iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** VII **Month of publication:** July 2019

DOI: http://doi.org/10.22214/ijraset.2019.7152

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VII, July 2019- Available at www.ijraset.com

Survey on Spam Reviews and Spammer Community Detection Technique

Kajal Sonawane[1], Prof. Prashant Yawalkar[2]
[1, 2]*Dept. of Computer Engineering, Bhujbal Knowledge City, Nashik,*

Abstract: Online reviews and feedback of a product plays a vital role in human tendency to purchase those products. To affect the product sale spammer generates fake reviews on online social media platform. To identify spam reviews and spammer communities is the area of interest of this research work. In literature work, various spam detection techniques are proposed based on Review-Behavioral (RB) Based features, Review-Linguistic (RL) Based Features, User-Behavioral (UB) Based Features but none of the technique provide a simultaneous study of these features and weighting of the features along with finding the relationship among the spam users. The proposed work generates a heterogeneous network having users and reviews are the node type and the spam detection technique is mapped to the heterogeneous network problem. Spam reviews are detected using metapath. A feature weighting method is proposed to detect the relative feature importance. Using generated metapath for similar products spammer community detection is proposed.

Keywords: heterogeneous network, spam detection, fake reviews, social network.

I. INTRODUCTION

Customer judge the product based on online reviews. The sale of product depends on these online reviews. To stand in market with positive impression, growing number of businesses tries to receive online praises from their consumers. On the other hand to reach to superior position negative marketing of opponent product is promoted. To float positive or negative reviews based on the business needs artificial reviews are generated and posted on various online social media sites. As per the survey[4] 1/3 customer reviews on social media sites are suspicious. To generate fake reviews and float those reviews on social media site is called as spamming. In the proposed work analysis of this spam reviews is done using 3 techniques: Review-Behavioral (RB) Based feature, Review-Linguistic (RL) Based Features, User-Behavioral (UB) Based Features. The spam detection probem is converted in to analysis of heterogeneous information network problem. A network is established using users and their reviews as a node and metapaths are identified to detect the network spam reviews.For each type of review weight is assigned. Based on the feature weight importance of feature is calculated in spam detection technique.

II. LITERATURE WORK

In the literature work, various spam detection techniques are introduced. These techniques are mainly classified into three categories:

A. Linguistic Based

In linguistic approach natural language processing technique is used to identify similarity among multiple reviews. Feng et al.[3] uses n-gram and their composition. Some studies [2][4]Language modeling also include study for features between multiple reviews like capital words in statements. Lai et al.[5] proposes the probabilistic language modeling technique to find similarity between multiple reviews.

B. Behavioural Based

This technique is based on metadata analysis of a review. Metadata includes user behaviour and review behaviour analysis. Feng et al.[6] proposes a technique that studies metadata of review based on distribution of user rating on different products. 36 different behavior analysis techniques are proposed by Jindal et.al[7] with supervised learning mechanism.

C. Graph Based

Network based algorithms can be applied for spam detection. In this techniques heterogeneous network is established between reviews and users. Fei et al. in [8] proposed a network based Loopy Belief Propagation (LBP) algorithm to find burstiness in reviews to find spam reviews. Li et al. in [10] proposes a technique to analyze a review from multiple users from same IP address. For this heterogeneous network is established between users ,reviews and user IPs.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VII, July 2019- Available at www.ijraset.com

The study of all categories is done independently. Netsapm[1] is the technique proposed by Saeedreza Shehnepoor, Mostafa Salehi, Reza Farahbakhsh, and Noel Crespi. In this technique simultaneous study of Behavioral (RB) Based, Linguistic (RL) Based and graph based approach is proposed. EuijinChoo, Ting Yu , and Min Chi [9] detects the spammer groups in review systems. This is done using sentiment analysis on user interactions and graph theory. It analyses user relationship graph and annotating the graph by sentiment analysis and then pruning is done. According to the studies in literature, a common platform is required that make the study of spam reviews and relationship among various spam detection techniques along with the spammer community identification.

III. ANALYSIS AND PROBLEM FORMULATION:

Online product reviews on social media sites impacts on user buying behaviour of product. Product sale depends on the product reviews. To improve the product sale or to decrease the opponent product sale spammer generates spam review about the product. To analyse online reviews programmatically and identify the spam reviews is the domain of work.

Spam Reviews and Spammer Community Detection using Heterogeneous Network Information Spam detection can be done using various techniques called as review features such as analysing review Linguistic, time frame and threshold, Burstiness of reviews written by a single user, average of a users negative ratio, etc. To provide simultaneous study of reviews features and apply feature weighting to analyze feature importance in spam detection along with the identification of spammer community detection.

IV. PROPOSED SYSTEM

The following figure 1 shows the steps involved in finding the spam reviews and spammer community:

When data is arrived from given training dataset, then first step is it will generate prior knowledge and depending on the prior knowlwdge, the incoming review dataset is processed in 3 steps: Network schema, metapath definition and creation and classification.

A. Generate Prior Knowledge

The first step is computing prior knowledge i.e. the initial probability of review ‘u’ being spam which is denoted as yu. It uses two versions as semi-supervised learning and unsupervised learning.

In the semi-supervised method,

yu = 1, if review u is labeled as spam in pre-labeled review.

yu = 0, if label of review is unknown.

In the unsupervised method, prior knowledge is realized by using,

$$y_u = \left(\frac{1}{L}\right)\sum_{l=1}^{L} f(xl_u) \qquad (1)$$

Where, $f(xl_u)$ is the probability of review ‘u’ being spam.

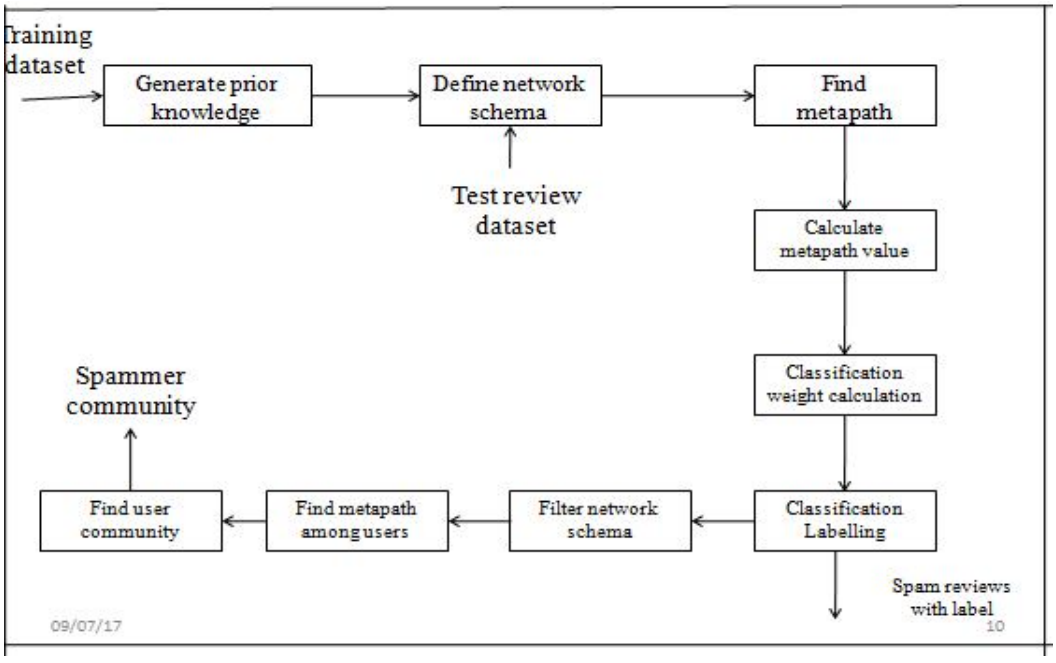


Fig 1: System Architecture

B. Network Schema

The next step is defining network schema based on a given list of spam features which determines the features engaged in spam detection. For example, If the list of features includes NR, ACS, PP1 and ETF, the output schema is as shown in below figure 2.

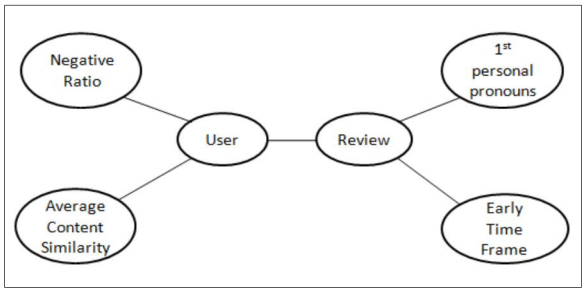


Fig.2: Example of Network Schema generated based on a given spam features list

C. Metapath Definition and Creation

A metapath is defined by a sequence of relations in the network schema. Table 1 shows all the metapaths used in the proposed framework. The length of user-based metapaths is 4 and the length of review-based metapaths is 2. There are different levels of spam certainty. A step function is used to determine these levels. Given a review 'u', the levels of spam certainty for metapath p_l is calculated using,

$$m_u^{p_l} = \frac{\lceil s * f(xl_u) \rceil}{s} \qquad (2)$$

Where, s is the number of levels.

After computing, two reviews u and v with the same metapath values for metapath p_l are connected to each other through that metapath and create one link of review network.

The proposed framework uses s = 20 levels.

Notation	Type	Semantic
R-DEV-R	RB	Reviews with same rate deviation from average item rate
R-U-NR-U-R	UB	Reviews written by different users with same negative ratio
R-ETF-R	RB	Reviews with same released date related to item
R-U-BST-U-R	UB	Reviews written by different users in the same burst
R-RES-R	RL	Reviews with same number of exclamation sentences containing '!'
R-PP1-R	RL	Review with same number of first Person Pronouns
R-U-ACS-U-R	UL	Reviews written by different users with same average content similarity using cosine similarity score
R-U-MCS-U-R	UL	Reviews written by different users with same maximum content similarity using cosine similarity score

Table 1: Metapaths used in the proposed framework

D. Classification

The classification part includes two steps as,

1) *Weight Calculation:* It determines the importance of each feature in spotting spam reviews. This step computes the weight of each metapath. This step will be able to compute the weight of each relation path, which will be used in the next step to estimate the label of each unlabeled review.

To compute the weight of metapath p_i, for i = 1,...,L where L is the number of metapaths, the following equation is used,

$$W_{p_i} = \frac{\sum_{r=1}^{n}\sum_{s=1}^{n} mp_{r,s}^{p_i} * y_r * y_s}{\sum_{r=1}^{n}\sum_{s=1}^{n} mp_{r,s}^{p_i}} \qquad (3)$$

Where, n denotes number of reviews.

2) *Labelling:* Let, $Pr_{u,v}$ be the probability of unlabeled review u being spam by considering its relation with spam review v. To estimate Pru, following equation is used,

$$p_{r_{u,v}} = 1 - \pi_{i=1}^{L} 1 - mp_{u,v}^{p_i} * W_{p_i}$$

$$p_{r_u} = avg\left(p_{r_{u,1}}, p_{r_{u,2}}, \dots, p_{r_{u,n}}\right) \qquad (4)$$

Where, n denoted the number of reviews.

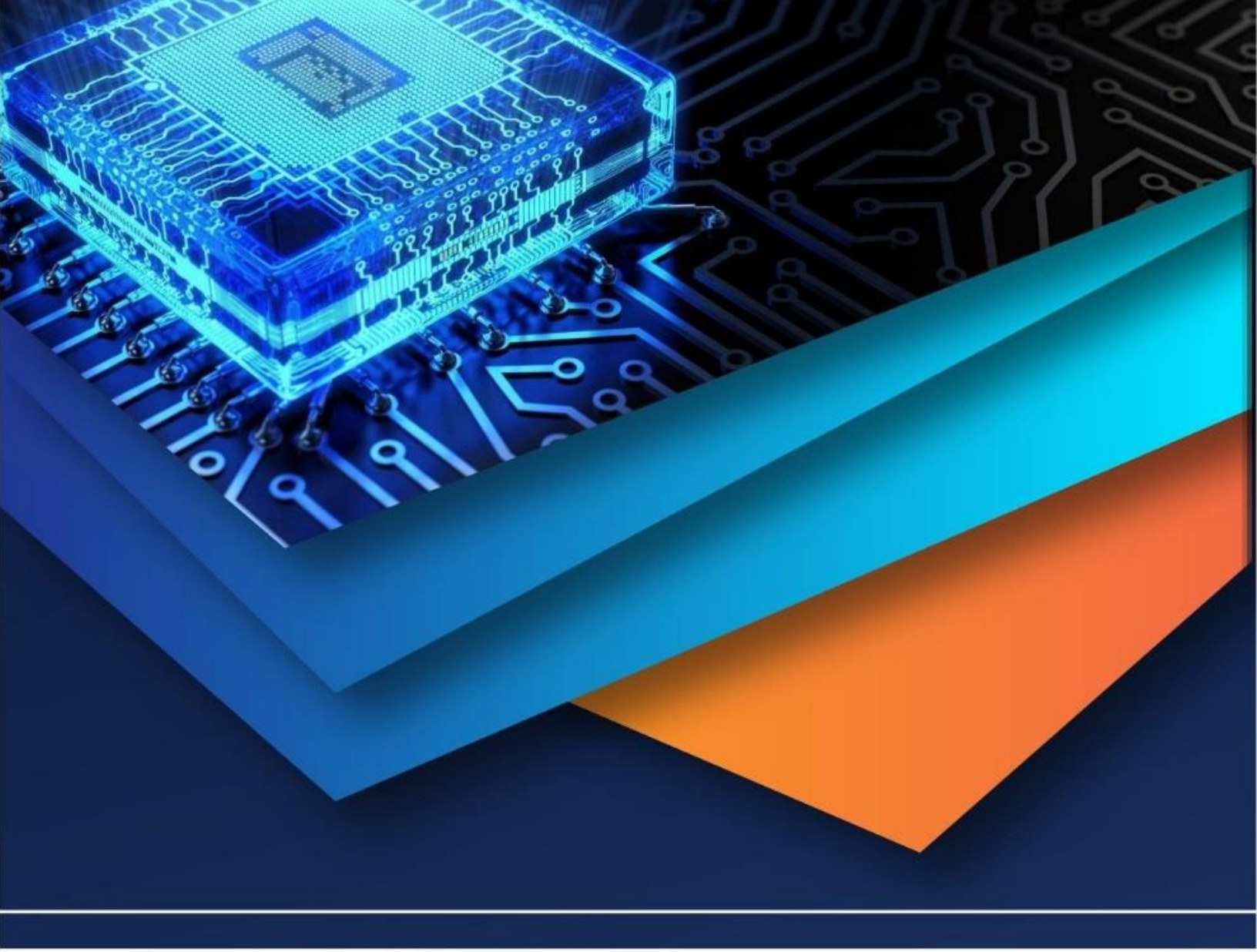
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue VII, July 2019- Available at www.ijraset.com

V. CONCLUSION

The proposed framework is designed for detection of spam reviews and spammer communities using Heterogeneous Information Network, which is based on metapath concept and graph based methods to label reviews. Four types of features will be used for detection of spam as Review-Linguistic Based features, Review-Behavioral Based features and User-Behavioral Based features. We can determine the importance of each feature in classification of review and calculate the weight of each review. We can also label each review by calculating the probability of each review being spam or not. The spammer communities will also detected by the framework.

REFERENCES

[1] Saeedreza Shehnepoor, Mostafa Salehi*, Reza Farahbakhsh, Noel Crespi NetSpam:a Network-based Spam Detection Framework for Reviews in Online Social Media IEEE Transactions on Information Forensics and Security 2017.
[2] J. Donfro, A whopping 20 percent of yelp reviews are fake.
[3] M. Ott, C. Cardie, and J. T. Hancock. Estimating the prevalence of deception in online review communities. In ACM WWW, 2012.
[4] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock. Finding deceptive opinion spam by any stretch of the imagination.In ACL, 2011.
[5] Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. SIAM International Conference on Data Mining, 2014.
[6] N. Jindal and B. Liu. Opinion spam and analysis. In WSDM, 2008.
[7] F. Li, M. Huang, Y. Yang, and X. Zhu. Learning to identify review spam. Proceedings of the 22nd International Joint Conference on Artificial Intelligence; IJCAI, 2011.
[8] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013.
[9] Choo E., Yu T., Chi M. (2015) Detecting Opinion Spammer Groups Through Community Discovery and Sentiment Analysis. In: Samarati P. (eds) Data and Applications Security and Privacy XXIX. DBSec 2015. Lecture Notes in Computer Science, vol 9149. Springer, Cham.
[10] H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY