IJMTST is UGC Approved Journal with Journal Id: 43137

International Journal for Modern Trends in Science and Technology, 8(S08): 141-144, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST08S0825
Available online at: https://www.ijmtst.com/vol8si08.html

Novel Factor Graph Based Model to Prevent Fraudulent Behaviour in MSN's

Dr. D J Samatha Naidu[1] | G.Maheswari[2]

Principal, Annamacharya PG college of computer studies, Rajampet.
Annamacharya PG college of computer studies, Rajampet.

To Cite this Article
Dr. D J Samatha Naidu and G.Maheswari. Novel Factor Graph Based Model to Prevent Fraudulent Behaviour in MSN's. International Journal for Modern Trends in Science and Technology 2022, 8(07), pp. 141-144. https://doi.org/10.46501/IJMTST08S0825

Article Info
Received: 26 May 2022; Accepted: 24 June 2022; Published: 28 June 2022.

ABSTRACT

The rapid development of modern communication technologies-in particular, (mobile) phone communications-has largely facilitated human social interactions and information exchange. However, the emergence of telemarketing frauds can significantly dissipate individual fortune and social health, resulting in potential slow down or damage to economics. Fraudulent activities are increasing rapidly with technology development of global communication in recent years. Millions of people suffer with frauds terribly. In existing works the problem of mining fraudsters and fraudulent strategies in a large-scale mobile network. By analyzing a one-month complete dataset of telecommunication metadata in shanghai with 698 million call logs between 54 million users, we find that fraudsters and non-fraudsters behave differently on communicating with others. In addition, fraudsters have preferences over users age and activity in phone communications when they choose targets. In proposed work, exploratory analysis, considered and proposed a novel semi-supervised model to distinguish fraudsters from non-fraudsters. The proposed model demonstrates that achieves a significant improvement comparing with several state-of-the-art baseline methods. The first challenge is caused by data sensitivity. How to unveil fraudulent strategy to better understand fraud is the second challenge. The third challenge is the label imbalance. To address the first and second challenge, we design and construct several exploratory analyzis on our real mobile network to study the behavior patterns of fraudsters. We disclose several fraudulent strategies based on our experiments. This project can potentially inform policymaking for government and mobile service providers

1. INTRODUCTION

Fraudulent review activities are increasing rapidly with the technology development of global communication in recent years. Millions of people suffer with frauds terribly. For instance, in China, phone fraud has been acknowledged as a significant problem. Estimations by both Qihoo1 and Tencent2 show that there are over 500 million phone frauds in 2016, which causes financial loses around 16.4 billion USD. Meanwhile, less than 3% of these cases are resolved. On August 29th of 2016, a college professor at Beijing was reported to have lost 2.67 million USD to a phone fraudster who claimed himself as a judicial officer. Besides the financial impact on individuals, the consequences of phone frauds have been even more tragic, even life-threatening. Fraud detection has attracted lots of efforts. However, the data availability and high sensitivity have caused this domain to be largely untouched by academia. Most existing work on fraud detection [1], [2], [3], [4] construct experiments on synthetic data or real-world data with limited scale. In this paper, we study on a large-scale mobile social network in real world, which covers a complete set of call logs in Shanghai and spans 30 days from September 1st to 30th, 2016. For each call log, the anonymous phone numbers, along with the starting and ending time of the conversation, are recorded. We also obtain annotations of

fraudsters made by crowd. Still, there are many other challenges remained. The first challenge is caused by data sensitivity, which forbids us to access the content information of each call log. It would be easier to detect fraudsters by monitoring particular topics in calls' content such as financial transfer. Without content information, due to privacy issue, we are forced to use meta information to make the inference. How could well educated people, like college professor in the above case, be swindled? Through our study, we show that user's information might has been seriously leaked and fraudsters select targets according to some strategy, instead of randomly (See details in Section 3). How to unveil **fraudulent strategy to better understand fraud is the second challenge.**It is worthwhile to highlight our contributions as follows: Based on a real phone-communication data, we disclose how fraudsters and non-fraudsters behave differently in mobile network. We study the "precisefraudulent strategy" and appeal to everyone to make the sure the protection of personal information has been brought to the forefront.We propose a novel framework to distinguish fraudsters from others in given mobile network. We validate the effectiveness of our model on alargescale mobile network in real world.

The proposed method is based on the probabilistic graphical model, which has also been applied in fraud reviews detection [31], social event extraction [32] , signal processing [33], etc. Many graph-based methods only consider limited types of features, which are mostly call frequency and call duration. In this work, by studying and distinguishing fraudsters from non-fraudsters, we propose several general and effective features. We believe that our features can benefit other work on fraud detection. Another difference between our work and others is that, to the best of our knowledge, we are the first to study fraudulent strategy.

2. EXISTING WORK

In many literatures, fraud detection is formulated as a binary classification problem. That is, given a set of phone numbers, predict whether each number is normal or fraudulent. For example, Weatherford et al. [23] utilize user profiles that store long-term information and train neural networks to differentiate fraud behavior andnormal one. Instead f neural networks, Yusoff [4] propose a model based on Gaussian mixed model (GMM) as the classifier. Dominik uses a threshold-type classification algorithm [24]. The major limitation of classifier-based methods is that, its performance is heavily influenced by annotations, and will be hurt when the label is sparse. In this work, we propose a semi-supervised learning framework to further utilize unknown labels and improve the performance.

DISADVANTAGES

- The system is not implemented by Ego-network characteristics.
- The system doesn't find the concept in which Fraudsters and non-fraudsters behave differently on communicating with others.

2.3.2. PROPOSED WORK WITH ADVANTAGES

The system designs and construct several exploratory analysis on our real mobile network to study the behavior patterns of fraudsters. We disclose several fraudulent strategies based on our experiments. For example, we find that fraudsters have preference on young people, and ones who are active in phone communications. We also find that it is better for us to hang up the fraudulent phone call immediately, instead of spending time on slagging off the fraudster to avoid receiving more fraudulent calls.

Based on our discoveries, we design a novel factor graph based model, FFD, to distinguish fraudsters. More informationand preference on choosing targets. We further propose a semi-supervised learning framework to utilize both the known and unknown labels and address the label sparsity challenge. According to our experiments, we see that our model achieves an improvement on F1 of 0.278 comparing with several state-of-the-art methods. It is worthwhile to highlight our contributions as follows:

ADVANTAGES

- Based on a real phone-communication data, we disclose how fraudsters and non-fraudsters behave differently in mobile network.
- The system studies the "precise fraudulent strategy" and appeal to everyone to make sure the protection of personal information has been brought to the forefront.

- The system proposes a novel framework to distinguish fraudsters from others in a given mobile network.
- The system validates the effectiveness of our model on a large scale mobile network in real world.

_

3. MODULES AND ITS DESCRIPTION

1.Admin

In this module, the Admin has to login by using valid user name and password. After login successful he can do some operations such as View All Users and Authorise,Add Products, View All Products, View All Products Reviews, View All Reviews by Rates, View All Products Recommendations.

2.User

In this module, there are n numbers of users are present. User should register before doing some operations. After registration successful he has to login by using authorized user name and password. Login successful he will do some operations like View My Profile, My Accounts, View All Products, View All My Purchased Items and Products, View All Recommended Items and Products for Me

3.Data Preparation

In this project, we use the way that websites like Olo and Travels guide was used. A review is associated with a rating score in a five-star scale. Each Products is associated with a category label and a textual description Loading All the Reviews of Customers for Analysis .

4.Analyzing Review

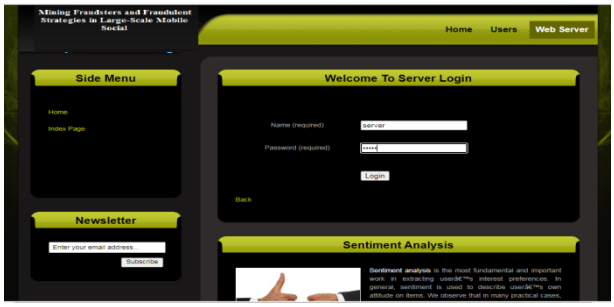
For detection of fraud online reviews, we start with raw text data. A dataset which was already labeled by the previous researchers, remove unnecessary texts like article and prepositions in the data. Then these text data are converted into numeric data for making them suitable for the classifier. Important and necessary features are extracted and then classification process took place Reviewers Information is taken and analyzed by EM Algorithm.

5.Detection of Fraud

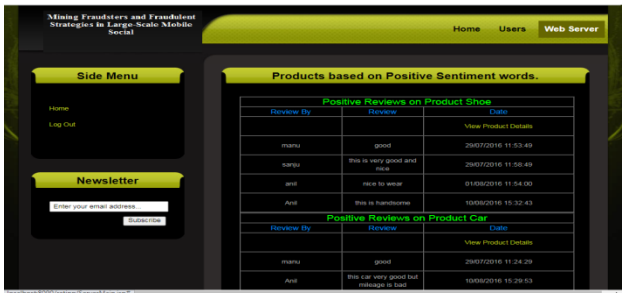
A.Each review goes through tokenization process first. Then, unnecessary words are removed and candidate feature words are generated. Each candidate feature words are checked against the dictionary and if it's entry is available in the dictionary then it's frequency is counted and added to the column in the feature vector that corresponds the numeric map of the word. Alongside with counting frequency, The length of the review is measured and added to the feature vector. Finally, sentiment score which is available in the data set is added in the feature vector. We have assigned negative sentiment as zero valued and positive sentiment as some positive valued in the feature vector

4. Simulation Results:

Screen 1: home page



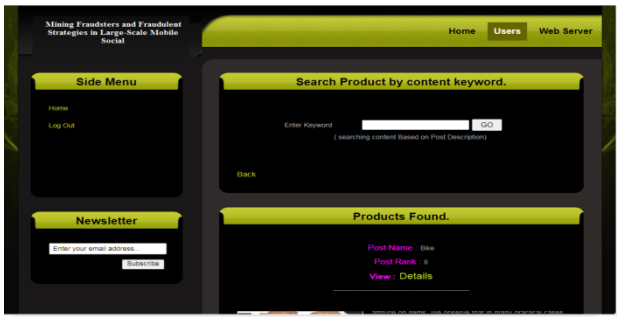
Screen 2:admin page



Screen 3 Report Showing Reviews on Products

Screen 4 View all Fraud

Screen 5:View Sentiments



Screen 6 View Products and Review IT

4. CONCLUSION

In this paper, we study the problem of mining fraudsters and fraudulent strategies in a large-scale mobile network. By analyzing a one-month complete dataset of telecommunication metadata in Shanghai with 698 million call logs between 54 million users, we find that fraudsters and nonfraudsters behave differently on communicating with others. In addition, fraudsters have preferences over users' age and activity in phone communications when they choose targets. Inspired by our exploratory analysis, we then propose a novel semi-supervised model to distinguish fraudsters from non-fraudsters.Experimental results demonstrate that our model achieves a significant improvement comparing with several state-of-the-art baseline methods. As for the future work, it is interesting to think about how to discover a fraud group, instead of an individual fraudster, consists of fraudsters with different roles and duties. Based on this, the collaboration patterns of different fraud groups can be disclosed.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Chengai Sun, Qiaolin Du and Gang Tian, "Exploiting Product Related Review Features for Fraud Review Detection," Mathematical Problems in Engineering, 2016.

[2] [2] A. Heydari, M. A. Tavakoli, N. Salim, and Z. Heydari, "Detection of review spam: a survey", Expert Systems with Applications, vol. 42, no. 7, pp. 3634–3642, 2015.

[3] [3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT), vol. 1, pp. 309–319, Association for Computational Linguistics, Portland, Ore, USA, June 2011.

[4] [4] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic Inquiry and Word Count: Liwc," vol. 71, 2001.

[5] [5] S. Feng, R. Banerjee, and Y. Choi, "Syntactic stylometry for deception detection," in Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers, Vol. 2, 2012.

[6] [6] J. Li, M. Ott, C. Cardie, and E. Hovy, "Towards a general rule for identifying deceptive opinion spam," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL), 2014.

[7] [7] E. P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, "Detecting product review spammers using rating behaviors," in Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM), 2010.

[8] [8] J. K. Rout, A. Dalmia, and K.-K. R. Choo, "Revisiting semi-supervised learning for online deceptive review detection," IEEE Access, Vol. 5, pp. 1319–1327, 2017.

[9] [9] J. Karimpour, A. A. Noroozi, and S. Alizadeh, "Web spam detection by learning from small labeled samples," International Journal of Computer Applications, vol. 50, no. 21, pp. 1–5, July 2012.