Detection Of Social Network Spam Based On Improved Extreme Learning Machine

Rongala Durga Pavani, Sri.G.Ramesh Kumar, Sri.V.Bhaskara Murthy,
MCA Student, Assistant Professor, Associate Professor
Dept Of MCA
B.V.Raju College, Bhimavaram

ABSTRACT
With the rapid advancement of the online social network, social media like Twitter has been increasingly critical to real life and become the prime objective of spammers. Twitter spam detection refers to a complex task for the involvement of a range of characteristics, and spam and non-spam have caused unbalanced data distribution in Twitter. To solve the mentioned problems, Twitter spam characteristics are analyzed as the user attribute, content, activity and relationship in this study, and a novel spam detection algorithm is designed based on regularized extreme learning machine, called the Improved Incremental Fuzzy-kernel-regularized Extreme Learning Machine (I2FELM), which is used to detect the Twitter spam accurately. As revealed from the experience validation results, the proposed I2FELM can efficiently identify the balanced and unbalanced dataset. Moreover, with few characteristics taken, the I2FELM can more effectively detect spam, which proves the effectiveness of the algorithm.

I. INTRODUCTION

Over the past few years, the Internet has been leaping forward, and the intelligent terminals have been progressively popularized. Under such background, Online Social Networks (OSN) turns out to be a critical channel for people to acquire information, disseminate information, and make friends and get entertained. For the complexity of the online social network structure, the large-scale nature of the group, and the massive, rapid, and difcult traceability of information generation, the effects of user adoption, content creation, group interaction and information dissemination on online social networks thoroughly impact social stability, organizational management models, as well as people's daily work and life [1], [2].

Take Twitter for an example, the detection of Twitter spam can facilitate the process of analyzing, guiding and monitoring social network events, as well as regulating the management of networks. At present, the research challenges of Twitter spam are presented as follows, namely the feature selection and detection algorithm selection. The details are characterized below: 1) in feature selection, predecessor research often selects the identical type of characteristics e.g., content based and user prole-based characteristics for detection.

On the whole, since many types of characteristics of social network abnormal users are different from those of normal users, and it is not enough to accurately express the state of the data. 2) In algorithm selection, researchers primarily use

supervised machine learning algorithms to deal with spam detection in social networks. Based on the idea of classification, the researchers have designed numerical form characteristics to identify spam users. The supervised machine learning algorithm can be split into a single classification algorithm and an integrated classification algorithm (e.g., Support Vector Machine (SVM) [3], [8][11], [13], [14], meta-classifiers (Decorate, Logit Boost) [4], Naive Bayesian (NB) [6], [9], [11], Back Propagation Neural Network (BP) [16], Radial Basis Function (RBF) [18], Extreme Learning Machine (ELM) [8], [22], K-nearest Neighbor (KNN) [9], [19], Decision Tree (DT) [9], [20], Random Forest (RF) [5], [7][9], [23][26] and extreme Gradient Boosting (XGBoost) [31], [32]). 3) The real dataset of social networks exerts a long tail effect, i.e., it is an unbalanced dataset with a number of non-spam far exceeding the spam. When those supervised machine learning algorithms are detected on unbalanced dataset, their performance will decline.

Accordingly, an algorithm capable of effectively exploiting multi-dimensional characteristics and exhibiting continuous feasibility in the face of imbalance datasets should be adopted. By understanding and summarizing the research achievements of predecessors, four novel characteristics are proposed to express the Twitter datasets accurately and improve supervised machine learning algorithm to deal with unbalanced datasets to detect Twitter spam effectively. The details are illustrated below: 1) How to select the full category feature and pay attention to the correlation between the characteristics of the social network account helps enhance the accuracy of identifying spam users.

This study considers the Twitter spam attributes composed by the user attribute, content, activity and relationship to express the user characteristic and detect the spam accurately. 2) This study proposes a novel incremental Twitter spam assessment algorithm, termed as the Improved Incremental Fuzzy-kernel-regularized Extreme Learning Machine (I2FELM) to enhance the accuracy in dealing with the unbalanced data. 3) I2FELM is capable of enhancing the performance using Cholesky factorization without square root and composite kernel function. Besides,, it can automatically determine the optimal number of hidden layer nodes by gradually adding new hidden nodes one by one. 4) The I2FELM introduces the fuzzy weight as a method to address the unbalanced problem, which can apply to each input and facilitate the learning of output weights. 5) On the public dataset and the collected dataset, a range of index parameters and experimental verification methods are adopted to ascertain the performance of I2FELM, and spam is assessed based on the imbalance data problem and few characteristics. The article structure is arranged as follows. Section II presents the relevant work. Section III illustrates the novel Twitter spam detection model. Section IV discusses the experimental procedure, and Section V draws the conclusion of the study.

II. EXISTING SYSTEM

Benevenuto *et al.* [3] considered two attribute sets, namely, content attributes and user attributes, to distinguish one user class from the other and exploited the mentioned characteristics as attributes of SVM process to classify users as either spam or non-spam. Lee *et al.* [4] conducted the statistical analysis of the properties of the mentioned spam profiles to create spam classifier to actively filter out existing and novel spam. Based on the mentioned profile characteristics, the authors developed meta-classifiers (Decorate, LogitBoost, etc.) to identify previously unknown spam. Stringhini *et al.* [5] initially created a set of honey net accounts
(honey-profiles) on Twitter and then identified multiple characteristics that allow authors to detect spam. Lastly, the RF.model was built to detect spam and employed in a Twitter dataset. Wang [6] developed the novel content-based characteristics and graph-based characteristics to facilitate spam detection; besides, a Bayesian classification algorithm was adopted to distinguish the suspicious behaviors from normal ones.

Chu *et al.* [7] presented the collective perspective and focused on identifying spam campaigns that manipulate multiple accounts to spread spam on Twitter. An automatic classification system was designed based on RF and a variety of characteristics, i.e., individual tweet/account levels to classify spam campaigns. In Meda *et al.*'s work [8], a standard Principal Component Analysis (PCA) algorithm was exploited to reduce the dimensionality of the 62 feature to the 20 characteristics, 10 characteristics, and 5 characteristics, and then three different machine learning algorithm (SVM, ELM, RF) were adopted to support spam detection in Twitter. Wang *et al.* [9] studied the suitability of five classification algorithms of Bayesian, KNN, SVM, DT, and RF at the detection stage; they took four different feature sets of user characteristics, content characteristics, n-grams, and sentiment characteristics to the social spam detection task. Zheng *et al.* [10] extracted a set of characteristics from content-based and user-based feature and applied into SVMbased spam detection algorithm. Chen *et al.* [11] built a hybrid model that uses SVM and NB to distinguish suspect users from normal ones based on the user-based characteristics and content-based characteristics. During the assessment, the authors assessed the impact of different factors on spam detection performance, covering discretization of functionality, size of learning data, and data related to time. Chen *et al.* [12] proposed an Lfun approach to identify the ``Spam Drift'' problem in statistical features based Twitter spam detection. They compared Lfun to four traditional machine learning algorithms and evaluated the performance of Lfun approach in terms of overall accuracy, F-measure and Detection Rate. He *et al.* [13] proposed an analysis approach based on information entropy and incremental learning to study how various features affect the performance of an RBF-based SVM spam detector, through this effort, they attempted to increase the awareness of a spam by sensing the features of a spam.

Disadvantages

- In the existing work, the system considers user profile features, which can easily be modified by malicious spams.
- This system less effective due to absence of Extreme learning machine (ELM).

III. PROPOSED SYSTEM

By understanding and summarizing the research achievements of predecessors, four novel characteristics are proposed to express the Twitter datasets accurately and improve supervised machine learning algorithm to deal with unbalanced datasets to detect Twitter spam effectively. The details are illustrated below:

1) How to select the full category feature and pay attention to the correlation between the characteristics of the social network account helps enhance the accuracy of identifying spam users. This study considers the Twitter spam attributes composed by the user attribute, content, activity and relationship to express the user characteristic and detect the spam accurately.

2) This study proposes a novel incremental Twitter spam assessment algorithm, termed as the Improved Incremental Fuzzy-kernel-regularized Extreme Learning Machine (I2FELM) to enhance the accuracy in dealing with the unbalanced data.

3) I2FELM is capable of enhancing the performance using Cholesky factorization without square root and composite kernel function. Besides,, it can automatically determine the optimal number of hidden layer nodes by gradually adding new hidden nodes one by one.

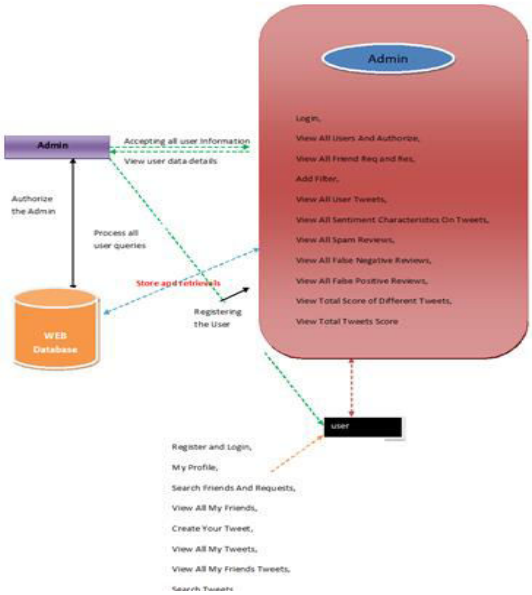
4) The I2FELM introduces the fuzzy weight as a method to address the unbalanced problem, which can apply to each input and facilitate the learning of output weights.

5) On the public dataset and the collected dataset, a range of index parameters and experimental verification methods are adopted to ascertain the performance of I2FELM, and spam is assessed based on the imbalance data problem and few characteristics.

Advantages

- ❖ The system is more effective due to implementation of Extreme learning machine (ELM).
- ❖ The system is more accuracy due to the assessment accuracy of the SVM.

IV. SYSTEM ARCHITECTURE



V. IMPLEMENTATION

- **Admin Server**

In this module, the Admin has to login by using valid user name and password. After login successful he can perform some operations such as View All Users And Authorize,View All Friend Req and Res,Add Filter,View All User Tweets,View All Sentiment Characteristics On Tweets,View All Spam Reviews,View All False Negative Reviews, View All False Positive Reviews, View Total Score of Different Tweets,View Total Tweets Score.

Friend Request & Response

In this module, the admin can view all the friend requests and responses. Here all the requests and responses will be displayed with their tags such as Id, requested user photo, requested user name, user name request to, status and time & date. If the user accepts the request then the status will be changed to accepted or else the status will remains as waiting

- **User**

In this module, there are n numbers of users are present. User should register before performing 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 can perform some operations like My Profile,Search Friends And Requests,View All My Friends,Create Your Tweet,View All My Tweets,View All My Friends Tweets,Search Tweets.

Searching Users to make friends

In this module, the user searches for users in Same Network and in the Networks and sends friend requests to them. The user can search for users in other Networks to make friends only if they have permission.

VI. CONCLUSION

This study presents a novel Twitter spam detection method, in which the feature set consists of user attribute, content, activity and relationship in the online social network for identifying the real spam. Moreover, the spam assessment algorithm is I2FELM, which uses fuzzy weights to resolve an unbalanced data problem for the accuracy enhancement. Furthermore, Cholesky factorization without square root and composite kernel function are employed to enhance performance. Also, the reasonable number of hidden nodes can be automatically determined. By the validation of experience, the proposed I2FELM can apply to the multi-dimension balanced or unbalanced datasets, and it has achieved high performance to assess the spam in the online social network. In the subsequent study, the emphasis will be placed on the following research directions. First, more factors will be considered to identify spam precisely (e.g., semantic analysis and emotion analysis). Also, we plan to exploit feature selection method and oversampling [21], [28], [29] to select a proper feature sets and improve model adaptation. On the other hand, to address insufficient labeled data in the social network, semi supervised learning method will be substituted to I2FELM model to detect Twitter spam automatically based on a small amount of labeled data.

REFERENCES

[1] M. Chakraborty, S. Pal, R. Pramanik, and C. Ravindranath Chowdary,``Recent developments in social spam detection and combating techniques: A survey," Inf. Process. Manage., vol. 52, no. 6, pp. 10531073, Nov. 2016.

[2] R. K. Dewang and A. K. Singh, ``State-of-art approaches for review spammer detection: A survey," J. Intell. Inf. Syst., vol. 50, no. 2, pp. 231264, Apr. 2018.

[3] F. Benevenuto, G. Magno, T. Rodrigues, and V. Almeida, ``Detecting spammers on Twitter," in Proc. CEAS, vol. 6, 2010, p. 12.

[4] K. Lee, J. Caverlee, and S. Webb, ``Uncovering social spammers: Social honeypots + machine learning," in Proc. 33rd Int. ACM SIGIR Conf. Res.Develop. Inf. Retr. (SIGIR), 2010, pp. 435442.

[5] G. Stringhini, C. Kruegel, and G. Vigna, ``Detecting spammers on social networks," in Proc. 26th Annu. Comput. Secur. Appl. Conf. (ACSAC), 2010, pp. 19.

[6] A. H. Wang, ``Don't follow me: Spam detection in Twitter," in Proc. Int. Conf. Secur. Cryptogr. (SECRYPT), Jul. 2010, pp. 110.

[7] Z. Chu, I. Widjaja, and H. Wang, ``Detecting social spam campaigns on Twitter," in Proc. Int. Conf. Appl. Cryptogr. Netw. Secur. Cham, Switzerland: Springer, 2012, pp. 455472.

[8] C. Meda, F. Bisio, P. Gastaldo, and R. Zunino, ``A machine learning approach for Twitter spammers detection," in Proc. Int. Carnahan Conf. Secur. Technol. (ICCST), Oct. 2014, pp. 16.

[9] B. Wang, A. Zubiaga, M. Liakata, and R. Procter, ``Making the most of tweet-inherent features for social spam detection on Twitter," 2015, arXiv:1503.07405. [Online]. Available: http://arxiv.org/abs/1503.07405

[10] X. Zheng, Z. Zeng, Z. Chen, Y. Yu, and C. Rong, ``Detecting spammers on social networks," Neurocomputing, vol. 159, pp. 2734, Jul. 2015.

[11] C. Chen, J. Zhang, Y. Xie, Y. Xiang,W. Zhou, M. M. Hassan, A. AlElaiwi, and M. Alrubaian, ``A performance evaluation of machine learning-based streaming spam tweets detection," IEEE Trans. Comput. Social Syst., vol. 2, no. 3, pp. 6576, Sep. 2015.

[12] C. Chen, Y. Wang, J. Zhang, Y. Xiang, W. Zhou, and G. Min, ``Statistical features-based real-time detection of drifted Twitter spam," IEEE Trans. Inf. Forensics Security, vol. 12, no. 4, pp. 914925, Apr. 2017.

[13] H. He, A. Tiwari, J. Mehnen, T. Watson, C. Maple, Y. Jin, and B. Gabrys, ``Incremental information gain analysis of input attribute impact on RBFkernel SVM spam detection," in Proc. IEEE Congr. Evol. Comput. (CEC), Jul. 2016, pp. 10221029.

[14] S. Teng, N.Wu, H. Zhu, L. Teng, andW. Zhang, ``SVM-DT-based adaptive and collaborative intrusion detection," IEEE/CAA J. Automatica Sinica, vol. 5, no. 1, pp. 108118, Jan. 2018.

[15] T. Wu, S. Wen, Y. Xiang, and W. Zhou, ``Twitter spam detection: Survey of new approaches and comparative study," Comput. Secur., vol. 76, pp. 265284, Jul. 2018.