Ping HE, Xuhong ZHANG, Changting LIN, Ting WANG, Shouling JI, 2024. Towards understanding bogus traffic service in online social networks. *Frontiers of Information Technology & Electronic Engineering*, 25(3):415-431. https://doi.org/10.1631/FITEE.2300068

Towards understanding bogus traffic service in online social networks

Key words: Online social networks; Measurement; Bogus traffic; Black market

Corresponding author: Shouling JI
E-mail: sji@zju.edu.cn
ORCID: https://orcid.org/0000-0003-4268-372X

Motivation

- Critical functionality and huge influence of the hot trend/topic page (HTP) in microblogging sites have driven the creation of a new kind of underground service called the bogus traffic service (BTS). BTS provides a kind of illegal service which hijacks the HTP by pushing the controlled topics into it for malicious customers with the goal of guiding public opinions.

Promotional profile portraits about BTS

Threat model

The threat model consists mainly of two parts: the physical world and the online social network (OSN) world. This study focuses on attacks in the OSN world. We define the attack process as the following: BTAs (attacker/attack entity) massively tweet and retweet the target topics (hashtags) until the site's HTP (victim) includes the target topics. Meanwhile, the BTAs adopt complex tactics to impersonate benign accounts to evade detection.

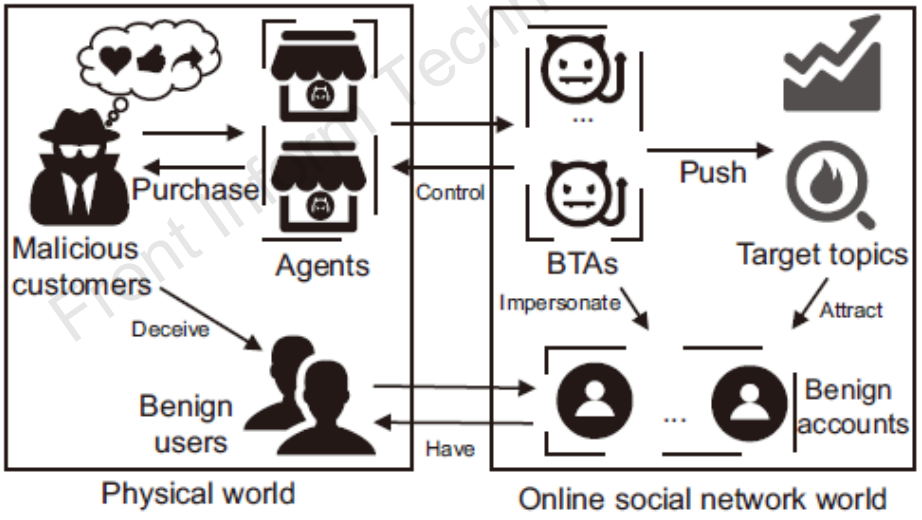

Fig. 1 Threat model of our research consisting of malicious customers, agents, benign users, BTAs, target topics, and benign accounts

Methodology

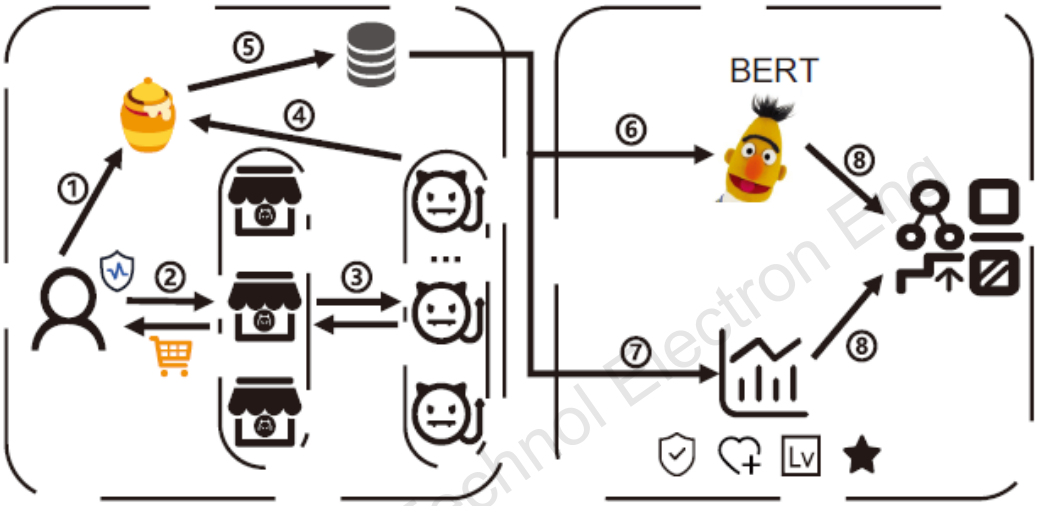

Fig. 3 Overview of our detection methodology

①: setting up a honeypot account to capture BTAs. ②: purchasing BTAs from agents. ③: agents manipulate BTAs to conduct an attack. ④: BTAs' attack is captured by the honeypot account. ⑤: building the ground-truth dataset. ⑥: using BERT to extract semantic features. ⑦: analyzing profile-based features. ⑧: XGBoost classifier

Infiltration results

Table 1 Purchase solution for BTAs

Agent	Source	Number of accounts ordered	Number of accounts obtained
GYCM87	Head portraits	50	27
KunMaiFen	Head portraits	1500	1004
June Flower Outdoor Store	E-commerce	700	468
Yetian Outdoor Store	E-commerce	700	489
Yesheng Outdoor Store	E-commerce	500	473
fsj3.com	Website	2000	1378
www.niufenba.com	Website	2000	1203
Total		7450	5042

Table 2 Datasets gathered in our infiltration process

Dataset	Type	Number of accounts	Time
Ground-truth dataset	BTAs	5042	Aug. 2019–Dec. 2019
	Benign users	6652	Aug. 2019–Dec. 2019
Candidate dataset	Fans of BTAs	523 323	Jan. 2020–May 2020

As shown in Table 1, we can see that the number of actually delivered BTAs is much smaller than the number of accounts ordered. As a result, we successfully obtain 5042 BTAs out of the 7450 ordered BTAs.

Detection results

Table 3 Performance of the BTA detection task on different methods

NLP model	Classifier	Precision	Recall	F1
BERT	XGBoost	0.972	0.959	**0.965**
	SVM	0.950	0.957	0.954
	RF	0.971	0.945	0.960
	MLP	0.958	0.957	0.957
word2vec	XGBoost	0.948	0.973	0.960
	SVM	0.868	0.911	0.889
	RF	0.961	0.937	0.949
	MLP	0.877	0.911	0.894

Best result is in bold

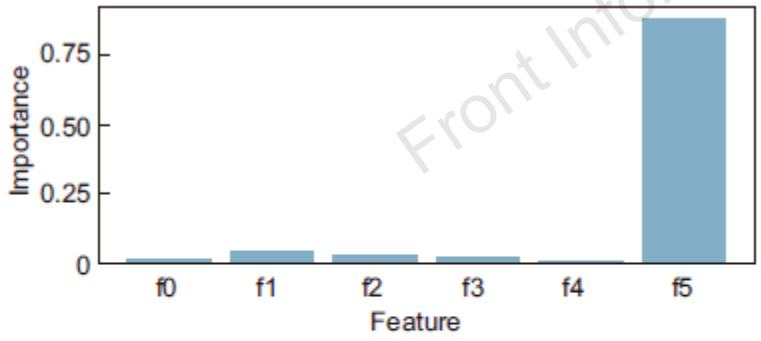

Fig. 7 Feature importance of our BTA detector

f0 is the authentication, f1 is the account level, f2 is the number of followings, f3 is the number of followers, f4 is the number of tweets, and f5 is the percentage of evasive tweets

BERT + XGBoost achieves the best performance.

Linguistic feature is the most influence feature.

Detection results

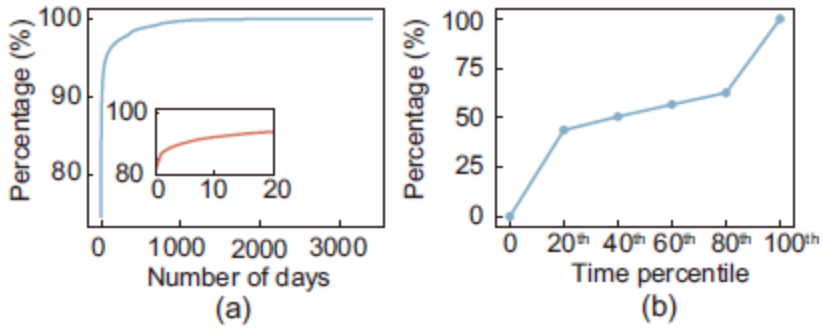

Fig. 8 Attack cycle measurement: (a) CDF of attack cycle duration; (b) CDF of bogus traffic volume within an attack cycle

Most of the bogus traffic concentrates around the beginning and end of an attack cycle.

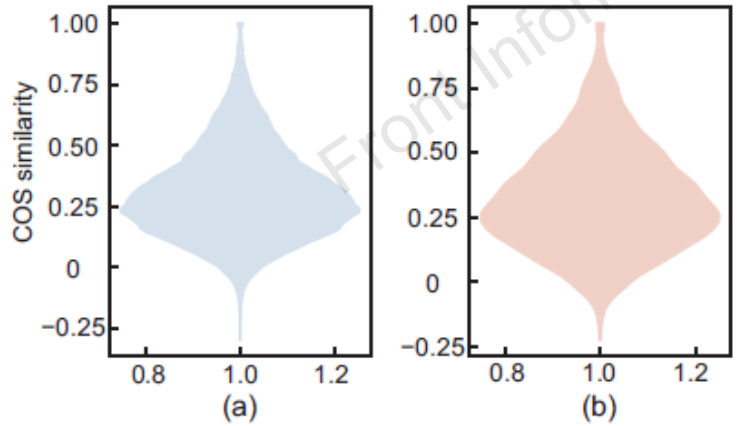

Fig. 12 Semantic similarity between the additional comments and the original tweets in the retweets: (a) BTAs; (b) benign users

The similarity distribution of BTAs is almost identical to that of benign users.

Conclusions

- This research conducts the first large-scale analysis on the bogus traffic service in online social networks.

- We design an NLP-based method to detect the bogus traffic accounts by capturing their linguistic differences.

- By deploying the method, we uncover 296 916 topics potentially linked to the bogus traffic. Moreover, we elucidate the operational mechanisms of bogus traffic services, examining both the attack cycle and the entities involved.