

SPAM REVIEW DETECTION USING USER AND FEATURE REVIEWS

Ms. Vaishnavi P. Kshirsagar¹, Mrs. D. K. Jadhav²

¹Post Graduate student of KIT's college of engineering Kolhapur, ²Assistant Professor of KIT's college of Engineering Kolhapur
^{1,2}Department of Computer science & Engineering KIT's College of Engineering Kolhapur Maharashtra, India.

Abstract – web portals play an important function in information growth which is considered as a significant fountainhead for manufacturers in publicity/promotion campaigns also for buyers in choosing services and products. Written reviews/Ratings are considered by companies as a product feedback. The reviews are of two types: 1)text review 2)rating. Actuality that anybody using any identity can give comments as review, which offers a golden window of opportunity for spammers to write fake reviews designed to mislead user's opinion. Reviews given for online stores are significant resource to assist buyers/customers can make smart choice for their purchase. Hence due to this reason, spammer acts the review system these spammers are usually invited or hired or enticed by businesses/companies to write fake/fraud reviews to promote or fame their services and products, and/or to divert customer's attention from their other companies or competitors.

Index Terms—Spam Reviews, Internet, Spammer, Social Media Text Review, Fake Review, Ratings, meapath identification/creation

1. Introduction:

Online web portals plays a significant part in human life. Specifically Online Social Media portals plays important part in today's digital world where it is an important platform for sellers for advertising campaigns also it is an important platform for buyers for choosing services and products. From the past few years, buyers/ customer's decisions-making processes depends on the reviews written by other users, and positive review and negative reviews motivating and demotivating people in their purchasing of products and services. Also, reviews written by user helps sellers and service providing companies to enhance the quality of their products and services. Thus written reviews and ratings of products and services becomes a significant part in success or progress of a business. Positive reviews gives benefits for a supplier and company, where negative reviews leaves bad effect and because of this cause business losses. Also these positive or negative reviews helps users/buyers in decision making. Reviews considered by company as a product feedback. There are two types of reviews: 1) Text Reviews 2) Rating. The fact that people with any identity can write reviews, this offers a golden opportunity for spammers.

Spammer intentionally writes fake reviews to mislead user's opinion or choice. 90 % people make choices for their purchases on the bases of reviews written by buyers. Therefore spammers are hired or enticed by companies to write fake reviews to promote their products and services review system has become a target of spammers who are usually hired or enticed by companies to write fake reviews to promote their products and services, and/or to distract customers from their competitors. Due to this, the review system has become easy target for spammers to mislead customers

2. Literature Review:

Reviews are used tremendously by users, customers or buyer and companies or organizations to make purchase or buy products and to make business decisions [1]. Some reviews are written about products or services like how good a product or service to change user's or customer's perception. These kind of reviews are considered as a spam reviews[2].One of them is a classifier that can calculate feature weights that show each feature's level of importance in determining spam reviews. The general raw concept

of our proposed framework is to model a given review dataset as a Heterogeneous Information Network (HIN) and to map the problem of spam detection into a HIN classification problem [4].

There are generally three types of spam reviews:

- 1) Untruthful opinion spam
- 2) Reviews on brands only
- 3) Non-Reviews.

Spam detection can be regarded as a classification problem with two classes, spam and non-spam [6].

There are three main types of information related to a review:

- 1) The content of review,
- 2) The reviewer who wrote the review,
- 3) The product being reviewed.

There are three types of features:

- 1) Review centric features,
- 2) Reviewer centric features,
- 3) Product centric features [9].

Further these types are classified as behavioral and linguistic based features. Content spam tries to add irrelevant or remotely relevant words in target pages to fool search engines to rank the target pages high [7]. There are large no of duplicate and near-duplicate reviews. The detection of duplicate and near-duplicate reviews are done by using machine learning algorithm [8].

3. Research Gap

Based on literature review following research gaps are identified:

1. It is hard to identify the singleton review as a spam or non-spam [4]
2. The classification of the users is difficult as one user has more than one account [6].
3. The reviews given in the form of ratings (star) are difficult to recognize as fake[10]
4. The review given by the spammer which is true-positive is not classified as a spam review [11].

4. Research Objectives

Considering stated research gaps, following objectives are defined in proposed study:

To organize user reviews on the basis of behavioral and linguistic features.

1. To implement generic graph based algorithm to determine the weights of features.
2. To classify test reviews into spam and non-spam labeling categories.
3. To test and analyze the performance against standard benchmarks.

5. Proposed Methodology

The proposed methodology has been described in 4 phases as follows.

- a. Prior Knowledge
- b. Network Schema Definition

- c. Metapath Definition and Creation
- d. Classification

1) Prior Knowledge:

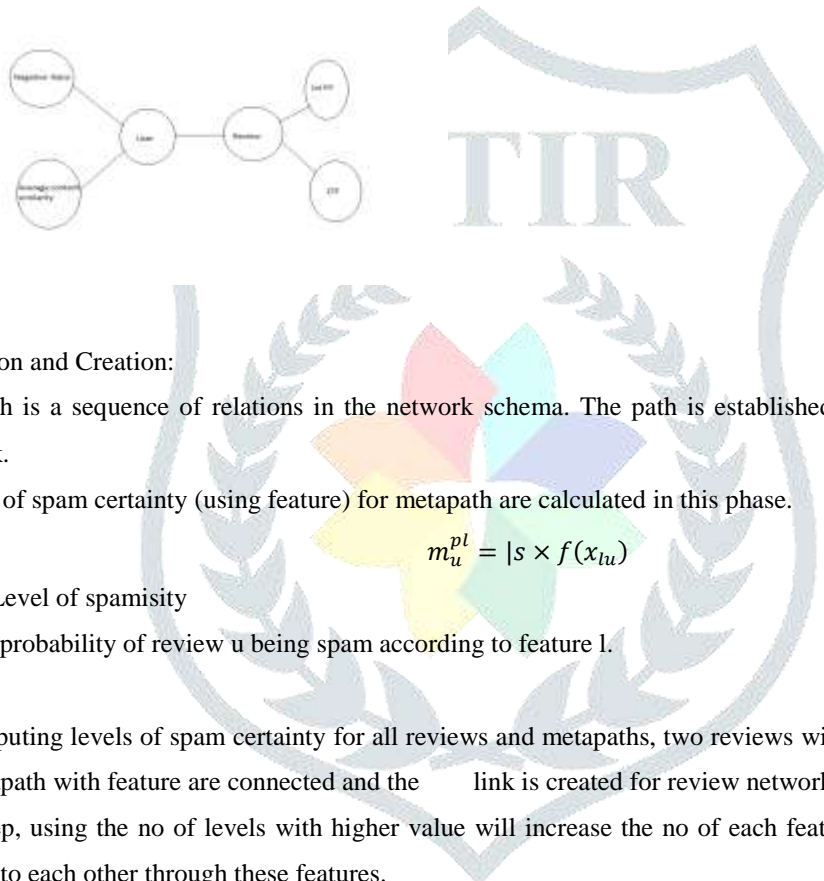
This phase computes the probability of review being spam. The proposed version works in two versions:

- Supervised Learning

In Supervised learning method, the initial probability of review being spam according to feature which is from set of features is calculated.

2) Network Schema:

- 1) The list of spam features which determines the features engaged in spam detection is used to design network schema.
- 2) The metapath is calculated at this phase.



3) Metapath Definition and Creation:

- 1) A metapath is a sequence of relations in the network schema. The path is established using the features used in the framework.
- 2) The levels of spam certainty (using feature) for metapath are calculated in this phase.

$$m_u^{pl} = |s \times f(x_{lu})$$

Where s=Level of spamisity

f(x_{lu})= probability of review u being spam according to feature l.

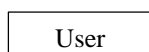
- 3) After computing levels of spam certainty for all reviews and metapaths, two reviews with the same metapath values for some metapath with feature are connected and the link is created for review network.
- 4) In next step, using the no of levels with higher value will increase the no of each feature's metapath. Reviews can be connected to each other through these features.
- 5) The spamicity of the review with maximum no of levels is calculated.

4) Classification:

It consists two steps:

- 1) Weight calculation which determines the importance of each spam feature in spotting spam reviews.
- 2) Labeling which calculates the final probability of each review being spam.

6. Block Diagram:



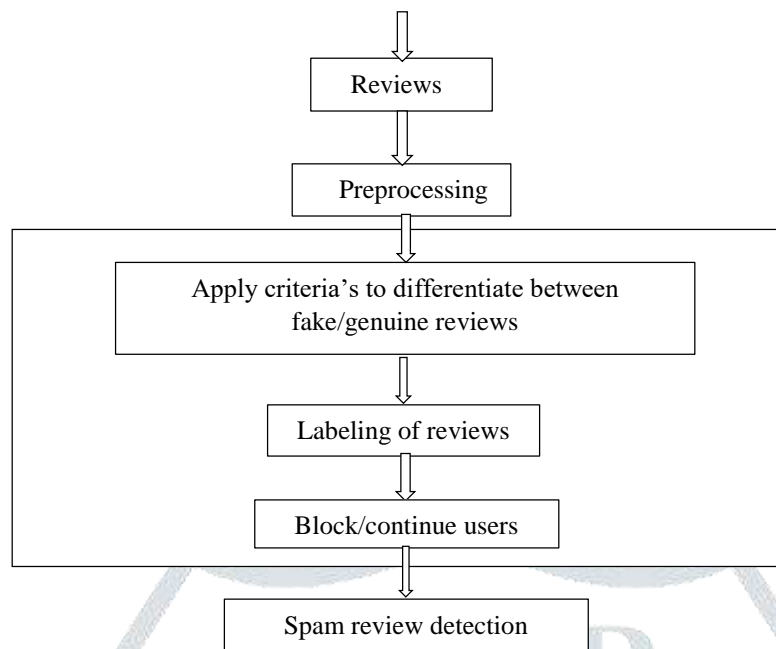


Fig1: Raw system architecture

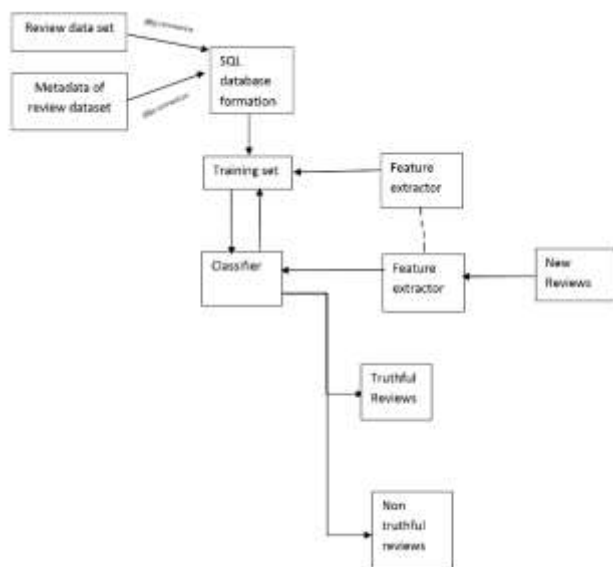


Fig2: Detailed working of Spam Review Detection System

7. Disadvantages of Existing System

1. There is no information filtering concept in online social network.
2. People believe on the written reviews in their decision-making processes and positive or negative reviews encouraging or discouraging them in their selection of products and services.
3. Anyone create registration and gives comments as reviews for spammers to write fake reviews designed to misguide users' opinion.
4. Less accuracy.
5. More time complexity

8. Advantages of Proposed System

1. We can identify the spam user using positive and negative reviews in online social media.
2. Display only trusted reviews to the users.
3. Our observations show that weights calculated by methapath concept can be significantly effective in discovering spam reviews and gives to a better performance.
4. It will helps to prevent misguide users by giving fake reviews.

9. Resources Required

Research Laboratory

(i) Software:

Operating System: Windows XP/7.

Coding Language: JAVA/J2EE

Data Base: MYSQL

Application development tool: Java

IDE: Net Beans

(ii) Hardware:

PC with standard configuration, min.4 GB RAM, 80 GB HDD.

Library Facilities: E-Journals, IEE

10. Conclusion

In this paper the concept introduces a spam detection system namely **Spam Review Detection: Using user and Feature reviews** based on a metapath algorithm also a graph based method to mark/label reviews depends on a rank based labeling approach. Proposed frameworks performance is calculated by using labeled dataset of Yelp website which is a dataset of reviews of NewYork based restaurants. The observations show that weights calculated by metapath concept can be very effective to identify spam/fake reviews and leads to a better performance. Also in addition, Spam Review Detection calculate the importance of each feature without even a train set, and it gives better results in the features'. After identifying four main categories for features observations show that the reviews behavioral category performs better than other categories.

Future work includes, metapath concept can be applied to other problems. For example, framework which are similar to each other can be used to find spammer communities. For finding spammer community reviews connected through spammer features group and reviews with highest similarity based on metapath concept are known as communities. In addition, utilizing the product features is an interesting future work on this study as we used features more related to spotting spammers and spam reviews.

References

1. 'Saeedreza Shehnepoor, Mostafa Salehi*, Reza Farahbakhsh, Noel Crespi NetSpam: a Network-based Spam Detection Framework for Reviews in Online Social Media <https://ieeexplore.ieee.org/document/7865975> DOI 10.1109/TIFS.2017.2675361'
2. 'J. Donfro, A whopping 20 % of yelp reviews are fake <http://www.businessinsider.com/20-percent-of-yelp-reviews-fake-2013-9>. Accessed: 2015-07-30.'
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. In 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. 'A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. Trueview: Harnessing the power of multiple review sites. In ACM WWW, 2015'.
10. 'B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah, K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Towards detecting anomalous user behavior in online social networks. In USENIX, 2014'.
11. H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.