SOCIAL SPAMMER DETECTION VIA CONVEX NONNEGATIVE MATRIX FACTORIZATION

[1]Mrs R.Jhansi Rani
[1]Assistant Professor
[1]Department of Computer Applications
[1]Chadalawada Ramanamma Engineering College
(Autonomous),Tirupathi

[2] PUJARI HEMALATHA
[2] Student
[2] Department of Computer Applications
[2] Chadalawada Ramanamma Engineering College
(Autonomous), Tirupathi

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 behaviors vary significantly among different users, which leads to a large yet sparse feature space to be modeled 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.

Keywords : Social networking (online),Feature extraction, Blogs, Support vector machines, Deep learning, Social factors

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 They try to seize our privacy, send us unwanted information, publish malicious content and links and promote commodity information, which thoroughly impacts social stability and organizational management models According to a study by Nexgate 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 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.

In the past decade, researchers have tried different techniques to detect spammers, such as link analysis and content analysis. The methods of content-based detection of spammers mainly focus on analyzing and extracting users' features and then directly applying existing classification approaches such as support vector machines (SVM) to detect spammers Recently, more advanced deep learning-based approaches have been proposed to detect social spammers only based on contentHowever, 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 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 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 behaviors 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 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.

To address these challenges, we propose to take advantage of both social network interaction information and content information.

In this paper, we study the problem of social spammer detection with social interaction and content information. In essence, we investigate: how to model the social interaction information and content information properly; and how to seamlessly utilize both social interaction and content information for the problem we are studying. Our solutions to these two challenges result in a novel spammer detection framework name Convex-NMF based Supervised Spammer Detection with Social Interaction (CNMFSD). Based on statistical analysis, we observe that spammers and legitimate users have different characteristic distributions. Thus, we use a matrix factorization model to collaboratively induce a succinct set of latent features for spammers and legitimate users, respectively, and this latent feature learning process is guided by the label information. The generated features are then used as

input representation for a spammer classification model. Then we refine the latent features with predicted label information and social interaction information. Finally, the refined latent features are used as the input representation for the final classification. The main contributions of this paper are outlined as follows:

- We propose 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.

- We propose 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.
- We evaluate our method on a large-scale real-world social 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.

The structure of this paper is organized as follows. In Section II, we review existing work in social spam detection. In Section III, we formally define the problem of social spammer detection with content and social interaction information. In Section IV, we propose a new model to integrate both content and social interaction information for spammer detection. In Section V, we report empirical results on a real world dataset. Finally, we conclude and present the future work in Section VI.

2. LITERATURE SURVEY

2.1 DIFFERENT AUTHORS DISCUSSION

we study the problem of social spammer detection with social interaction and content information. In essence, we investigate: how to model the social interaction information and content information properly; and how to seamlessly utilize both social interaction and content information for the problem we are studying. Our solutions to these two challenges result in a novel spammer detection framework name Convex-NMF based Supervised Spammer Detection with Social Interaction (CNMFSD). Based on statistical analysis, we observe that spammers and legitimate users have different characteristic distributions. Thus, we use a matrix factorization model to collaboratively induce a succinct set of latent features for spammers and legitimate users, respectively, and this latent feature learning process is guided by the label information. The generated features are then used as the input representation for a spammer classification model.

2.2 DOMAIN DESCRIPTION

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

3. PROBLEM STATEMENT

3.1 EXISTING SYSTEM

Spammers since Heymann firstly surveyed potential solutions and challenges in social spammer detection. Masood 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 jointly modeled user activities' information and the social following information to learn a classifier. SSDM incorporated users' text information and social following information into an efficient spare supervised model for spammer detection. Mateen proposed a hybrid technique that utilizes user-based, content-based, and graph-based characteristics for spammer profiles detection. Gupta presented a policy for the detection of spammers on Twitter and used the popular techniques, i.e., Naive Bayes, clustering, and decision tree.

An important line of research in spam detection relies on analyzing the tweet content, as shown in where suspicious use of hashtags or URLs is traced. The main objective in is to study the semantics of short texts or messages in contrast with a set of Wikipedia text pages that are modeled and used as an aggregation of entities. The work presented in stresses the need for efficient URL detection schemes utilizing different features such as lexical ones and dynamic behaviors.

Other directions adopted in detecting Twitter spammers focus on discovering traits or patterns that best describe the spammer's behavioral profile. In such works like the main contribution is to determine deceptive double characters for user profiles, which is done by analyzing nonverbal behavior variables as a function of time, such as follows and retweets. Also, Sumner follow a similar technique. Direct approaches to checking up the user's portfolio include, but are not limited to, the notion of having no profile photo/biography/personal tweets or a suspiciously high/low number of followers/followees. Examples of different profile-based behavior analysis activities are demonstrated.

Different from discovering traits or patterns, some work considers social network information to identify spammers. Ghosh investigated link farming on Twitter and proposed a ranking scheme to deter spam. Yang proposed a criminal account inference algorithm by exploiting criminal accounts' social relationships. Cao presented the SybilRank algorithm relying on social graph properties to rank users. Cui proposed a Hybrid Factor Non-Negative Matrix Factorization method to incorporate the predictive factors for user-post specific social influence prediction

3.2 DISADVANTAGE OF EXISTINTG SYSTEM:

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

4.PROPOSED SYSTEM

4.1 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 social 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.

4.2 ADVANTAGE OF PROPOSED SYSTEM:

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

5.IMPLEMENTATION

5.1 Service Provider

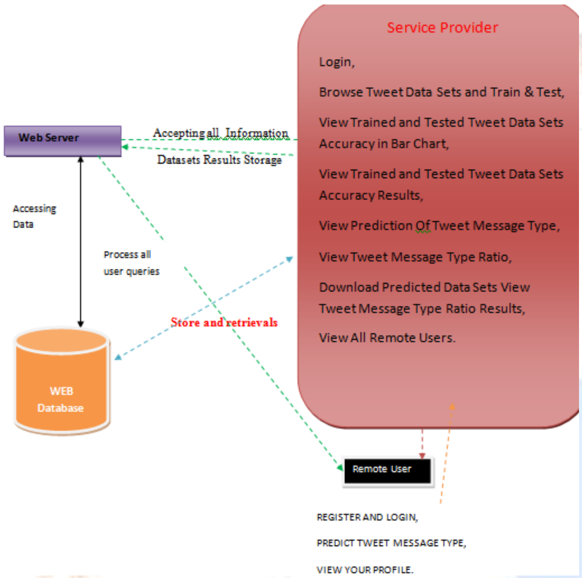
In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Browse and Train & Test Data Sets, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Prediction Of Attack Status, View Attack Status Ratio, Download Trained Data Sets, View Attack Status Ratio Results, View All Remote Users**5.2 View and Authorize Users**

In this module, the admin can view the list of users who all registered. In this, the admin can view the user’s details such as, user name, email, address and admin authorizes the users.

5.3 Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT ATTACK STATUS TYPE, VIEW YOUR PROFILE.

6. SYSTEM ARCHITECTURE


Service Provider
Login,
Browse Tweet Data Sets and Train & Test,
View Trained and Tested Tweet Data Sets Accuracy in Bar Chart,
View Trained and Tested Tweet Data Sets Accuracy Results,
View Prediction Of Tweet Message Type,
View Tweet Message Type Ratio,
Download Predicted Data Sets View Tweet Message Type Ratio Results,
View All Remote Users.
Web Server
Accepting all Information
Datasets Results Storage
Accessing Data
Process all user queries
Store and retrievals
WEB Database
Remote User
REGISTER AND LOGIN,
PREDICT TWEET MESSAGE TYPE,
VIEW YOUR PROFILE.

8.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 tf-idf to extract the user content matrices. However, spammer always posts some normal tweets to imitate the behavior of legitimate users. Thus, it is essential to distinguish the importance of tweets when we extract the user content matrix.

In future work, we will directly use raw tweets as the model input to learn user representations by distinguishing the importance of each tweet via deep learning techniques. After that, we plan to use graph neural networks to model social interactions among users.

9. FUTURE ENHANCEMENT

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.

10.REFERENCES

Aliaksandr Barushka 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.

Qiang Fu, Bo Feng, Dong Guo, and Qiang Li. Combating the evolving spammers in online social networks. Computers & Security, 72:60–73, 2018.

Zhijie Zhang, Rui Hou, and Jin Yang. Detection of social network spam based on improved extreme learning machine. Access, 8:112003– 112014, 2020.

Nexgate2013. 2013 state of social media spam. uploads/2013/09/Nexgate-2013-State-of-Social-Media-Spam-Research-Report.pdf.

Dehai Liu, Benjin Mei, Jinchuan Chen, Zhiwu Lu, and Xiaoyong Du. Community based spammer detection in social networks. In International

Conference onWeb-Age Information Management, pages 554–558. Springer, 2015.

Faiza Masood, Ahmad Almogren, Assad Abbas, Hasan Ali Khattak, Ikram Ud Din, Mohsen Guizani, and Mansour Zuair. Spammer detection and fake user identification on social networks. Access, 7:68140– 68152, 2019.

Sanjeev Rao, Anil Kumar Verma, and Tarunpreet Bhatia. A review on social spam detection: Challenges, open issues, and future directions. Expert Systems with Applications, 186:115742, 2021.

Chao Chen, Jun Zhang, Yi Xie, Yang Xiang, Wanlei Zhou, Mohammad Mehedi Hassan, Abdulhameed AlElaiwi, and Majed Alrubaian. A performance evaluation of machine learning-based streaming spam tweets detection. Transactions on Coputational social systems, 2(3):65– 76, 2015.

Xianghan Zheng, Zhipeng Zeng, Zheyi Chen, Yuanlong Yu, and Chunming Rong. Detecting spammers on social networks. Neurocomputing, 159:27–34, 2015.

Chao Yang, Robert Harkreader, and Guofei Gu. Empirical evaluation and new design for fighting evolving twitter spammers. Transactions on Information Forensics and Security, 8(8):1280–1293, 2013.

Zi Chu, Indra Widjaja, and Haining Wang. Detecting social spam campaigns on twitter. In International Conference on Applied Cryptography and Network Security, pages 455–472. Springer, 2012. [12] Mohd Fazil, Amit Kumar Sah, and Muhammad Abulaish. Deepsbd:

A deep neural network model with attention mechanism for socialbot detection. Transactions on Information Forensics and Security, 16:4211–4223, 2021.

Zulfikar Alom, Barbara Carminati, and Elena Ferrari. A deep learning model for twitter spam detection. Online Social Networks and Media, 18:100079, 2020.

Xinbo Ban, Chao Chen, Shigang Liu, Yu Wang, and Jun Zhang. Deeplearnt features for twitter spam detection. In 2018 International Symposium on Security and Privacy in Social Networks and Big Data (SocialSec), pages 208–212., 2018.

Saptarshi Ghosh, Bimal Viswanath, Farshad Kooti, Naveen Kumar Sharma, Gautam Korlam, Fabricio Benevenuto, Niloy Ganguly, and Krishna Phani Gummadi. Understanding and combating link farming in the twitter social network. In Proceedings of the 21st international conference on World Wide Web, pages 61–70, 2012.

Yin Zhu, Xiao Wang, Erheng Zhong, Nathan Liu, He Li, and Qiang Yang. Discovering spammers in social networks. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 26, pages 171–177, 2012.

Xia Hu, Jiliang Tang, Yanchao Zhang, and Huan Liu. Social spammer detection in microblogging. In Twenty-third international joint conference on artificial intelligence. Citeseer, 2013.

David M Beskow and Kathleen M Carley. Bot conversations are different: leveraging network metrics for bot detection in twitter. In 2018 /ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pages 825–832., 2018.

Jianshu Weng, Ee-Peng Lim, Jing Jiang, and Qi He. Twitterrank: finding topic-sensitive influential twitterers. In Proceedings of the third ACM international conference on Web search and data mining, pages 261–270, 2010.

Christian Thurau, Kristian Kersting, Mirwaes Wahabzada, and Christian Bauckhage. Convex non-negative matrix factorization for massive datasets. Knowledge and information systems, 29(2):457–478, 2011.

Chris HQ Ding, Tao Li, and Michael I Jordan. Convex and seminonnegative matrix factorizations. transactions on pattern analysis and machine intelligence, 32(1):45–55, 2008.

Paul Heymann, Georgia Koutrika, and Hector Garcia-Molina. Fighting spam on social web sites: A survey of approaches and future challenges. Internet Computing, 11(6):36–45, 2007.

Malik Mateen, Muhammad Azhar Iqbal, Muhammad Aleem, and Muhammad Arshad Islam. A hybrid approach for spam detection for twitter. In 2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST), pages 466–471., 2017.

Arushi Gupta and Rishabh Kaushal. Improving spam detection in online social networks. In 2015 International conference on cognitive computing and information processing (CCIP), pages 1–6., 2015.

Sangho Lee and Jong Kim. Warningbird: A near real-time detection system for suspicious urls in twitter stream. transactions on dependable and secure computing, 10(3):183–195, 2013.