ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 3, 2024

Social Spammer Detection via Convex Nonnegative Matrix Factorization

[1] K VENKATESH, [2] M. SRAVANI

[1](Assistant Professor), MCA, **DANTULURI NARAYANA RAJU COLLEGE(A) PG COURSES, BHIMAVARAM ANDHRA PRADESH**

[2]MCA, scholar, **DANTULURI NARAYANA RAJU COLLEGE(A) PG COURSES, BHIMAVARAM ANDHRA PRADESH**

ABSTRACT

With the increasing popularity of social network platforms such as Twitter and Sina Weibo, a lot of malicious users, also known as social spammers, disseminate illegal information to normal users. Several approaches are proposed to detect spammers by training a classifier with optimization methods and mainly using content and social following information. Due to the development of spammers' strategies and the courtesy of some legitimate users, social following information becomes vulnerable to fake by spammers. Meanwhile, the possible social activities and behaviours vary significantly among different users, which leads to a large yet sparse feature space to be modelled by existing approaches. To address issues, in this paper, we propose a new approach named CNMFSD for spammer detection in social networks, which exploits both content information and users' interaction relationships in an innovative manner. We have empirically validated the proposed method on a real-world Twitter dataset, and experimental results show that the proposed CNMFSD method improves the detection performance significantly compared with baselines.

1.INTRODUCTION

SOCIAL networks, such as Twitter, Face book, and Sina Weibo, are increasingly used to disseminate and share information easily and quickly. However, it is a double-edged sword since the success of social networks also attracts more social spammers [1]. They try to seize our privacy, send us unwanted

information, publish malicious content and links [2], and promote commodity information, which thoroughly impacts social stability and organizational management models [3]. According to a study by Negate [4], the number of social spammers grows so fast that one in two hundred social messages is spam. Meanwhile, to increase their influence and be undetected, spammers collude with each other to construct the criminal communities [5]. Thus, social spammer detection is a challenging task for researchers. Successful social spammer detection presents its significance to improve the quality of user experience, and positively impact the overall value of the social systems going forward [6].

In the past decade, researchers have tried different techniques to detect spammers, such as link analysis [7] and content analysis [8], [9]. The methods of content-based detection of spammers mainly focus on analysing and extracting users' features and then directly applying existing classification approaches such as support vector machines (SVM) to detect spammers [9]–[11]. Recently, more advanced deep learning-based approaches have been proposed to detect social spammers only based on content [12]–[14]. However, with the development of spamming strategies, these methods could not accurately detect spammers with new strategies, only relying on the extracted features. Another category of methods is proposed to detect spammers via social network analysis [15]. These methods assume that spammers cannot establish an arbitrarily large number of social trust relations with legitimate users. The users, who have relatively low social influence or social status in social networks, will be determined as spammers. Unfortunately, only depending on network information, these methods are hard to distinguish between legitimate users and spammers.

Some approaches [16][18] have been proposed to detect spammers via both content and network analysis, which identify spammers more accurately than the traditional approaches. The main challenge in detecting social spammers is that the possible social activities and behaviours are more varied and complex, and they constitute a much larger feature space. As a result, spammers are more challenging to detect. Therefore, it is crucial to design more

effective methods for extracting users' features. Meanwhile, the reflexive reciprocity [19] indicates that many users simply follow back when they are followed by someone for the sake of courtesy. It is easier for spammers to acquire a large number of follower links in social networks. Thus, with the perceived social influence, they can avoid being detected. However, the interactions between spammers and legitimate users are usually unilateral. In most cases, spammers share a message and then mention (i.e., @) legitimate users. On the contrary, legitimate users constantly interact with legitimate users but have few interactions with spammers. Consequently, it is more reasonable to take the interactions among users into consideration when detecting spammers.

2. EXISTING SYSTEM

Spammers since Heymann et al. [22] firstly surveyed potential solutions and challenges in social spammer detection. Masood et al. [6] elaborated a classification of spammer detection techniques, including fake content, URL-based spam detection, detecting spam in trending topics, and fake user identification. In this paper, we only focus on the binary classification task, i.e., spammer or legitimate user identification. Many approaches employed machine learning methods to train a classifier to detect spammers. SMFSR [16] jointly modelled user activities' information and the social following information to learn a classifier. SSDM [17] incorporated users' text information and social following information into an efficient spare supervised model for spammer detection. Mateen et al. [23] proposed a hybrid technique that utilizes user-based, content-based, and graph-based characteristics for spammer profiles detection. Gupta et al. [24] presented a policy for the detection of spammers on Twitter and used the popular techniques, i.e., Naive Bayes, clustering. decision tree. An important line of research in spam detection relies on analysing the tweet content, as shown in [25] and [26] where suspicious use of hashtags or URLs is traced. The main objective in [26] is to study the semantics of short texts or messages in contrast with a set of Wikipedia text pages that are modelled and used as an aggregation of entities. The work presented in [25] stresses the need for efficient URL detection schemes utilizing different features such as lexical ones and dynamic behaviours.

Disadvantages

The system is not implemented Convex-NMF based Supervised Spammer Detection with Social Interaction (CNMFSD). ϖ The system is not implemented any ml classifier for test and train the datasets.

3. PROPOSED SYSTEM

• The system proposes a three-stage optimization model that conducts feature extraction and classifier learning simultaneously. First, we use Convex Non-negative Matrix Factorization (CNMF) and Non-negative Matrix Factorization (NMF) to induce latent feature from content information, then train an SVM classifier and finally, refine latent features using social interaction information as the input representations of the classifier. Through iteratively learning among content information, social interaction regularization, and classification model, the proposed method can train an accurate classifier.

• The system proposes a novel method to induce latent features and a novel social interaction regularization term. Using CNMF, we get the latent content matrix of spammers and legitimate users, respectively, and then obtain the user feature latent matrix by NMF according to the latent content matrix. The latent feature refine process is guided by the social interaction relationship matrix and the label information.

• The proposed system evaluates our method on a large-scale real-world socii al network data set from Twitter, one of the largest social networks in the world. The experimental results show that the proposed framework can identify more spammers compared with baseline approaches. We conduct experiments to demonstrate the significance of using CNMF to induce latent features for spammers and legitimate users, respectively, and validate the effectiveness of the new social interaction regularization term.

Advantages ¬

The proposed system refines the latent features with predicted label information and social interaction information with the help of some classifier. ¬ The proposed system implemented UNLABELED USER CONTENT FACTORIZATION.

4.OUTPUTSCREENS

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN2454-9940

www.ijasem.org

Vol 18, Issue 3, 2024

Fig.1. Admin login pa

Fig.2. User details

Fig.3. Output graph

Fig.4. Output graph

Fig.5. Output graph

Fig.6. Spammer detected

Fig.7. Spammer ratio details

5.CONCLUSION

In this paper, we propose a new framework by taking advantage of content and social interaction information for social spammer detection. Different from existing methods that utilize users' the following

information, the proposed method CNMFSD integrates users' interaction information based on the trained classification model. In addition, we introduce a new strategy to induce latent features using CNMF in spammers and legitimate users' space for improving the performance of detecting spammers. Experimental results on a real dataset show that CNMFSD obtains better detection performance compared with existing methods.

In this work, we employ Convex-NMF to learn latent user features for legitimate users and spammers, respectively. Such a fine-grained learning strategy makes the proposed model obtain accurate latent user representations, which further helps the model to achieve better performance. Besides, introducing social interaction into this task can also improve prediction performance.

Although the proposed model outperforms baselines, it also has some disadvantages. First, in the classifier training stage, we do not consider the social interaction graph, which is trained solely based on the outputs from CNMF. Second, we use tied to extract the user content matrices. However, spammer always posts some normal tweets to imitate the behaviour of legitimate users. Thus, it is essential to distinguish the importance of tweets when we extract the user content matrix.

6. REFERENCES

1] Aliaksandr Babushka and Petr Hajek. Spam detection on social networks

using cost-sensitive feature selection and ensemble-based regularized deep

neural networks. Neural Computing and Applications, 32(9):4239–4257,

2020.

[2] Qiang Fu, Bo Feng, Dong Guo, and Qiang Li. Combating the evolving

spammers in online social networks. Computers & Security, 72:60–73,

2018.

[3] Zhijun Zhang, Rui Hou, and Jin Yang. Detection of social network spam

based on improved extreme learning machine. IEEE Access, 8:112003–

112014, 2020.

[4] Nexgate2013. 2013 state of social media spam. http://nexgate.com/wpcontent/

uploads/2013/09/Nexgate-2013-State-of-Social-Media-Spam-

Research-Report.pdf.

[5] Dehair Liu, Benjin Mei, Jinchuan Chen, Zhiwen Lu, and Xiaoyong Du.

Community based spammer detection in social networks. In International

Conference on Web-Age Information Management, pages 554–558.

Springer, 2015.

[6] Faiza Masood, Ahmad Almgren, Assad Abbas, Hasan Ali Khattak,

Ikram Udi Din, Mohsen Guiana, and Mansour Zubair. Spammer detection

and fake user identification on social networks. IEEE Access, 7:68140–

68152, 2019.

[7] Sanjeev Rao, Anil Kumar Verma, and Tarun Preet Bhatia. A review on

social spam detection: Challenges, open issues, and future directions.

Expert Systems with Applications, 186:115742, 2021.

[8] Chao Chen, Jun Zhang, Yi Xie, Yang Xiang, Wan lei Zhou, Mohammad

Mehedi Hassan, Abdulhameed Allawi, and Majed Arabian. A

performance evaluation of machine learning-based streaming spam tweets

detection. IEEE Transactions on Computational social systems, 2(3):65–

76, 2015.

[9] Xiangshan Zheng, Zhipeng Zeng, Sheyi Chen, Yunlong Yu, and Chumming

Rong. Detecting spammers on social networks. Neurocomputing,

159:27–34, 2015.

[10] Chao Yang, Robert Harkreader, and Goofier Gu. Empirical evaluation and

new design for fighting evolving twitter spammers. IEEE Transactions on

Information Forensics and Security, 8(8):1280–1293, 2013.