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

ISSN 2454-9940
www.ijsem.org
Vol 11, Issuse.3 july 2023

SOCIAL SPAMMER DETECTION IN SOCIAL NETWORK

Dr.AR.SIVAKUMARAN[1], B.HIMANI[2], C. CHANDANA[3], CH. CHAITANYA JYOTHI[4], D. VARSHITHA[5]

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. 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: CNMFSD, Spammer, weibo, cloud, network.

INTRODUCTION

Social spammer detection in social networks, particularly focusing on platforms like Twitter, Facebook, and Sina Weibo. It emphasizes the negative impact of social spammers who invade users' privacy, disseminate unwanted information, share malicious content and links, and promote commodities. The rapid growth of social spammers and their collusion to form criminal communities pose significant challenges to researchers. Effective social spammer detection is crucial for improving user experience and the overall value of social systems. Over the past decade, researchers have explored different techniques to detect spammers, including link analysis and content analysis. Contentbased methods

[1]Associate Professor, [2,3,4,5]UG Students, Department of CSE Malla Reddy Engineering College for Women (UGC-Autonomous) Maisammaguda,Hyderabad-500100

ISSN 2454-9940
www.ijsem.org
Vol 11, Issuse.3 july 2023

typically involve analyzing and extracting user features and applying classification approaches such as support vector machines (SVM) to identify spammers. More recently, deep learning-based approaches have emerged, focusing solely on content for spammer detection. However, as spammers employ new strategies, these methods struggle to accurately detect them solely based on extracted features. Another category of methods relies on social network analysis to detect spammers. These methods assume that spammers cannot establish an excessively large number of social trust relations with legitimate users. Users with lower social influence or status in the network are classified as spammers. Unfortunately, relying solely on network information makes it difficult to distinguish between legitimate users and spammers.

PROJECT IDEA

The proposed system introduces a three-stage optimization model for social spammer detection, combining feature extraction and classifier learning. It utilizes Convex Non-negative Matrix Factorization (CNMF) and Non-negative Matrix Factorization (NMF) to generate latent features from content information, followed by training an SVM classifier. The system further refines these latent features using social interaction information as input for the classifier. Through iterative learning involving content information, social interaction regularization, and the classification model, the system achieves accurate spammer identification. The method is evaluated on a largescale real-world dataset from Twitter, demonstrating its ability to identify more spammers compared to baseline approaches.

EXISTING SYSTEM

Various machine learning approaches have been employed, such as SMFSR, which considers user activities and social following information, and SSDM, which incorporates text information and social following into a supervised model. Content analysis plays a crucial role, like focusing on detecting suspicious hashtags and URLs. Behavioral profiling is another direction, where deceptive user traits and profile characteristics are analyzed. Social network information is also utilized, exploring link farming and criminal account inference, respectively. Overall, the research

ISSN 2454-9940
www.ijsem.org
Vol 11, Issuse.3 july 2023

landscape encompasses a range of techniques targeting different aspects of social spammer detection.

PROPOSED SYSTEM

The system proposes a three-stage optimization model that conducts feature extraction and classifier learning simultaneously. We use Convex Non-negative Matrix Factorization (CNMF) and Nonnegative 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 new social interaction regularization term.

IMPLEMENTATION

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 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.

View and Authorize Users:

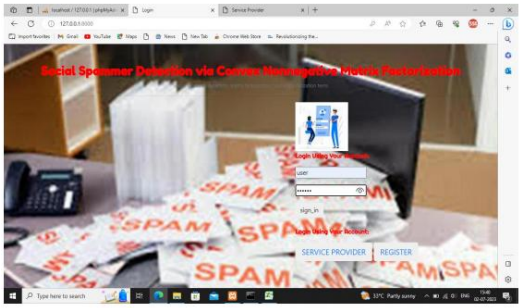
INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940
www.ijsem.org
Vol 11, Issuse.3 july 2023

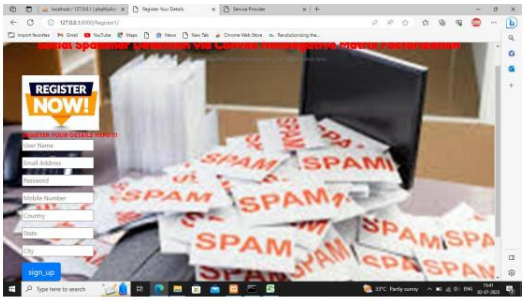
In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User:

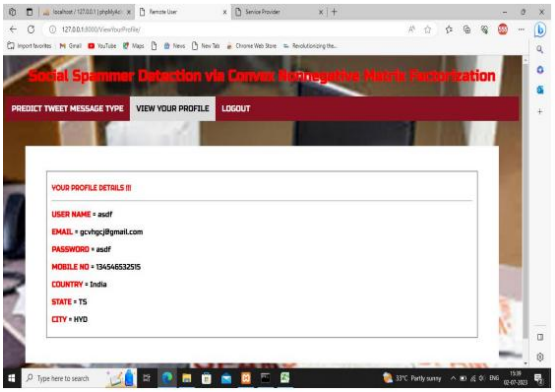
In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will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



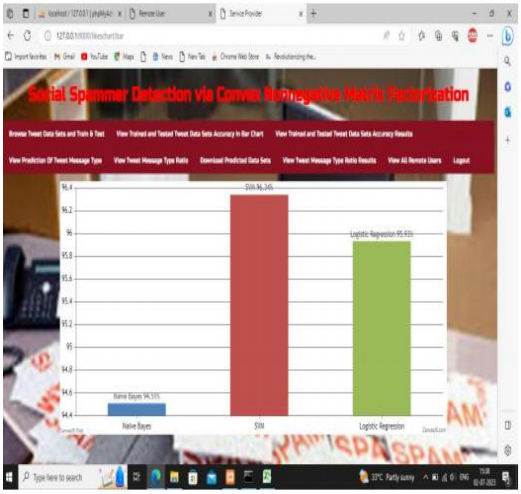
In above screen click on 'Registration' and fill the details if you don't already have a account then you can sgn_in using 'User Id' and 'password'.



In above screen we can see registration form fill the details in correct order so that you can register by giving details and sign_up



Here you can paste the tweet and see that it is 'spam tweet' or normal tweet' you can predict.



Conclusion

ISSN 2454-9940
www.ijsem.org
Vol 11, Issuse.3 july 2023

In this project, proposes a new framework called CNMFSD for social spammer detection by leveraging content and social interaction information. The method integrates users' interaction information and employs Convex-NMF to learn accurate latent user features for both legitimate users and spammers. Experimental results demonstrate that CNMFSD outperforms existing methods in detecting spammers. However, there are some limitations, such as the absence of social interaction graph consideration during classifier training and the use of tf-idf for user content matrix extraction, which may not effectively distinguish the importance of tweets. In future work, the authors plan to explore deep learning techniques to directly use raw tweets and employ graph neural networks to model social interactions among users.

REFERANCES

[1] Aliaksandr Barushka and Petr Hajek. Spam detection on social networksusing cost-sensitive feature selection and ensemble-based regularized deepneural 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] Zhijie 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] 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.

[6] 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. IEEE Access, 7:68140–68152, 2019.

[7] Sanjeev Rao, Anil Kumar Verma, and Tarunpreet Bhatia. A review on social spam detection: Challenges, open issues, and future

ISSN 2454-9940
www.ijsem.org
Vol 11, Issuse.3 july 2023

directions.Expert Systems with Applications,186:115742,2021.