
Havana	Administrator, Member, Titleless, Vendor	Banned

[1] These users do not have any title.
[2] This name is curtailed because it was very long.
[3] Not a member.

TABLE 4. List of known interventions by law enforcement

Adverse Event	Date of Event	Agency Involved	Event Breakdown
E2: Silk Road Shutdown	October 2013	FBI	Seizure of US$3.6 million of funds in escrow. Arrest of the founder and chief operator of the site, Ross Ulbricht (undercover name Dread Pirate Roberts) [29].
E3: Arrests of Silk Road Admins	December 20, 2013	FBI	Arrest of three admins – Two were working in Silk Road 2.0. The new Dread Pirate Roberts surrendered control of the site. Defcon took over the site and promised to bring it back to working order [31].
E4: Silk Road 2.0 escrow accounts compromised	February 13, 2014	Silk Road Rivals	Bitcoin in escrow accounts worth US$2.7 m, reported stolen [31].
E4: Operation Commodore	February 2014	Dutch police	Launch and closure of Utopia. Servers located in Germany. Five people were arrested [16].
E5: Operation Onymous	November 2014	Europol's EC3, FBI, ICE, HIS, Eurojust	410 hidden services including Silk Road 2.0 servers taken down. 17 vendors and administrators arrested. US$1 m Bitcoin, EUR 180K cash, drugs, narcotics, gold and silver seized [12].

TABLE 5. List of minor adverse events

Adverse Event	Date of Event	Agency Involved	Event Breakdown
E1: Silk Road Notoriety	June 2011	Gawker Blog	Publication of article in GAWKER blog on 1st of June 2011 US Senetor Charles Schumer asked federal authorities to shut down the market on 5th of June 2011. Publication of article in The Sydney Morning Herald on 12th of June 2011.
E6: Conviction of Silk Road Founder	February 2015	Federal Court Manhattan.	Ross Ulbricht got 30 years in prison Sentence.
E7: Arrest of two Federal Agents	March 2015	US District Court Northern California	Carl Mark and Shaun Bridges were arrested for working as informants of Ulbricht.
E8: Conviction of Silk Road Drug Dealer	May 2015	Fedral Court Chicago	Cornelis Jan "Maikel" Slomp sentenced to 10 years in prison.
E9: Ulbricht appeal denied	May 2017	US Court of Appeals for the 2nd Circuit	Ulbricht appeal was declined.

active days cannot go above number of days in the sample period (30 in our case).

- **Number of Active Users:** This feature is measured in different ways based on the degree of activeness of individual users for network analysis [51] and in the identification of roles [44]. However, since we are not performing any network analysis, we do not measure degree of activeness of users. We simply consider users as active users if they post a message during a sample period.
- **Number of Memberships:** Membership of the forum is a mandatory requirement for any user of the forum. This feature is used by the forum administration for self-regulation [48]. The number of memberships is the count of new members registered during a sample period.
- **Average Message Length:** This is an average number of characters in a message. This feature was extracted to measure users' engagement with the forum. Characters rather than words are used as units of length because words are unequal and messages are not properly punctuated in darknet datasets.

3.3. Model Description

The anomaly detection approach offers a model that can be used for training purposes in order to detect anomalies in a darknet forum. The model is characterised by feature values, and consists of their weighted sum. Weighted sum is a useful model, which can be used for classification [36], clustering [42], and regression [33] – herein used for regression with regard to prediction of the effect of adverse events for anomaly detection by exploiting the differences along the time series in activity of DNMs. The weighted sum model was selected, because it is the most widely used standard linear representative model for vast variety of real-world applications, such as mechanics [28], business [10], Chemistry [23], Decision Support Systems [32], and statistics [25].

The model is trained by optimising the weights of a feature set through minimisation of a standard error of the model obtained through linear regression. This is done through the application of an evolutionary algorithm. Evolutionary algorithms are a useful optimisation procedure and have been used previously to optimise weighted sum in other applications [38]. This algorithm has proven its efficiency to optimise weighted sum models, for example, in crop planning [41], environment and economy [27], and engineering [34]. Fig. 1 depicts a flowchart of the overall procedure.

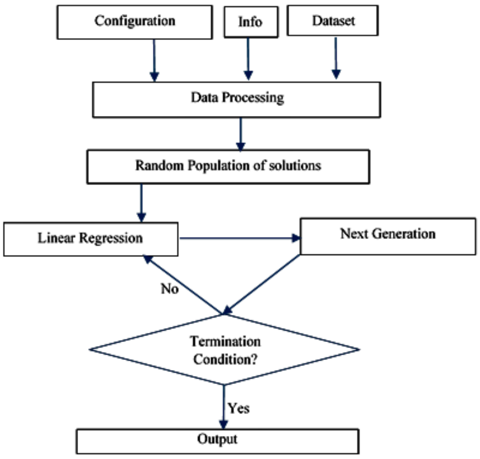

Figure 1. Flowchart of Anomaly Detection Algorithm.

The learning process starts by taking data from the forum. The input data includes a dataset, which consists of two files, one file representing the activity of users and another file representing users' metadata. The input also includes the configuration parameters, which consist of model type, feature types, sample size, threshold limit and output options. The input also includes other info such as a list of coding terms [19] and a list of hacking terms [45]. These lists are used to extract the terms present in the user messages. The input data is then processed, i.e., feature set samples are collected from the two files representing the dataset. The data is then used in the evolutionary algorithm (EA) which generates a random population of solutions, which means a random generation of weights of features. The weights are uniformly generated between 1.0 and 2.0. There was no upper limit on weights as far as evolutionary process is concerned. They could be evolved up to any limit. However, their lower limit was fixed to 1.0. This is because the evolutionary process was tailored to minimise the standard error. If weights were allowed to go below the limit of 1.0, all the weights would have evolved to 0.0 ending up in the model exactly over the x-axis with 0.0 standard error. Hence, controlling the lower limit of weights was essential. The EA then applies linear regression to each individual solution to compute slope, y-intercept and standard error. Next, the EA reproduces the next generation by applying the reproduction procedure on the most promising individuals – the solutions with comparatively lesser standard error among the individuals within the population. The reproduction procedure consists of the application of genetic operators, i.e., mutations and crossovers, over the promising solutions selected from the population. Linear regression is applied again on the new generation to produce the next generation. This iterative procedure continues until a termination condition is met. The termination condition is met when there is no improvement (no further minimisation of standard error) in any of the solutions within a population for a certain number of generations. This is a very efficient method to train the weights of the feature set. This trained set of feature weights represents a linear model of user activity of the forum. Any sample activity that deviates from this model, beyond a threshold limit, can be categorised as *anomalous activity*.

The activity of a sample is considered anomalous if it surpasses the threshold limit. The threshold limit is measured in terms of standard error units. If the weighted sum of a feature set which represents activity of the month is above two standard units away from the model estimate, then the sample value is beyond the threshold limit of the model value and it is considered anomalous. The threshold limit of two standard error units was chosen after experimentation with different DNMs, where we observed distance of data points from the model estimate against well-known adverse events. We observed that at two standard error units, the model yielded minimum false positive outliers and maximum true positive outliers. It should be noted that minor adverse events 5 were not part of this analysis. For the darknet fora with fewer samples standard error is normally large. Hence, anomalies are rarely found in such communities.

It should be noted that point anomalies are caused by the action or inaction of only one individual (see Section 3.1 and [50]). Therefore, they are not considered a consequence of any adverse event. However, group anomalies are caused by the collective action or inaction of individuals in a group, therefore they are considered a consequence of the adverse events. Hence, in experiments it is expected that point anomalies may not align with the dates of the adverse events in most of the cases. However, group anomalies should. Some exceptions can be allowed due to unexpected anomalies that can happen due to events not directly relevant to DNMs. The proposed model is designed to comply with this criterion, where fulfilment of this test can be considered a success of the model.

4. Experiments and Results

Our approach was applied on 35 DNM communities along with their several user types (see Section 3.1), which resulted in the identification of the 10 adverse events labeled (E1-E9) in Tables 4 and 5. There are nine labels for 10 events because two events have the same label (E4) due to their occurrence in the same sample period. The approach is based on unsupervised learning as described in Section 3.3.

In our first attempt, we tried different feature combinations and different weighted sum models, such as a model based on relative weighted sum as compared to past average. This model has been applied in different application areas such as text mining [46]. All these attempts were aimed at finding anomalies against baseline adverse events. However, we found that no specific combination of features and parameters could be achieved on the wide spectrum of datasets. Therefore, we decided to use all the features described in section 3.2 and their absolute weighted sum as a model to represent the amount of activity. We also decided to look for other adverse events which might be affecting the activities of users causing anomalies elsewhere. Furthermore, we investigated potential cascading effects, i.e., shutdowns and startups of DNMs triggering anomalies in other forums.

Figure 2 lists anomalies detected in all 35 communities, along with the 10 adverse events E1-E9. Events E2-E5 are baseline adverse events listed in Table 4 while the rest of the events are minor adverse events listed in Table 5. Two types of anomaly are represented: anomalous behaviour of the whole community is represented by the larger filled circles, while anomalous behaviour of one or more defined subgroups is represented by smaller points. The green circles represent anomalies that are directly aligned with either known or minor adverse events, whereas red circles represent anomalies not directly aligned with any adverse event. However, several of those are found aligned to some adverse event indirectly through a cascading effect. The startup and shutdown of a market place is represented through arrows <—> and the start and end point of data availability are represented through a vertical bar |. The dates of startups and shutdowns of market places were taken from [17]. The data available is not always in accordance with the startup and shutdown of a marketplace. This is because, in such cases, a forum starts before the start of a marketplace and ends after the end of a marketplace.

As can be seen in Figure 2, our model has detected anomalies that are aligned with the both major and minor adverse events (E1-E9). It can be seen that most of the anomalies are found in the context of major adverse events (E2-E5), while minor adverse events have lesser effect. This can be observed in the graph against minor adverse events (E6-E9), where anomalies are represented by the smaller circle. This shows minor events only cause anomalies in user types rather than in whole communities. It is also interesting to see that lots of startups and shutdowns are aligned with these adverse events. There are also several anomalies and startups and shutdowns that are not exactly aligned with these adverse events, e.g., anomalies, startups and shutdowns between E2 and E3. These can be explained by the delayed effect of adverse event E2. Similarly, anomalies, startups and shutdowns between events E3 and E4 can be seen as the delayed effect of events at E3 and a cascading effect of events at E2. This reasoning is based on the evidence from the output of our model.

As an example, Fig. 3 depicts the level of activity in terms of the weighted sum of the feature set in each

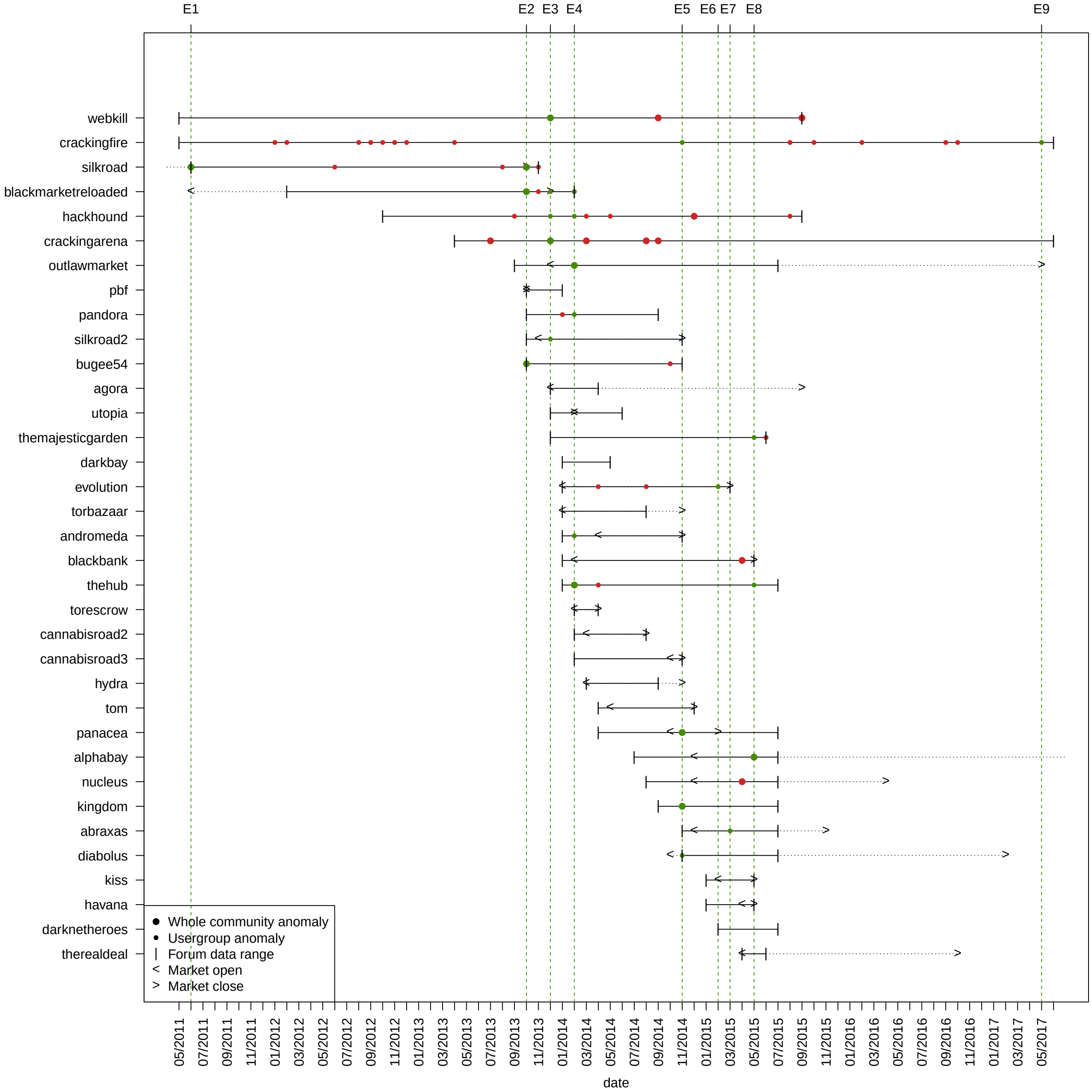

Figure 2. Anomalies detected within darknet markets, and known adverse events **E1**: Silk Road notoriety raised via GAWKER, Sen. Charles Schumer requests action from federal authorities. **E2**: Silk Road shutdown. **E3**: Arrests of Silk Road admins. **E4**: Silk Road 2.0 escrow compromise; Operation Commodore. **E5**: Operation Onymous. **E6**: Silk Road founder Ulbricht sentenced. **E7**: Arrest of federal agents bribed by Ulbricht. **E8**: Dutch Silk Road vendor sentenced. **E9**: Ulbricht appeal denied.

month for the Black Market Reloaded DNM. It shows a considerable increase in activity in October 2013. This aligns with event E2. Fig. 4 demonstrates the distance of activity points from the model in terms of standard error units. It can be seen that activity in October 2013 crosses the threshold line of two standard units. Therefore, this is anomalous activity caused by event E2. Table 6 represents the use of each feature during the anomalous month compared to its long term average. It can be seen that the largest increase is in the use of hacking terms i.e., 718%, followed by new memberships 549%, use of coding terms 515%, number of posts 500%, number of threads 464%, and number of active users 444%. This is an extraordinary increase in activity with the largest portion taken by use of hacking terms escalated by new memberships. This made administrators of Black Market Reloaded suspicious. In November 2013, they announced that the marketplace would be closed soon and advised members to close their crypto currency escrow accounts. Eventually, the market was closed in December 2013. This was a direct effect of event E2.

Additionally, the BMR shutdown has produced anomalies in Hack Hound, Cracking Arena and Silk Road 2.0, as shown in Fig 2. However, the event E3, i.e., the compromise of Silk Road escrow accounts also occurred during this month. Hence, this event may also have contributed to these anomalies. The data inside these anomalies may give some indication of their cause. According to this data, in Cracking Arena Market in December 2013, the number of memberships were increased by 254% compared to long term average. This effect can be attributed to both events, i.e. BMR shutdown and compromise of accounts in Silk Road 2.0. The members from BMR and Silk Road 2.0 may have joined Cracking Arena. In Hack Hound, the anomaly is only found in one user type (small circle). The membership increase here is 132%, which can also be caused by both events. Similarly, in Silk Road 2.0, there is 152% membership increase in its user type Newbie. It stands to reason that this is likely to be an impact of the BMR Shutdown.

As far as Event E3 is concerned it is unlikely that the compromise of escrow accounts could contribute to the increase in membership of the same market where the negative incident had happened. However, none of these anomalies were fatal enough to cause shutdown of these markets. The reason behind this is that Cracking Arena and Hack Hound were not Crypto Currency Markets and Silk Road 2.0 was a freshly opened market, so its administrators did not have any long term past record to judge changes in the data. Unfortunately, we do not have data of a well established crypto currency market during that time that could be tested for anomalies. However, a Europol report [11] shows that there was a well established market 'Buy It Now', at the time of the BMR shutdown, which voluntarily closed a couple of months after the BMR shutdown. Such volunteer shutdowns are a typical consequence of an unusual hype in memberships, as we already highlighted in the example of BMR at the time of Silk Road Shutdown. It is likely that the 'Buy it Now' shutdown is a cascading effect of Silk Road shutdown through anomaly infliction effect reaching 'Buy It Now' via anomalies in BMR.

The second impact of site shutdown, i.e., start up of new markets can also be witnessed from October 2013 (Silk Road Shutdown) and onwards in Fig. 2. Therefore, our analysis shows that the shutdown of any site has a two-pronged effect. It gives birth to new sites and it also produces anomalies in the current sites. The anomalies may cause shutdown of those sites and the shutdowns again, followed by the same two-pronged effect. This results in the chain reaction of startups and shutdowns which can be witnessed in the Fig. 2. Keeping in mind this observation, we have proposed some suggestions in the conclusion section for disruption of the markets by emulating effect of market shutdowns.

TABLE 6. ANOMALIES IN BLACK MARKET RELOADED (WHOLE COMMUNITY)

Feature Name	10/2013	Long Term Average
Number of Hacking Terms	423.871	51.8179
Number of Coding Terms	161968	26348.9
Average Message Length	347.277	335.547
Number of Posts	18982.3	3165.34
Number of Active Days	30	28.9457
Number of Attachments	0	0.0774194
Number of Threads	2060.32	365.568
Number of Memberships	1800	277.268
Number of Active Users	2559.68	470.32

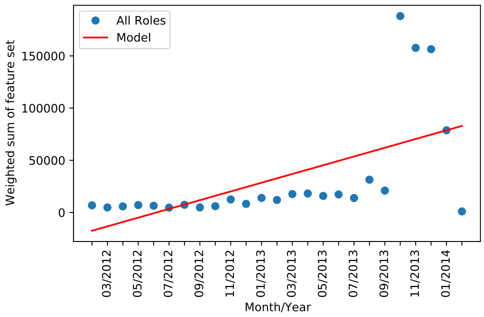

Figure 3. Graph of Black Market Reloaded (Whole Community).

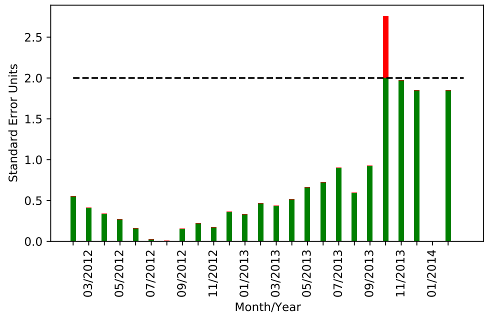

Figure 4. Barchart of Black Market Reloaded (Whole Community).

5. Conclusion and Future Work

The concept of anomaly detection is well established in the areas of network security and social networks. To the authors' knowledge this is first application of anomaly detection to DNMs to identify and understand the impact of adverse events. In so doing, a machine learning method has been developed which is based on unsupervised learning. The scope of the study is defined in respect to DNMs, adverse events and user types. An evolutionary algorithm is used to train the model, which is based on a weighted sum of a feature set. Features are taken from the literature, where they were used in different contexts.

The implementation of an anomaly detection approach is described and sample size and threshold limit are discussed. User types are also studied and categorised as group and individual user types. This categorisation helps us to differentiate between group and point anomalies, which greatly increases our ability to analyse the impact of adverse events by omitting point anomalies from analysis, which we believe are superfluous indicators. Therefore, we are able to concentrate on only the anomalies which are potentially the consequence of the adverse events.

We analysed the effectiveness of the approach by identifying anomalies and comparing them against well known adverse events where we could connect the timeline of anomalies with the timeline of adverse events across most of the forums and their user types. Where anomalies did not align with major adverse events, our manual search showed that other minor events (Table 5) may underlie those anomalies, indicating the value that an anomaly detection approach can bring to the analysis of DNMs. However, it was noticed that minor events mostly cause anomalies only in user types rather than whole community. Where our analysis identified anomalies that did not directly align with the timing of the events, we analysed if they represented cascading effects. We discovered a two-pronged effect of shutdowns which causes alternating startups and shutdowns of forums. This two-pronged effect of shutdown of market can be exploited by law enforcement agencies. While law enforcement agencies can do little about new startups, they can exacerbate anomalous effect of shutdown by introducing fake memberships and use other disruption techniques like rumours, spam and DDOS attacks.

Our current approach is based on user types which depends on titles assigned to users during their registration. These titles sometimes do not reflect the actual roles users play in the market. Our future work will focus on refining the the anomaly detection approach by replacing user types with actual roles of the users which could be identified by utilising social network analysis techniques [35] which are currently under study. Further to this, the approach requires detailed analysis regarding parametric optimisation with specific emphasis on smaller sample sizes to see the impact of adverse events at a micro level, in order to further our understanding of darknet ecosystem.

6. Acknowledgement

This work was supported by the UK Home Office through a National Cyber Security Programme research grant, *AMoC: Advanced Modelling of Cyber Criminal Careers*. The authors would also like to thank all members of the Advisory Group that have contributed to this project, our shepherd, Sergio Pastrana, and the reviewers for their thoughtful comments and efforts towards improving our manuscript[1].

References

[1] A. Alnafessah and G. Casale. "Artificial neural networks based techniques for anomaly detection in Apache Spark." In: *Cluster Comput 23, 1345–1360* (2020). DOI: https://doi.org/10.1007/s10586-019-02998-y.

[2] Monowar H Bhuyan, Dhruba Kumar Bhattacharyya, and Jugal K Kalita. "Network Anomaly Detection: Methods, Systems and Tools". In: *IEEE Communications Surveys & Tutorials* 16.1 (2014), pp. 303–336.

[3] Roderic Broadhurst et al. "Impact of darknet market seizures on opioid availability". In: *Australian Institute of Criminology* (2021).

[4] Jason Chan et al. *Shedding Light on the Dark: The Impact of Legal Enforcement on Darknet Transactions*. Tech. rep. 2019.

[5] Adrian Chen. *The Underground Website Where You Can Buy Any Drug Imaginable*. Wired. June 2011. [Online]. URL: www.wired.com/2011/06/silkroad-2.

[6] Janis Dalins, Campbell Wilson, and Mark Carman. "Criminal motivation on the dark web: A categorisation model for law enforcement". In: *Digital Investigation* 24 (2018), pp. 62–71.

[7] S. Das, M. Gui, and A Pahwa. "Artificial Immune Systems for Self-Nonself Discrimination: Application to Anomaly Detection." In: *Advances of Computational Intelligence in Industrial Systems.* Ed. by Y. Liu et al. Vol. 116. Studies in Computational Intelligence. Berlin, Heidelberg.: Springer, 2008. DOI: https://doi.org/10.1007/978-3-540-78297-1_11.

[8] David Décary-Hétu and Luca Giommoni. "Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of Operation Onymous". In: *Crime, Law and Social Change* 67.1 (2017), pp. 55–75.

[9] Benoit Dupont. "Bots, cops, and corporations: on the limits of enforcement and the promise of polycentric regulation as a way to control large-scale cybercrime". In: *Crime, law and social change* 67.1 (2017), pp. 97–116.

[10] Moses Olabhele Esangbedo and A. Che. "Grey Weighted Sum Model for Evaluating Business Environment in West Africa". In: *Mathematical Problems in Engineering* 2016 (2016), pp. 1–14.

[11] Europol. "Drugs and the darknet: Perspectives for enforcement, research and policy". In: *European Monitering Center for Drugs and Drug Addiction* (2017), pp. 1–90.

[12] Europol. *Global Action Against Dark Markets: On Tor Network*. Europol. Nov. 2014 [Online]. URL: www.europol.europa.eu/newsroom/news/global-action-against-dark-markets-tor-network.

[13] Europol. *OPERATION ONYMOUS*. Europol. Nov. 2014. [Online]. URL: www.europol.europa.eu/activities-services/europol-in-action/operations/operation-onymous.

[14] Cyrus Farivar. *Just a month after shutdown, Silk Road 2.0 emerges*. July 2019. [Online]. URL: www.arstechnica.com/information-technology/2013/11/just-a-month-after-shutdown-silk-road-2-0-emerges.

[15] Yaokai Feng et al. "A behavior-based method for detecting distributed scan attacks in darknets". In: *Journal of information processing* 21.3 (2013), pp. 527–538.

[16] Swansea University Global Drug Policy Observatory. *Law enforcement is currently not the greatest threat to the survival of Darknet drug markets*. Global Drug Policy Observatory. May 2014. [Online]. URL: www.swansea.ac.uk/media/Law-enforcement-is-not-the-greatest-threat-to-survival-of-Darknet-drug-markets.pdf.

[17] Gwern. "Darknet Market mortality risks". In: *https://www.gwern.net/DNM-survival* ().

[18] Gwern Branwen. *Dark Net Market archives, 2011-2015*. July 2015. URL: www.gwern.net/DNM-archives.

[19] Hackterms. *All terms related to coding*. URL: www.hackterms.com/about/all/.

[20] Martin Horton-Eddison and Matteo Di Cristofaro. "Hard interventions and innovation in crypto-drug

[1] As part of the open-review model followed on WACCO this year, all the reviews for this paper are publicly available at https://github.com/wacco-workshop/WACCO

markets: The escrow example". In: *Policy Brief* 11 (2017).

[21] Alice Hutchings, Richard Clayton, and Ross Anderson. "Taking down websites to prevent crime". In: *2016 APWG Symposium on Electronic Crime Research (eCrime)*. IEEE. 2016, pp. 1–10.

[22] Meng Jiang. "Behavior Modeling in Social Networks". In: *Encyclopedia of Social Network Analysis and Mining*. Ed. by Reda Alhajj and Jon Rokne. New York, NY: Springer New York, 2017, pp. 1–11. ISBN: 978-1-4614-7163-9. DOI: 10.1007/978-1-4614-7163-9_110203-1. URL: https://doi.org/10.1007/978-1-4614-7163-9_110203-1.

[23] Robert Johansson et al. "Account for variations in the H2O to CO2 molar ratio when modelling gaseous radiative heat transfer with the weighted-sum-of-grey-gases model". In: *Combustion and Flame* 158.5 (2011), pp. 893–901. ISSN: 0010-2180. DOI: https://doi.org/10.1016/j.combustflame.2011.02.001. URL: https://www.sciencedirect.com/science/article/pii/S0010218011000423.

[24] Seidel Jon. *World's most prolific online drug dealer 'Supertrips' gets 10 years*. Chicago Suntimes. May 2015. [Online]. URL: www.chicago.suntimes.com/2015/5/28/18423277/world-s-most-prolific-online-drug-dealer-supertrips-gets-10-years.

[25] Behzad. Kamgar-Parsi, Behrooz. Kamgar-Parsi, and Menashe. Brosh. "Distribution and moments of the weighted sum of uniforms random variables, with applications in reducing monte carlo simulations". In: *Journal of Statistical Computation and Simulation* 52.4 (1995), pp. 399–414. DOI: 10.1080/00949659508811688. eprint: https://doi.org/10.1080/00949659508811688. URL: https://doi.org/10.1080/00949659508811688.

[26] Ravneet Kaur and Sarbjeet Singh. "A survey of data mining and social network analysis based anomaly detection techniques". In: *Egyptian Informatics Journal* 17.2 (2016), pp. 199–216. ISSN: 1110-8665. DOI: https://doi.org/10.1016/j.eij.2015.11.004. URL: https://www.sciencedirect.com/science/article/pii/S1110866515000651.

[27] R.T.F.A. King and H.C.S. Rughooputh. "Elitist multiobjective evolutionary algorithm for environmental/economic dispatch". In: *The 2003 Congress on Evolutionary Computation, 2003. CEC '03.* Vol. 2. 2003, 1108–1114 Vol.2. DOI: 10.1109/CEC.2003.1299792.

[28] V. P. K. Kunisetti, Raviteja Meesala, and Vinay Kumar Thippiripati. "Improvised predictive torque control strategy for an open end winding induction motor drive fed with four-level inversion using normalised weighted sum model". In: *Iet Power Electronics* 11 (2017), pp. 808–816.

[29] Wesley Lacson and Beata Jones. "The 21st Century DarkNet Market: Lessons from the Fall of Silk Road." In: *International Journal of Cyber Criminology* 10.1 (2016).

[30] Getoor Lise and Diehl Christopher P. "Link mining: a survey." In: *ACM SIGKDD Explorations Newsletter* 7 (2 December 2005), pp. 3–12. DOI: https://doi.org/10.1145/1117454.1117456.

[31] Linda Marric. *The Underground Website Where You Can Buy Any Drug Imaginable*. Newscientist. Mar. 2021. [Online]. URL: www.newscientist.com/article/mg24933260-400-silk-road-review-the-true-story-of-the-dark-webs-illegal-drug-market.

[32] Ammar Naufal, Amelia Kurniawati, and Muhammad Azani Hasibuan. "Decision support system of SMB telkom university roadshow location prioritization with weighted sum model method". In: *2016 2nd International Conference of Industrial, Mechanical, Electrical, and Chemical Engineering (ICIMECE)*. 2016, pp. 107–111. DOI: 10.1109/ICIMECE.2016.7910428.

[33] Fernandez-Gonzalez Pablo, Bielza Concha, and Pedro Larranaga. "Random Forests for Regression as a Weighted Sum of k -Potential Nearest Neighbors". In: *IEEE Access* 7 (2019), pp. 25660–25672.

[34] Sidhartha Panda. "Multi-objective evolutionary algorithm for SSSC-based controller design". In: *Electric Power Systems Research* 79.6 (2009), pp. 937–944. ISSN: 0378-7796. DOI: https://doi.org/10.1016/j.epsr.2008.12.004. URL: https://www.sciencedirect.com/science/article/pii/S0378779608003222.

[35] Ildiko Pete et al. "A Social Network Analysis and Comparison of Six Dark Web Forums". In: *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE. 2020, pp. 484–493.

[36] Anthony Quinn, Andrew Stranieri, and John Yearwood. "Classification for accuracy and insight: A weighted sum approach". In: *Proceedings of the sixth Australasian conference on Data mining and analytics-Volume 70*. Citeseer. 2007, pp. 203–208.

[37] Toms Rekˇsņ a. "Complex Network Analysis of Dark-net Black Market Forum Structure". In: *Master thesis Leiden University Student Repository* (2017).

[38] Wang Rui et al. "Localized Weighted Sum Method for Many-Objective Optimization". In: *IEEE Transactions on Evolutionary Computation* 22.1 (Feb. 2018), pp. 3–18.

[39] Sagar Samtani, Ryan Chinn, and Hsinchun Chen. "Exploring hacker assets in underground forums". In: *2015 IEEE international conference on intelligence and security informatics (ISI)*. IEEE. 2015, pp. 31–36.

[40] Sagar Samtani, Hongyi Zhu, and Hsinchun Chen. "Proactively Identifying Emerging Hacker Threats from the Dark Web: A Diachronic Graph Embedding Framework (D-GEF)". In: *ACM Transactions on Privacy and Security (TOPS)* 23.4 (2020), pp. 1–33.

[41] Ruhul Sarker and Tapabrata Ray. "An improved evolutionary algorithm for solving multi-objective crop planning models". In: *Computers and Electronics in Agriculture* 68.2 (2009), pp. 191–199. ISSN: 0168-1699. DOI: https://doi.org/10.1016/j.compag.2009.06.002. URL: https://www.sciencedirect.com/science/article/pii/S0168169909000969.

[42] Weiguo Sheng et al. "A weighted sum validity function for clustering with a hybrid niching ge-

netic algorithm". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 35.6 (2005), pp. 1156–1167.

[43] K.G. Srinivasa. "Application of Genetic Algorithms for Detecting Anomaly in Network Intrusion Detection Systems." In: *Advances in Computer Science and Information Technology.* Ed. by N. Meghanathan, N. Chaki, and D. Nagamalai. Vol. 84. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Berlin, Heidelberg.: Springer, Networks and Communications. CCSIT 2012. DOI: https://doi.org/10.1007/978-3-642-27299-8_61.

[44] Yuan Taiquan. "Identifying Hackers' Roles and Examining the Evolution Pattern of Hackers on Hacker Forums by Clustering and Semi-supervised Learning Algorithms". In: *Masters Thesis, University of Bristol* (2020).

[45] techopedia. *IT Terms tagged with 'Hacking'*. Mar. 2021. [Online]. URL: www.techopedia.com/dictionary/tags/hacking.

[46] Fritsche Ulrich and Puckelwald Johannes. "Deciphering Professional Forecasters' Stories - Analyzing a Corpus of Textual Predictions for the German Economy". In: *Macroeconomics and Finance Series, University of Hamburg, Department of Socioeconomics* (April 2018).

[47] M. Vanetti et al. "Content-Based Filtering in On-Line Social Networks." In: *Privacy and Security Issues in Data Mining and Machine Learning.* Ed. by C. Dimitrakakis et al. Vol. 6549. Lecture Notes in Computer Science. Berlin, Heidelberg.: Springer, 2011. DOI: https://doi.org/10.1007/978-3-642-19896-0_11.

[48] Frank Wehinger. "The dark net: Self-regulation dynamics of illegal online markets for identities and related services". In: *2011 European Intelligence and Security Informatics Conference*. IEEE. 2011, pp. 209–213.

[49] Xiaowei Ying, Xintao Wu, and Daniel Barbará. "Spectrum based fraud detection in social networks". In: *2011 IEEE 27th International Conference on Data Engineering*. 2011, pp. 912–923. DOI: 10.1109/ICDE.2011.5767910.

[50] Rose Yu et al. "A survey on social media anomaly detection". In: *ACM SIGKDD Explorations Newsletter* 18.1 (2016), pp. 1–14.

[51] Maryam Zamani et al. "Differences in structure and dynamics of networks retrieved from dark and public web forums". In: *Physica A: Statistical Mechanics and its Applications* 525 (2019), pp. 326–336.