Blackmarket-Driven Collusion Among Retweeters–Analysis, Detection, and Characterization

Hridoy Sankar Dutta and Tanmoy Chakraborty

***Abstract*— The growth of online social media has led to a huge increase in the number of users who want to share and publicize various kinds of information. Twitter, the most popular micro-blogging platform, has become a hotbed for users who are involved in different activities such as news publishing, job hunting, recruiting, advertising and publicity. `Retweeting` a tweet is a major action to broadcast a user's message out to millions of users. Retweet action has two major advantages: (i) gaining quick exposure to the content, and (ii) increasing likelihood of gaining new Twitter followers in return. The organic way of gaining a larger number of retweets is a time consuming process, which leads to the creation of unfair methods to gain retweets. Thus, Twitter users often approach various blackmarket services to gain retweets inorganically in a short duration. Blackmarkets spread their collusive ecosystem in such a way that Twitter is unable to detect them even after devoting significant effort to purge the platform off bots, trolls, and fake accounts. One major reason behind the evasion is that the collusive users involved in blackmarket services exhibit a mix of organic and inorganic behavior – they organically reweet some genuine tweets; at the same time, they inorganically retweet tweets submitted to blackmarket services. This paper is the first attempt to provide a thorough study of the collusive users involved in two types of blackmarket services – Premium and Freemium. We collect a novel dataset of collusive users comprising of users from both types of blackmarket services. We provide network-centric, profile-centric, timeline-centric and retweet-centric characteristics of these users and show how users involved in premium blackmarket services exhibit diverse behavior as compared to those involved in freemium services. We further employ human annotators to label collusive users into three types: bots, promotional customers, and normal customers. We then curate 63 novel features to run state-of-the-art classifiers in two settings – binary classification (collusive vs. genuine) and multi-class classification (bot, promotional customers, normal customers, and genuine users). Bagging achieves the best accuracy (macro F1-score of 0.892) in the former setting, whereas Random Forest outperforms others (macro F1-score of 0.791) in the latter setting. We also develop a chrome extension, `SCoRe++` which can detect collusive retweeters in real time.**

***Index Terms*— Retweets, collusion, blackmarkets, Twitter, online social networks.**

Manuscript received December 4, 2018; revised May 19, 2019 and September 23, 2019; accepted November 3, 2019. Date of publication November 13, 2019; date of current version January 27, 2020. This work was supported in part by the Ramanujan Fellowship (SB/S2/RJN-073/2017) and in part by the Early Career Research Award (ECR/2017/001691). This article was presented in part at the ASONAM [1]. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Qian Wang. *(Corresponding author: Hridoy Sankar Dutta.)*

The authors are with the Department of Computer Science and Engineering, IIIT Delhi, New Delhi 110020, India (e-mail: hridoyd@iiitd.ac.in; tanmoy@iiitd.ac.in).

This article has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/TIFS.2019.2953331

I. Introduction

Online Social Networks (OSNs) attract social entities (users or organizations) connected by social relationships such as followers, friends, etc. to establish a platform for information exchange. Statistics reported that in year 2018, 3.196 billion users were engaged with OSNs, which is a 13% increase over the statistics of 2017.[1] The major reason for the increase in attraction of users towards OSNs is the availability of online news, job hunting, recruitment, advertisement, networking opportunities, and many more. OSNs not only spread these messages but also help in targeting audiences that might have potential interest in the same. OSNs help in connecting individuals whose likes and dislikes are similar, thereby creating a network around those commonalities. One of the major advantages of Twitter over other social networks in gaining such attraction is the availability of a large number of third-party tools for accessing the service. Moreover, there exist numerous tools for users to grow their presence on Twitter and some additional features such as analyzing brand performance, click-though rate, etc. There exists a bunch of tools to boost online presence for businesses on Twitter.[2] Tools such as Hootsuite[3] provide a common dashboard of multiple accounts managed by a single Twitter user and allow an easy way to retweet a tweet from multiple accounts. Advanced tools such as Tweriod[4] suggest users the best time to publish tweets by analyzing the followers of an account and determining the best time to post a tweet.

The tremendous increase in the involvement of users in OSNs has led to the rise of many fraudulent and spamming activities. In case of Twitter, a tweet's size limit is only 280 characters (previously 140 characters). Thus, Twitter users involved in the promotion of large campaigns/events/websites have to cram their thoughts within the tweet size limit. This becomes even more problematic as most of the tweets related to promotions consist of URLs which usually contain 50-80 characters. Many of these users resort to using various blackmarket services for these promotions. Moreover, most of these promoters try to gain popularity and visibility from specific target audience, which in turn helps them gain more retweets/followers to their tweets/accounts. This leads to users choosing artificial ways (through blackmarket services) to

[1] https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/

[2] http://blog.digitalinsights.in/twitter-tools-for-business-boost-your-brand/0572587.html

[3] https://hootsuite.com

[4] https://www.tweriod.com/

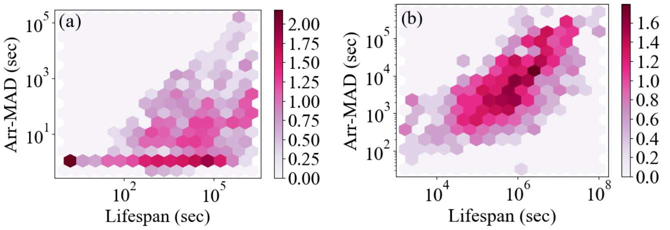


Fig. 1. (b) Collusive retweeters show asynchronous behavior as opposed to the (a) synchronous behavior shown by normal retweet fraudsters mentioned in Giatsoglou *et al.* [2].

boost their social growth. The benefits of using these services are manifold: (i) making tweets popular by providing retweets, (ii) gaining new followers, and (iii) boosting social credibility. The main attraction of these services is that they provide 'real' retweets/followers/likes, where 'real' means that the users involved in supporting these services are active Twitter users but not bots or inactive users.[5] These fake engagements on Twitter are strictly against the Twitter Terms of Service (TOS), and participation in these services may lead to account suspension.[6] To counter this, the blackmarket services promise to deliver their facilities in such a way that Twitter will not catch their activities. The users engaged in collusive activities are operated by regular human beings expressing a mixed behavior of organic and inorganic retweeting activities – on one hand, they organically retweet as genuine users, and on the other hand, being a part of blackmarket services, they inorganically and randomly retweet other tweets. We show how collusive users exhibit asynchronous behavior as compared to the synchronous behavior of normal retweet fraudsters. We examine two features as possible indicators of retweet fraud by projecting retweet threads of a user in a 2-D feature subspace - Lifespan (time elapsed between the first and last retweets) and Arr-MAD (mean absolute deviation of retweets inter-arrival time) [1]. In Fig. 1, two binned 2-D heatmaps (in log scale) of Lifespan vs. Arr-MAD of retweet threads are shown for (a) normal retweet fraudsters (data taken from Giatsoglou et al. [2]) and (b) collusive retweeters (on our collected dataset). We can easily recognize micro-clusters of fraudulent retweet threads (at very low values of Arr-MAD) for normal retweet fraudsters in Fig. 1(a); whereas in Fig. 1(b), the retweet threads of majority of collusive users are not concentrated on any area of the feature subspace, indicating the asynchronous behavior of such users [3], [4]. Thus, these users cannot be detected accurately by bot detection or fake user detection algorithms. Results in Section VI empirically validate our claim. A bunch of academic research on Twitter studied the problem of bots [5]–[7] and spam [8], [9]; but the problem of detecting collusive activities has not been deeply investigated yet. We use the term *collusive users* and *collusive retweeters* interchangeably in this paper, both referring to collusive users.

Here, we focus on analyzing two types of collusive retweeters collected from two different types of blackmarket services – Premium and Freemium (see Supplementary for a detailed study on blackmarket services).

– *Premium services* - These services provide retweets when they receive payment from customers.

– *Freemium services* - These services are free to use, but also provide some subscription plans for higher usage. In this paper, we restrict our study on freemium services to only credit-based services. In credit-based services, users retweet others to gain retweets on their content, a type of 'give and take' policy.

Contributions of the paper. We extend our previous work [1], where our major focus was to detect collusive users involved only in credit-based *freemium* blackmarket services. In this paper, our major aim is to provide an in-depth analysis of collusive users involved in *both premium and freemium* services and explore different facets of blackmarket services. In particular, the main contributions of this article are summarized below:

– We create a dataset of collusive retweeters collected from multiple premium and freemium blackmarket services. We further collect a set of genuine users who are experts in the domain of Machine Learning/Information Retrieval/Data Mining. Each collusive user in the dataset is labeled into one of the following types: *Bots*, *Promotional Customers* and *Normal Customers*.

– We analyze the activities of collusive retweeters (both premium and freemium) based on retweet-centric, network-centric, profile-centric and timeline-centric properties. The major findings are as follows. In the retweet-centric study, we observe that the credit system in freemium services is the primary reason for active participation of the freemium users. In the network-centric study, we observe that the premium services control a limited set of accounts which they use to provide retweets for their customers. On the contrary, freemium services accommodate a large set of registered customers who use the credit system to gain appraisals for the content the customers' submit. In the profile-centric study, we observe that freemium users are more involved in using social media tools such as blackmarket services owned APIs to perform the retweet operation, which in turn indicates that these users are more interested in gaining the credits rather than the content of the tweet. In the timeline-centric study, freemium users are more involved in quoting rather than retweeting tweets as compared to the premium users. The reason is that the freemium users perform the retweet operation using the APIs which embed their own content (hashtags/urls) into the tweets.

– We evaluate the performance of the state-of-the-art machine learning algorithms to distinguish users of blackmarket services from genuine users. We conduct experiments in two settings: multi-class (bots, promotional customers, normal customers, and genuine users) and binary-class (collusive and genuine). In the first setting, we achieve a macro F1-score of 0.791 with Random Forest classifier. In the second setting, Bagging turns out to be the best performer with a macro F1-score of 0.892.

– Finally, we develop a chrome extension `SCoRe++` to detect collusive users in real time. `SCoRe++` is an updated version of `SCoRe` [1] where we incorporate the features of

[5] See Supplementary for details on why Twitter users opt for artificial boosting services.

[6] https://help.twitter.com/en/safety-and-security/fake-twitter-followers-and-interactions

both premium and freemium users. Note that SCoRe only had features of freemium users. We apply these features in the backend to train our best performing binary classification model (Bagging in our case).

II. Related Work

In this section, we introduce relevant prior studies on the detection and characterization of collusive entities in OSNs. Most of the studies deal with detecting fraudulent users, spammer groups, fake followers, etc. to guide a wide range of problems such as information propagation, user behavior, malicious group detection, etc. However, these approaches are not suitable for the collusive retweeter detection as collusive users show a mixture of inorganic and organic behavior (see Fig. 1). In our work, we provide a detailed overview of the realm of previous studies from two different facets: (i) malicious activities in OSNs, and (ii) research on blackmarket services. In the latter part of this section, we mention how our method differs from the relevant existing studies.

A. Malicious Activities in Online Social Networks

There has been a considerable amount of studies on the detection of malicious activities in OSNs. Chu *et al.* [5] conducted a series of experiments to characterize the differences among the behavior of humans, bots, and cyborgs on Twitter. They proposed an automatic classification system using tweeting behavior, tweet content and account properties to distinguish among these types of users. Gianvecchio *et al.* [7] developed a classification system to detect bots and humans in Yahoo! chat. Davis *et al.* [6] proposed Botornot to detect bots in Twitter. It leverages more than a thousand features to check whether a Twitter account is controlled by a human or a bot. Retweets/followers are used to increase the visibility of a tweet/user. Moreover, the usefulness of retweets is not only limited to Twitter. Search engines such as Google ranks websites based on tweets containing website URLs.[7] In other words, if a tweet is published with a URL, and the tweet is retweeted several times, it affects the ranking scheme of search engine. Giatsoglou *et al.* [10] proposed NDSync and observed how retweet threads of fraudulent retweeters are clustered together. NDSync also automatically detects fake retweeter groups by assigning a suspiciousness score to each retweeter of the group. Giatsoglou *et al.* [2] proposed RT-GEN to imitate retweet patterns of both fake and genuine users. They reported that fake retweeters retweet concurrently in a lockstep manner.

Multiple studies have focused on the detection of anomalous activities on different social media platforms. Beutel *et al.* [11] proposed a graph-based approach to detect lockstep behavior on Facebook pages searching for non-bipartite cores in a bipartite graph between users and pages. Leskovec *et al.* [12] determined the signs of links on multiple OSNs (Epinions, Slashdot, and Wikipedia). Several papers [13]–[15] investigated the problem of review spam on e-commerce websites. Gupta *et al.* [16] studied phone number based spam attack on Twitter.

B. Research on Blackmarket Services

There has been relatively less work on investigating OSN based blackmarket services. Stringhini *et al.* [17] examined multiple Twitter follower markets to understand the growth and dynamics of customers of these markets. Shah *et al.* [18] analyzed behaviors of link fraudsters. They created multiple honeypot accounts and purchased followers from several follower markets. They proposed different follower-centric features to differentiate between fraudsters and genuine users. De Micheli and Stroppa [19] studied the pros and cons of the follower and retweeter based blackmarket services. They also reported that all interactions in Twitter can be easily counterfeited. Cresci *et al.* [20] identified 49 distinct features to distinguish accounts of human and fake followers. Thomas *et al.* [21] developed a classifier to detect fraudulent accounts that blackmarket merchants sold on Twitter. They even reported that around 10-20% of all the spam accounts originate from the blackmarket services. Singh *et al.* [22] studied the behavioral patterns of Twitter follower markets and developed an automated approach to classify these users as spammers. Liu *et al.* [23] detected crowdturfing following activities on Twitter (voluntary following). They termed these users as ‘volowers’. They further proposed DetectVC to detect volowers using graph-based features and prior knowledge collected from the follower markets. There have been some recent efforts on collusion in various social media platforms such as live video-streaming [24], [25] and online recruitment [26].

C. Differences With Our Previous Work

We extend our previous work [1], where our primary goal was to detect collusive users involved in credit-based freemium blackmarket services. We divided the freemium blackmarket services into three categories: social-share services, credit-based services, and auto-time retweet services. We collected and annotated collusive users into three types: bots, promotional customers and normal customers, and ran several supervised methods for classification.

This paper differs from the previous work from various angles:

- We provide a thorough study of various kinds of blackmarket services, which to our knowledge is the first attempt of this kind.
- We examine the diverse aspects of collusive retweeting activities on Twitter based on premium and freemium blackmarket services.
- We extend our human-annotated dataset of collusive users by adding the data of premium collusive users.
- We offer an exploratory analysis of the dataset of collusive users based on network-centric, profile-centric, timeline-centric, and retweet-centric observations. This analysis aims to differentiate premium and freemium blackmarket customers.
- We release a new version of our chrome extension, SCoRe++, by adding features calculated from both premium and freemium blackmarket services.

To summarize, previous studies provide a profound explanation of the methodologies related to the detection of fake

[7] https://blog.twitter.com/marketing/en_us/a/2015/new-measurement-and-buying-tools-through-doubleclick-coming-soon.html

OSN entities. But these studies tend to avoid equally important question of detecting *collusive entities*, which may not be same as fake entities. To overcome the limitations of the previous studies, we provide an in-depth analysis of the collusive users and propose an automated strategy to detect them.

III. Dataset Collection

In this section, we present our effort to collect a large set of Twitter accounts that actively interact with blackmarkets. The data forms the ground-truth on customers who bought retweets from the blackmarkets (*premium retweeters*) and victims who were compromised by the blackmarket and traded as retweeters (*freemium retweeters*). We also collected a set of genuine users to make the data balanced.

1) Creating Honeypot Accounts and Adding to Blackmarket Services: To analyze the behavior of the blackmarket services, we created one Twitter account (honeypot) per service. All the honeypot accounts were created on the same day with empty timelines (tweets/retweets) and empty social networking features (followers/friends) after careful IRB approval. Hence, we asserted that all the retweeters of these honeypots accounts would be from blackmarket services. We conducted an initial experiment on two new honeypot accounts (one premium and one freemium) to check whether they retweet fresh Twitter accounts or not. It was surprising to see that most of the freemium services do not allow to add accounts/tweets if there is no profile-related information such as user bios, profile image, inactive accounts. We next created a set of unique tweets which looked attractive (to attract more retweeters), and posted one tweet from each account. We next added each account/tweet to one blackmarket service. Summarily, we created 8 honeypot accounts and added 4 of them to premium services and the remaining 4 to freemium services. As premium services offer retweets in tiers, we selected the package which provides 100 retweets, and the retweeters delivery time is within 48 hours.

2) Creating Two Sets of Users for Supervised Classification: **Blackmarket customers.** On adding the tweets to blackmarket services (premium and freemium), we implemented several tracking scripts using Twitter API to monitor changes in the tweets. For each tweet submitted to any premium service, we collected retweeters' information at 10 minutes interval. Similarly, for tweets submitted to any freemium service, we visited the earning area of each freemium service at 6 hours interval per day and retweeted tweets of all other customers present in the dashboard to earn maximum credits. We also kept track of the retweets which were not present in the next time frame (deleted retweets). Summarily, we took four premium services, namely SocioCube (SC), RedSocial (RS), SocialShop (SS), and SocialKing (SK), and four freemium services, namely KingdomLikes (KL), YouLikeHits (YH), Traffup (TU), and Like4Like (LL) (see Supplementary for details on how we located the blackmarket services).

Genuine users. We also collected a set of genuine users to compare them with the users of blackmarket services. We started by selecting a set of known users who are guaranteed not to have participated in any form of collusive activities before. To increase the set of genuine users, we added some well-known machine learning and deep learning scientists as suggested in [18]. From the set of genuine users, we discarded those users who had more than 1 million followers since such high follower count may resemble a celebrity-like users. This may further create a bias in our dataset.

Table II shows the statistics of the dataset. Overall, we collected 475 customers from the premium services (C_p), 196 customers from the freemium services (C_f) and 1000 genuine users. We further extended our set of customers from the freemium services by adding 279 collusive users from our previous dataset [1]. Interestingly, out of the overall 950 collusive users ($C_p + C_f$), 13 users turned out to have verified Twitter accounts. This clearly shows that Twitter in-house algorithms have been unsuccessful to detect such users. We further extracted profile information and timelines of customers of both types of service and genuine users using Tweepy[8] library. The summary of the entire dataset of this work is presented in Table I.

IV. Human Annotation of Collusive Users

So far, we have collected collusive users from multiple premium and freemium blackmarket services. Each collusive user was then categorized by three human annotators[9] into any of three classes (bots, promotional customers and normal customers) based on the instructions given to them. The annotators were also given complete liberty to search for any information on the web for proper annotation of the users. We developed the following set of rules to define each type of collusive user:

1) **Bots:** Twitter bots are user accounts that can post/retweet tweets or communicate with other users in an automated way without direct human input. The rules that should be followed by a Twitter bot are mentioned in Twitter Automation Rules for Bots released by Twitter.[10]
2) **Promotional customers:** Promotional customers are engaged in campaigns to extend the scope of their content to a relevant audience on Twitter. These users often use keywords such as 'giveaway', 'free', 'win', 'gain' in their tweets in order to gain quick popularity. Twitter also has a set of rules that should be followed by a user for content promotion.[11]
3) **Normal customers:** For all the users who do not fall under the above two categories, we label them as normal customers.

Out of 475 collusive users collected from premium blackmarket services, we found 83 bots, 129 promotional customers and 263 normal customers. In case of freemium blackmarket services, out of 475 collusive users, we found 97 bots, 202 promotional customers and 176 normal customers. The inter-annotator agreement was 0.73 based on Fleiss' κ.

Next, we present numerous insights of blackmarket retweeters based on four properties: retweet-centric, network-centric, timeline-centric and profile-centric.

[8] http://www.tweepy.org/

[9] Annotators were experts in social media with age range between 25-35.

[10] https://digitalinspiration.com/twitter-bots-automation-rules-5066

[11] https://help.twitter.com/en/rules-and-policies/twitter-contest-rules

TABLE I

SUMMARY OF DATA COLLECTED FROM BLACKMARKET SERVICES

Type	Service	# retweets promised	# retweets delivered	# retweets deleted
Premium	SC	100	320	222
Premium	RS	100	60	16
Premium	SS	100	28	21
Premium	SK	100	67	36
Freemium	KL	–	46	2
Freemium	YH	–	74	10
Freemium	TU	–	13	0
Freemium	LL	–	63	11

TABLE II

DATASET STATISTICS

Type	# users	# verified users
Premium collusive	475	0
Freemium collusive	475	13
Genuine	1000	28

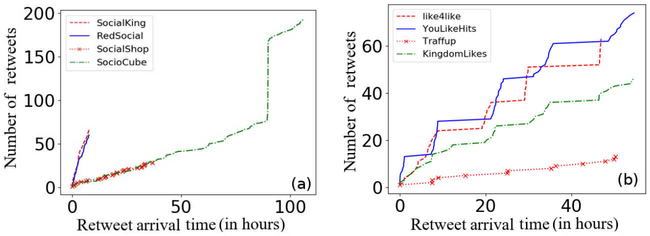


Fig. 2. Retweet arrival time - (a) premium and (b) freemium.

V. ANALYSIS OF BLACKMARKET RETWEETERS

We start our analysis on the data collected from four premium and four freemium blackmarket services. We discuss the retweet-centric, network-centric, timeline-centric and profile-centric characteristics of users.

A. Retweet-Centric Observations

1) Arrival Time of Customers: Here, we analyze the arrival time of customers for each of the blackmarket services. Fig. 2 shows the variation of the arrival of customers over time for different blackmarket services. Premium users have a steady increase whereas freemium users show stepwise behavior. One possible reason for such behavior in case of freemium services, is that the tweets of customers of these services only get retweeted, when the user has available credits in his/her account. A customer on adding his/her tweet to any of the credit-based freemium services has to set a credit/score for the tweet. When the tweet is retweeted by other customers, that score is deducted from the credit of the source user and is added to the credit of the retweeted user. When a user has no available credits, the freemium services stop working. The user can again earn credits by retweeting tweets of other customers in the earning area.

2) Active Time of Customers: The active time is measured as the number of retweets a tweet receives during a particular interval of a day. Fig. 3 shows the active time for the customers of the blackmarket services. Premium users are active almost every hour with a peak during office hours (9 a.m.-5 p.m.) whereas freemium users are active during non-office hours. One possible reason of premium services showing such behavior is that premium accounts are mostly controlled by the media marketing companies and not the owner himself.[12] The premium users are target-specific users (involved in brand endorsements, event campaigns, product launches, etc.), who want to entice more people to keep track of the company activities. The media marketing companies help the premium users to create eye-catching tweets to target the right audiences for faster channel growth and lead generation. On the other hand, customers of freemium services control their accounts and have to earn credits in order to get retweeted by other customers. Retweeting tweets of random users to gain credits is a tiresome job which requires immense manual effort and is usually done during free time.

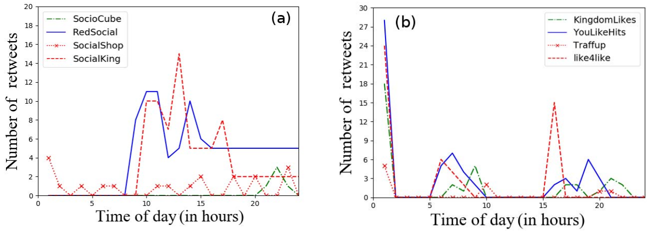


Fig. 3. Active time - (a) premium and (b) freemium.

3) Retweet Deletion Time of Customers: Even though the process of gaining artificial retweets in exchange of money or credits looks as simple as it sounds, customers tend to delete retweets after a certain point of time. Fig. 4 shows the number of retweets that we lost for a particular time interval. Premium retweeters tend to delete more retweets than freemium retweeters. The reason for such behavior by premium retweeters may be due to the irrelevancy of those tweets from all other tweets in their timeline. Moreover, deleting these retweets in quick time may help them evade the in-house fake detection algorithms deployed by Twitter. This also may be a trick deployed by the premium services to keep the customer in a vicious circle to buy more retweets from their service.

4) Followers Gained Upon Using Blackmarket Services: It is very interesting to see that customers who purchased retweets from blackmarket services also gain followers. As mentioned in Section III, we created all our honeypot accounts with empty social networking features (followers/followees). Fig. 5 shows the relation of the number of followers gained to the number of retweets for both premium and freemium blackmarket services. Though premium services do not help in gaining followers, it is very clear that customers of the freemium services gain followers on retweeting tweets of other customers. However, it is not guaranteed that these

[12] https://www.digitalmarketingagency.com/twitter-management-services/

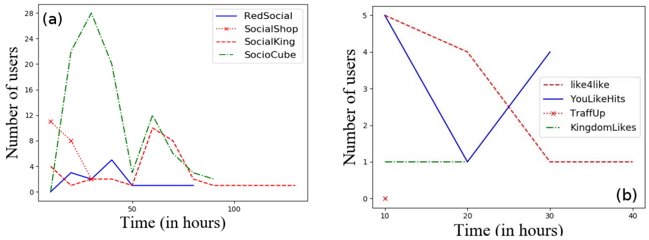


Fig. 4. Retweet deletion time - (a) premium and (b) freemium.

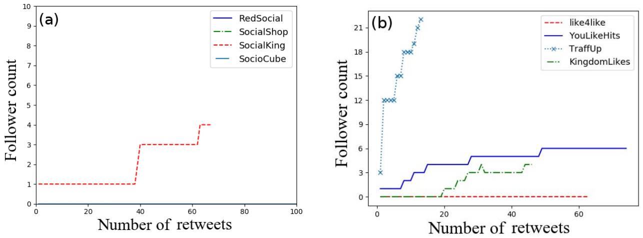


Fig. 5. Followers gained - (a) premium and (b) freemium.

followers belong to the blackmarket services or are real followers. We anticipate that some of the followers may be real users because the customers of the freemium services retweet tweets randomly without looking at the content which might create interest among genuine users following the real content to connect with these users.

B. Network-Centric Observations

In this section, we analyze the network generated from the data collected from the blackmarket services. We create a directed network for each type of services - nodes of each network are the honeypot accounts and the retweeters of their tweets, and edges represent retweets among these nodes. We use the metrics defined in [27], i.e., *density*, *transitivity* and *reciprocity* to describe the network structure. Here, *density* is the ratio of number of edges in the network to the maximum possible edges. *Transitivity* represents the clustering coefficient of the network. *Reciprocity* measures the bidirectional property of the network – value of 1 indicates that all the edges in the graph are bidirectional. Table III represents the network property of both types of collusive users involved in blackmarket services. Analysis of the network reveals noteworthy difference between users involved in premium and freemium services. Premium retweeters have high values of *density* and *transitivity* but low value of *reciprocity* as compared to freemium retweeters. Higher value of *density* and *transitivity* for premium service accounts indicates that the same limited set of accounts is used to provide retweets to users that avail the services, as opposed to freemium services where the retweets are provided by a larger number of regular users. Higher value of *reciprocity* for freemium retweeters is expected as customers of these services are involved in a 'give and take' relationship where users retweet others to gain retweets on their own content. Fig. 6 shows the visualization of the network formed by our honeypot accounts and their retweeters in case of premium and freemium blackmarket services. It represents how each user in the network is retweeted by other users. We observe that premium users are retweeted by a limited set of users. However, in the case of freemium users, there is no such retweeting pattern among users.

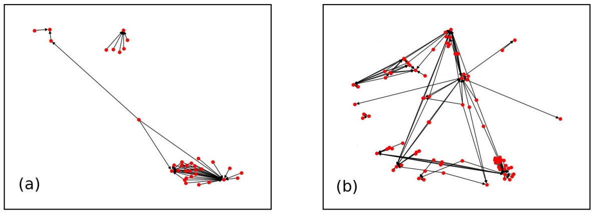


Fig. 6. Network visualization - (a) premium and (b) freemium.

TABLE III

SUMMARY OF THE NETWORK STATISTICS OF PREMIUM AND FREEMIUM BLACKMARKET SERVICES

Property	**Premium**	**Freemium**
Density	0.038	0.017
Transitivity	0.338	0.126
Reciprocity	0.348	0.383

C. Profile-Centric Observations

1) User Activeness per Day/Hour: Blackmarket users are active throughout the week publishing tweets/retweeting other content. Fig. 7 shows the activeness of both types of blackmarket users: (a) per day basis and (a) per hour basis. Both premium and freemium users are active equally on every day of a week. However, we see an increasing trend of activeness for these users on normal office hours (11 a.m. - 5 p.m.) as shown in Fig. 7.

2) Wordcloud of User Description/Bios: We also study the 160-character user description field for each of the blackmarket users. Users of the blackmarket services use targeted words to attract more followers/retweeters. Fig. 8 shows the wordcloud of the description/bios text for premium and freemium users generated after removing two-letter words and common stopwords. Here, the font size corresponds to the frequency of the text. We find that premium users have words in the description such as 'CEO', 'official', 'speaker', 'Founder'. These words appear to be associated with high profile accounts. However, freemium users have keywords such as 'like', 'agency', 'SocialMedia', 'YouTubeMarketing'. Both premium and freemium users have many keywords related to BitCoin such as 'blockchain', 'crypto', 'btc'. One possible reason of the presence of such keywords for freemium services could be the time when our dataset was collected; the news of the rise and fall of Bitcoin was primarily driven by OSNs.

3) Language Used by Blackmarket Retweeters: We study the language of each tweet/retweet in the timeline of blackmarket retweeters. Table IV lists the top 5 languages used by premium and freemium retweeters with the percentage of tweets for each language. We find that more than 50% of tweets posted/retweeted by both types of users are in English. Russian (Spanish) ranks second with 19.17% (10.46%) of all content being posted in the language by Premium (Freemium) users. Surprisingly, we notice around 9% users of both types whose language type is undetected by Twitter API. The primary aim of collusive users is to polarise discussion, influence and mislead users, boost campaigns, etc. where the primary mode of communication is in English.

4) API Clients Used by Blackmarket Retweeters: Twitter API can be accessed by an individual using various third-party APIs. As different blackmarkets promise that all their users are

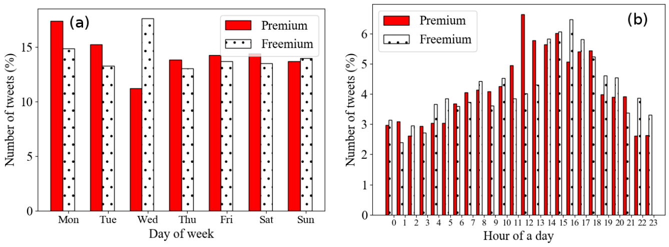


Fig. 7. Number of tweets (a) per day of a week and (b) per hour of a day.

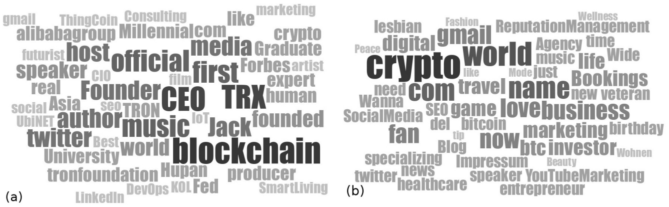


Fig. 8. Wordcloud of profile description - (a) premium and (b) freemium.

genuine users with activities resembling a normal Twitter user, they normally use web or mobile (Android/iPhone) interface to connect to Twitter. Table V shows the overall application usage for top 10 clients by premium and freemium blackmarket users. Unsurprisingly, we observe that freemium users use Twitter with the API provided by different blackmarket services, social media marketing websites, etc. Around 0.09% of tweets of freemium users are from the blackmarket services owned APIs such as *Easy Retweet API*, *NotFollow.me en Espaol* and *Pay with a Tweet*.

5) Account Dormancy: To measure account dormancy, we calculate the time difference between the creation date of the account and the first activity of the account. Unsurprisingly, 71.6% of the premium retweeters and 55% of the freemium retweeters perform their first activity on the day the account was created. However, 19.5% of premium retweeters and 17.5% of freemium retweeters remain inactive for atleast a month after creating their account.

6) URLs Shortening Services: Blackmarket users also use a set of URLs to promote their contents. Various URL shortening services such as *bit.ly*, *tinyurl.com* are used by the users to advertise/promote contents. Thomas *et al.* [28] mentioned that spam URLs are commonly shortened as compared to non-spam URLs. We find that 5.8% of tweets of premium users are from URL shortening services with 2.1% from *bit.ly*. Similarily, we find 5.1% of tweets of freemium users from URL shortening services with a similar percentage of tweets from *bit.ly*. To identify whether these shortened URLs are malicious or not, we run PhishTank[13] for each of the expanded URLs generated from shortened URLs. We observe a very small percentage of tweets (premium-0.0012% and freemium-0.0019%) detected by PhishTank as suspected phises. Although the URL shortening services deploy strong spam filtering algorithms in the backend, the presence of alternate shortening services is used to create multiple redirects to finally lead to the expanded URL.

TABLE IV
TOP 5 LANGUAGES USED BY PREMIUM AND FREEMIUM RETWEETERS

Premium	% of tweets	Freemium	% of tweets
English (en)	51.41	English (en)	67.21
Russian (ru)	19.17	Spanish (es)	10.46
No language detected (und)	9.05	No language detected (und)	9.54
Spanish (es)	4.40	Arabic(ar)	1.64
Portuguese (pt)	3.48	Portuguese (pt)	1.46

TABLE V
TOP 10 CLIENTS USED BY PREMIUM AND FREEMIUM USERS

Premium	% of tweets	Freemium	% of tweets
Twitter Web Client	92.44	Twitter Web Client	75.24
Twitter for iPhone	1.62	Twitter for Android	9.19
Google	1.43	Twitter Lite	5.93
Mobile Web (M2)	0.80	Google	3.94
Facebook	0.77	Buffer	1.10
TweetDeck	0.76	Twitter for iPhone	0.94
vk.com pages	0.60	TweetCaster for Android	0.86
GTX1024	0.38	Twitter for Nokia S40	0.78
Twitter Lite	0.31	Facebook	0.61
Instagram	0.20	TweetDeck	0.27
Twitter for Websites	0.19	Mobile Web	0.24

D. Timeline-Centric Observations

1) Embedded Mentions, Symbols and Urls: We detect the number of tweets in the timeline of each collusive user with no mentions, symbols or URLs. In case of premium users, we find 7.7%, 99.7%, 73.2% tweets to have no mentions, symbols and URLs, respectively. However, in case of freemium users, we find 37.8%, 99.2%, 46.9% tweets with no mentions, symbols and URLs, respectively. We observe that both types of users do not use many symbols in the tweets. Nevertheless, freemium users are more involved in sequence of mentioning other users in their tweets. The reason for such behavior by freemium users is that retweeting targeted tweets (tweets with mentions to other users) may help them gain free followers, thereby increasing social popularity.

2) Quoted Tweets: We also find the number of quoted tweets for both types of users. We find that only 1.56% of the tweets by premium users and 2.57% of the tweets by freemium users are quoted tweets. The reason for the higher value for freemium users can be because these users use third-party API to access Twitter. Using third-party API to retweet a tweet may add new entities (hashtags, mentions) to the retweet, making it a quoted tweet. Services providing third-party API will be interested to promote themselves by adding hashtags related to their services.

VI. CAN MACHINE LEARNING DETECT BLACKMARKET USERS?

All the analysis detailed in this paper so far are based on the data collected from the premium and freemium blackmarket services. However, one might be interested to devise an automated technique to categorize different types of users involved in blackmarket agencies. In this section, we describe features used for our experiments and the details of our experimental setup. Further, we analyze the experimental results and discuss the importance of the features. We use 63 features and divide them into two broad categories – profile-centric features and

[13] PhishTank is an anti-phishing site https://www.phishtank.com/

TABLE VI

PERFORMANCE OF DIFFERENT SUPERVISED CLASSIFIERS FOR MULTI-CLASS CLASSIFICATION

Classifier	**Micro**				**Macro**			
	Precision	**Recall**	**F1**	**ROC-AUC**	**Precision**	**Recall**	**F1**	**ROC-AUC**
Decision Tree	0.715	0.715	0.715	0.811	0.712	0.719	0.714	0.829
K-NN	0.627	0.627	0.627	0.739	0.584	0.567	0.569	0.669
Logistic Regression	0.662	0.662	0.662	0.749	0.622	0.585	0.584	0.671
Naive Bayes	0.534	0.534	0.534	0.631	0.529	0.495	0.494	0.592
SVM	0.323	0.323	0.323	0.441	0.329	0.256	0.133	0.391
Random Forest	**0.785**	**0.785**	**0.785**	**0.844**	**0.792**	**0.787**	**0.791**	**0.849**
Bagging	0.782	0.782	0.782	0.830	0.781	0.788	0.784	0.831
Boosting	0.323	0.323	0.323	0.441	0.329	0.256	0.133	0.391
MLP	0.708	0.708	0.708	**0.923**	0.737	0.718	0.712	**0.921**
NN	0.348	0.348	0.348	0.421	0.521	0.348	0.202	0.441

TABLE VII

PERFORMANCE OF DIFFERENT SUPERVISED CLASSIFIERS FOR BINARY CLASSIFICATION

Classifier	**Micro**				**Macro**			
	Precision	**Recall**	**F1**	**ROC-AUC**	**Precision**	**Recall**	**F1**	**ROC-AUC**
Decision Tree	0.825	0.825	0.825	0.846	0.845	0.846	0.846	0.846
K-NN	0.811	0.811	0.811	0.813	0.808	0.813	0.809	0.813
Logistic Regression	0.825	0.825	0.825	0.819	0.824	0.819	0.820	0.819
Naive Bayes	0.788	0.788	0.788	0.788	0.787	0.788	0.785	0.788
SVM	0.568	0.568	0.568	0.500	0.284	0.500	0.362	0.500
Random Forest	0.890	0.890	0.890	0.884	0.893	0.884	0.886	0.884
Bagging	**0.895**	**0.895**	**0.895**	**0.889**	**0.896**	**0.889**	**0.892**	**0.889**
Boosting	0.568	0.568	0.568	0.500	0.284	0.500	0.362	0.500
MLP	0.840	0.840	0.840	0.827	0.824	0.828	0.823	0.827
NN	0.676	0.676	0.676	0.504	0.637	0.639	0.638	0.504

timeline-centric. These features are used to classify users into collusive or genuine [1].

- **Profile-centric features (PF)**
  – *Account age*: Number of days between the day the data was collected and the day the account was created.
  – *Screen name length:* Length of the screen name.
  – *User description existence:* Whether the user has a description or not.
  – *User description length:* Length of user description.
  – *User URL existence:* Whether the user has a URL present in his/her profile or not.
  – *Follower count:* Number of users the account follows.
  – *Friend count:* Number of users the account is followed by.
  – *Bot score:* This feature is calculated using the Botometer API.[14] The score is based on how likely the account is to be a bot.
- **Timeline-centric features (TF):**
  – *Status count:* Total number of tweets the user has posted.
  – *Retweet count:* Total number of tweets the user has retweeted.
  – *Average mentions per tweet:* It is calculated as the ratio of total number of mentions to the total number of tweets the user has posted/retweeted.
  – *Average URLs per tweet:* It is calculated as the ratio of total number of URLs to the total number of tweets the user has posted/retweeted.
  – *Average hashtags per tweet:* It is calculated as the ratio of total number of hashtags to the total number of tweets the user has posted/retweeted.
  – *Average tweets per day:* It is the average number of tweets the user publishes in a day.
  – *Average retweets per tweet:* It is the average number of retweets the user retweets in a day.
  – *Average symbols per tweet:* It is calculated as the ratio of total number of symbols to the total number of tweets the user has published/retweeted.
  – *Tweeting likelihood ($TL_{1\text{-}7}$)*: It is calculated as the ratio of the per day tweet count of a user to the total number of tweets the user posted in a week. We calculate 7 different features for seven days of a week.
  – *Retweeting likelihood ($RL_{8\text{-}14}$)*: It is calculated as the ratio of the per day retweets of a user to the total number of retweets the user posted in a week. We calculate 7 different features for seven days of a week.
  – *Tweet regularity ($TR_{1\text{-}7}$)*: It is the fraction of tweets posted by the user at i^{th} hour of that day calculated as $-\sum_{i=1}^{24} p(x_i)\log p(x_i)$, where $p(x_i)$ is the fraction of tweets posted by the user at i^{th} hour of that day. We calculate 7 different features for seven days of a week.
  – *Reweet regularity ($RR_{1\text{-}7}$)*: It is the fraction of retweets posted by the user at i^{th} hour of that day calculated as $-\sum_{i=1}^{24} p(x_i)\log p(x_i)$, where $p(x_i)$ is the fraction of retweets posted by the user at i^{th} hour of that day. We calculate 7 different features for seven days of a week.
  – *Tweet steadiness*: Tweet steadiness is defined as $1/\sigma_t$ where σ_t is the standard deviation of time difference between consecutive user-generated tweets.
  – *Retweet steadiness*: It is the reciprocal of the standard deviation of time difference between consecutive user-generated retweets.
  – *Maximum tweet likelihood ($MTL_{1\text{-}7}$)* : It is the ratio of per-day number of tweets posted by a user to the

[14] https://botometer.iuni.iu.edu/

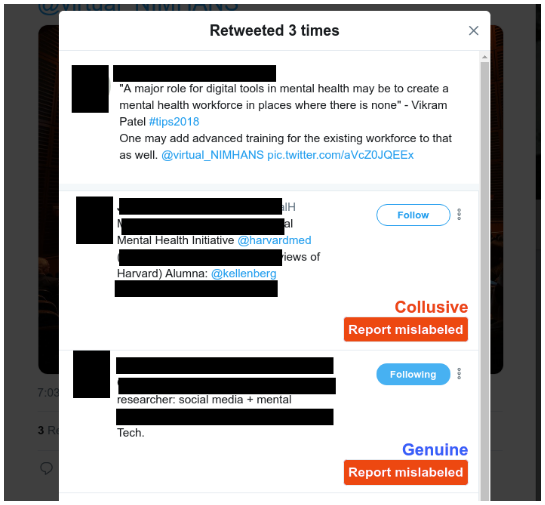
maximum number of tweets the user posted in a day of a week.

- *Maximum retweet likelihood (MRL$_{1\text{-}7}$)*: It is the ratio of per-day number of retweets done by a user to the maximum number of retweets the user does in a day of a week.
- *Retweet count deviation*: It is the standard deviation of number of retweets for all user-generated tweets.
- *Retweet time deviation average*: It is the mean of log-time difference between consecutive retweets.
- *Retweet time deviation standard deviation*: It is the standard deviation of log-time difference between consecutive retweets.

We use these features for our classification experiment. We first conduct a multi-class classification experiment with the following four classes - bots, promotional customers, normal customers and genuine users. We then perform a binary classification experiment with two classes - collusive (combining all types of customers) and genuine users.

We use several standard supervised classifiers – Support Vector Machine (SVM), Decision Tree (DT), Naive Bayes (NB), K-Nearest Neighbors (K-NN) and Logistic Regression (LR), and three ensemble classifiers – Random Forest (RF), Bagging (BG), Boosting (BO), Multi-layer Perceptron (MLP) and a customized neural network (NN). Hyper-parameter optimization is also performed to generate the best results using multiple parameters e.g., we use CART with Gini gain criteria for Decision Tree, multinomial logistic regression and SVM with linear kernel. We initialize MLP with 1 hidden layer of 100 dimensions using ReLU activation function with Adam adaptive update rule. For NN, in case of binary classification, we pass the extracted features through dense layers of dimensionality 4. We use ReLU as a activation function independently in each of the 4 nodes of first 2 dense layers. Finally, we set the final activation function to be sigmoid and minimize the binary cross entropy loss for 500 epochs using Adam adaptive update rule. In case of multi-class classification, we use softmax activation function in the output layer. The performance evaluation of each of the methods is measured using the following evaluation metrics: Precision, Recall, F1-score, and Area under the ROC curve. The evaluation metrics are reported after averaging the result of 10-fold cross-validation.

Table VI shows the result of the multi-class classification. With our feature set, Random Forest classifier turns out to be the best model with a macro F1-score 0.791. Using the MLP classifier, we achieve a very high ROC-AUC score of 0.923 in micro (0.921 in macro) setting. Further investigation reveals that the prediction results too many negative samples (most of the users are classified as genuine users) which is responsible for the high value of ROC-AUC score. Table VII shows the result of the binary classification. With our feature set, Bagging outperforms the other classifiers in detecting collusive users in terms of all the evaluation metrics (macro F1-score 0.892). We show in Supplementary the importance of each features for the multi-class and binary classification experiments.


Retweeted 3 times ×

[redacted] [redacted]
"A major role for digital tools in mental health may be to create a mental health workforce in places where there is none" - Vikram Patel #tips2018
One may add advanced training for the existing workforce to that as well. @virtual_NIMHANS pic.twitter.com/aVcZ0JQEEx

[redacted] [redacted]
Mental Health Initiative @harvardmed
[redacted]
Harvard) Alumna: @kellenberg
[redacted]

Collusive

[redacted] [redacted]
researcher: social media + mental
[redacted]
Tech.

Genuine

Fig. 9. Working of our chrome extension SCoRe++.

Next, we discuss the working principle of the chrome extension to detect collusive users in real time. The chrome extension will notify Twitter users whether a retweeter of a tweet is 'collusive' or 'genuine'.

VII. SCoRe++: Detecting Collusive Users in Real Time

In order to detect collusive users in real time, we develop a chrome extension SCoRe++ for end-users. The extension classifies a Twitter user into collusive or genuine and seamlessly integrates the result into user's Twitter pages. On installing the chrome extension, it automatically loads when an end-user opens a tweet page. SCoRe++ extracts the IDs of each retweeter of the tweet and makes a request to our backend server. The backend code then extracts the features using the Tweepy API and feeds the features into the best trained model (Bagging in our case) to return appropriate label for each retweeter. The returned label of each retweeter is showed in the tweet page. We also add a 'Report mislabeled' label for feedback from the end-user. Upon clicking on this button by the end-user, we store the retweeter ID and the label returned by SCoRe++ in our server. We retrain our classifier with the feedback from the end-user and check the confidence score of the new model. The current version of SCoRe++ is written in Javascript[15] and the backend code is written in Python (using Flask Framework[16]). The time taken to label each retweeter is dependent on various factors (internet bandwidth, Twitter API query time etc.). Fig. 9 shows the working of SCoRe++ on Chrome Browser.

VIII. Conclusion

In this study, we thoroughly investigated Twitter blackmarket services which provide retweets for a tweet. We conducted our study on two different types of blackmarket services: premium and freemium. We first followed an active-probing strategy to create a dataset of collusive users collected from both the blackmarket services. In case of premium services,

[15] https://en.wikipedia.org/wiki/JavaScript
[16] http://flask.pocoo.org/

we purchased a fixed set of retweeters by paying a certain amount to the service. To collect the users involved in freemium services, we chose only credit-based freemium services where we retweeted tweets of other users of the service to gain credits. We then provided a detailed analysis of these collusive users based on retweet-centric, profile-centric, timeline-centric and network-centric behavior. Following this, we ran several supervised classifiers to classify collusive users and genuine users based on a set of 63 features. We further developed a chrome extension SCoRe++ to detect collusive users in real time. We would like to reemphasize that this paper is the first attempt to present a rigorous analysis of various kinds of blackmarket services which, we believe, would help researchers understand the micro-level dynamics of these syndicates and lead to develop computational techniques to expose their activities.

In our continuing research, we are interested to explore the following avenues. First, we would like to detect suspicious tweets which are submitted to blackmarket services. Automatic detection of such tweets will surely trigger and fuel further research efforts for the detection of collusive users. Second, we will try to capture the interdependency of collusive users and the tweets which are submitted to the blackmarket services. Third, we also intend to expand our human-annotated dataset of collusive users (both premium and freemium users). Ultimately, our goal is to design a scalable real-world collusive user detection framework for Twitter.

Acknowledgment

T. Chakraborty would like to thank the support of Google India Faculty Award and the Infosys Centre for Artificial Intelligence, Indraprastha Institute of Information Technology, India.

References

[1] H. S. Dutta, A. Chetan, B. Joshi, and T. Chakraborty, "Retweet us, we will retweet you: Spotting collusive retweeters involved in blackmarket services," in *Proc. IEEE ASONAM*, Aug. 2018, pp. 242–249.

[2] M. Giatsoglou, D. Chatzakou, N. Shah, C. Faloutsos, and A. Vakali, "Retweeting activity on Twitter: Signs of deception," in *Proc. PAKDD*. Ho Chi Minh City, Vietnam: Springer, 2015, pp. 122–134.

[3] A. Chetan, B. Joshi, H. S. Dutta, and T. Chakraborty, "CoReRank: Ranking to detect users involved in blackmarket-based collusive retweeting activities," in *Proc. WSDM*, 2019, pp. 330–338.

[4] U. Arora, W. S. Paka, and T. Chakraborty, "Multitask learning for blackmarket tweet detection," 2019, *arXiv:1907.04072*. [Online]. Available: http://arxiv.org/abs/1907.04072

[5] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia, "Who is tweeting on Twitter: Human, bot, or cyborg?" in *Proc. 26th Annu. Comput. Secur. Appl. Conf.*, 2010, pp. 21–30.

[6] C. A. Davis, O. Varol, E. Ferrara, A. Flammini, and F. Menczer, "BotOrNot: A system to evaluate social bots," in *Proc. WWW*, 2016, pp. 273–274.

[7] S. Gianvecchio, M. Xie, Z. Wu, and H. Wang, "Measurement and classification of humans and bots in Internet chat," in *Proc. USENIX Secur. Symp.*, 2008, pp. 155–170.

[8] B. Wang, A. Zubiaga, M. Liakata, and R. Procter, "Making the most of tweet-inherent features for social spam detection on Twitter," 2015, *arXiv:1503.07405*. [Online]. Available: https://arxiv.org/abs/1503.07405

[9] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, "Detecting spammers on Twitter," in *Proc. CEAS*, vol. 6, 2010, p. 12.

[10] M. Giatsoglou, D. Chatzakou, N. Shah, A. Beutel, C. Faloutsos, and A. Vakali, "ND-Sync: Detecting synchronized fraud activities," in *Proc. PAKDD*. Ho Chi Minh City, Vietnam: Springer, 2015, pp. 201–214.

[11] A. Beutel, W. Xu, V. Guruswami, C. Palow, and C. Faloutsos, "CopyCatch: Stopping group attacks by spotting lockstep behavior in social networks," in *Proc. WWW*, 2013, pp. 119–130.

[12] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Predicting positive and negative links in online social networks," in *Proc. WWW*, 2010, pp. 641–650.

[13] N. Jindal and B. Liu, "Review spam detection," in *Proc. WWW*, 2007, pp. 1189–1190.

[14] S. Ahmad, A. Pathak, and S. Jaiswal, "A survey about spam detection and analysis using users' reviews," *Malaya J. Mat.*, no. 1, pp. 1–4, 2018. [Online]. Available: https://bit.ly/339lJKC

[15] S. Dhawan, S. C. R. Gangireddy, S. Kumar, and T. Chakraborty, "Spotting collusive behaviour of online fraud groups in customer reviews," in *Proc. IJCAI*, 2019, pp. 245–251.

[16] S. Gupta, A. Khattar, A. Gogia, P. Kumaraguru, and T. Chakraborty, "Collective classification of spam campaigners on Twitter: A hierarchical meta-path based approach," in *Proc. WWW*, 2018, pp. 529–538.

[17] G. Stringhini *et al.*, "Follow the Green: Growth and dynamics in Twitter follower markets," in *Proc. IMC*, 2013, pp. 163–176.

[18] N. Shah, H. Lamba, A. Beutel, and C. Faloutsos, "The many faces of link fraud," in *Proc. ICDM*, Nov. 2017, pp. 1069–1074.

[19] C. De Micheli and A. Stroppa, "Twitter and the underground market," in *Proc. 11th Nexa Lunch Seminar*, vol. 22, 2013, pp. 1–43.

[20] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "Fame for sale: Efficient detection of fake Twitter followers," *Decis. Support Syst.*, vol. 80, pp. 56–71, Dec. 2015.

[21] K. Thomas, D. McCoy, C. Grier, A. Kolcz, and V. Paxson, "Trafficking fraudulent accounts: The role of the underground market in Twitter spam and abuse," in *Proc. USENIX Secur. Symp.*, 2013, pp. 195–210.

[22] M. Singh, D. Bansal, and S. Sofat, "Followers or fradulents? An analysis and classification of Twitter followers market merchants," *Cybern. Syst.*, vol. 47, no. 8, pp. 674–689, 2016.

[23] Y. Liu, Y. Liu, M. Zhang, and S. Ma, "Pay me and i'll follow you: Detection of crowdturfing following activities in microblog environment," in *Proc. IJCAI*, 2016, pp. 3789–3796.

[24] S. Yang, H. Jin, B. Li, and X. Liao, "A modeling framework of content pollution in peer-to-peer video streaming systems," *Comput. Netw.*, vol. 53, no. 15, pp. 2703–2715, 2009.

[25] N. Shah, "FLOCK: Combating astroturfing on livestreaming platforms," in *Proc. WWW*, 2017, pp. 1083–1091.

[26] M. N. Reddy, T. Mamatha, and A. Balaram, "Analysis of e-recruitment systems and detecting e-recruitment fraud," in *Proc. Int. Conf. Commun. Cyber Phys. Eng.* Springer, 2018, pp. 411–417.

[27] N. Shah, H. Lamba, A. Beutel, and C. Faloutsos, "OEC: Open-ended classification for future-proof link-fraud detection," 2017, *arXiv: 1704.01420*. [Online]. Available: http://arxiv.org/abs/1704.01420

[28] K. Thomas, C. Grier, D. Song, and V. Paxson, "Suspended accounts in retrospect: An analysis of Twitter spam," in *Proc. IMC*, 2011, pp. 243–258.

Hridoy Sankar Dutta received the B.Tech. degree in computer science and engineering from the Institute of Science and Technology, Gauhati University, India, in 2013 and the M.Tech. degree in computer science and engineering from NIT Durgapur, India, in 2015. He is currently pursuing the Ph.D. degree in computer science and engineering with the Indraprastha Institute of Information Technology, Delhi (IIIT-D), India. His research interests include data-driven cyber security, social network analysis, natural language processing, and applied machine learning.

Tanmoy Chakraborty is currently an Assistant Professor and a Ramanujan Fellow with the Department of Computer Science and Engineering, IIIT Delhi, India. His primary research interests include social network analysis, data mining, and natural language processing. He has been serving as a Program Committee Member in several top conferences. He has received several awards, including the Google Indian Faculty Award, the Early Career Research Award, and the DAAD Faculty Award.