



## Supplementary materials for

Ping HE, Xuhong ZHANG, Changting LIN, Ting WANG, Shouling JI, 2024. Towards understanding bogus traffic service in online social networks. *Front Inform Technol Electron Eng*, 25(3):415-431.  
<https://doi.org/10.1631/FITEE.2300068>

### 1 Communication channel analysis

**Popularity.** QQ and WeChat are the most influential IM Apps in China. As of March 31, 2020, QQ and WeChat had 0.6935 and 1.023 billion monthly active users, respectively. We believe that one of the reasons that agents adopt these IM Apps is that they want to make it easy and convenient for customers to reach them.

**Security.** These two platforms ensure good privacy support for conversations, which naturally provides an umbrella for the agents to escape from detection.

**Convenience.** The BTS also needs a payment platform. QQ and WeChat provide a complete payment platform to satisfy this need. Additionally, the payment services in these Apps are widely used by the users; e.g., WeChat Pay has more than 0.8 billion active monthly users. Thus, these Apps provide a convenient payment method for both customers and agents to complete transactions.

### 2 Honeypot account

Fig. S1 shows the announcement in our honeypot account. In the announcement, we tell users that our account is being used for studying BTS and do not trust the fans of our account. Besides, we leave our contact method.

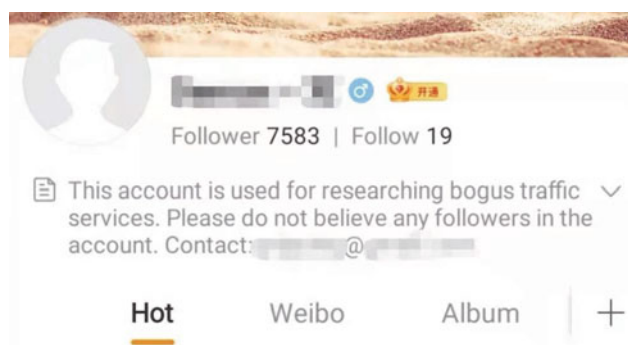


Fig. S1 The announcement in our honeypot account

### 3 Evasive tweets

Fig. S2 shows two examples of evasive tweets. They have legal sentence structure and meanings. However, they are quotations expressed in ancient Chinese.



Fig. S2 Two examples of evasive tweets in our dataset. The words in red are the translations for the Chinese characters of evasive tweets

#### 4 Profile-based features

**Authentication.** Authentication is a binary feature that represents whether a user passes the official verification. For example, Weibo authentication requires users to bind their mobile phones and have an avatar. In general, the authenticated users are unlikely to be BTAs. The detailed authentication rules of Weibo are described in Section 6.

**Number of of followings.** Fig. S3(a) shows the distribution of the number of of followings for the BTAs and benign users in our collected seed dataset. We can see that compared with the benign users, the BTAs tend to follow more users.

**Number of of followers.** Fig. S3(b) shows the distribution of the number of of followers for the BTAs and benign users. We can see that the number of followers of the benign users exhibits a power-law distribution while that of the BTAs concentrates at the interval of 10 to 1000 followers.

**Number of of tweets.** Fig. S3(c) shows the distribution of the number of tweets for the BTAs and benign users. We can see that the number of tweets of benign users also exhibits a power-law distribution while that of the BTAs concentrates at the interval of 10 to 1000 tweets.

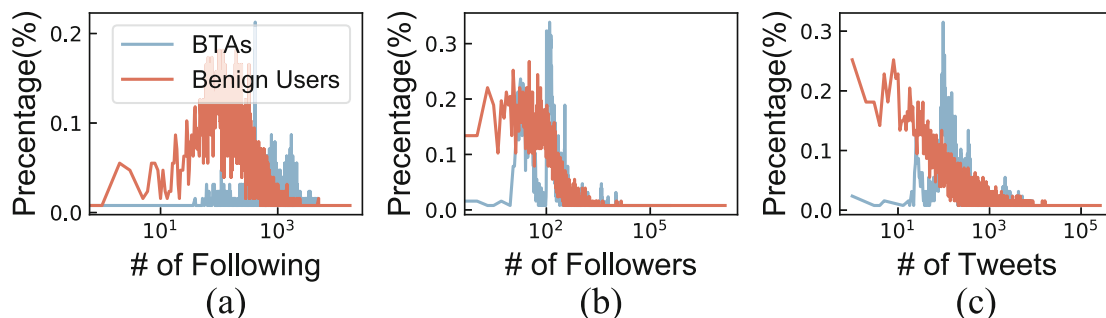


Fig. S3 Feature analysis of the profile-based features: (a) the distribution of the number of following; (b) the distribution of the number of followers; (c) the distribution of the number of tweets

**Account level.** Microblogging sites usually record the age of user accounts, and older accounts will be ranked at a high level. In general, accounts with high levels are unlikely to be BTAs.

We also consider other profile features (e.g., if a user has a self-description) and behavior-based features (e.g., the number of average likes per tweet). We do not take the graph-based features into consideration, as BTAs do not rely on social connections to conduct attack. In total, we analyze 12 profile-based and behavioral-based features, which are authentication, account level, if a user has self-description, number of followings, number of followers, the length of nick name, gender, number of tweets, the length of the self-description, number of average likes per tweet, number of average comments per tweet, and number of

average retweets per tweet. We manually check the distribution of each feature. At last, we find five features, i.e., authentication, account level, number of followings, number of followers, and number of tweets, which have significant different distributions between BTAs and benign users.

#### Our verification rules.

1. Check if the profile portrait contains promotional information about BTS. If it does, then label the user as BTA, and go to step 2 otherwise.
2. Check if 80% of the tweets are all copied from existing corpora, such as famous quotations. If they are then label them as BTAs.

To evaluate the accuracy of our verification rules, we apply our verification rules to 500 randomly sampled sybils and 500 randomly sampled benign users. The results show that our verification rules achieve an accuracy of 99.3%.

## 5 Case study

We investigate the topics related to superstars as a case to show the patterns of the attack cycle. We select three topics which are the names of these three superstars. For ethics and privacy reasons, we do not disclose the names of these three superstars. For each topic, we count the number of corresponding tweets in each month. The collection period spans from 2017 to 2020. As shown in Fig. S4, these superstars periodically buy the BTS and each time the attack lasts for a short period. Interestingly, the most intensive attack for the three superstars all happened around May 2018. More specifically, for superstar A, the attack lasts from May 10th to May 19th, 2018. By querying the Hot Search Topic history of Weibo, we find that from May 11, 2018 to May 14, 2018 and from May 16, 2018 to May 19, 2018, this superstar’s topic appeared in the Hot Search Topic list, lasting about 46 hours in total. The reason for the attack could be that a talent show was held during that time and that these three superstars all participated in it. They may want to attract more public attention and thus purchase the BTS.

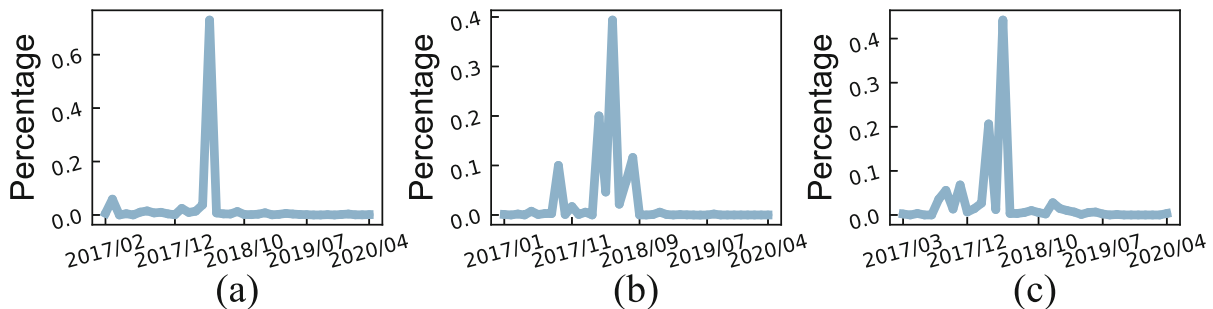


Fig. S4 Bogus traffic distributions for three superstars: (a) superstar A; (b) superstar B; (c) superstar C

## 6 Weibo authentication rules

1. Binding a mobile phone number, having a clear profile photo, the number of followers is not less than 50, the number of followings is not less than 50, and having mutual fans with at least 2 authenticated users.
2. The content of the posted tweets could reflect an active real individual.