International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

# Detection of Spam Reviews in Online Social Media Using Graph Based Methods

Renuka[1], Nandini Prasad K. S[2]
[1,2]Department of Information Science and Engineering, Dr. Ambedkar Institute of Technology, Bangalore
Email address: rrsuryawanshi11@gmail.com, iseofficial123@gmail.com

***Abstract*****—***These days, a major piece of individuals trusts on content in online social media like opinions and feedbacks on a subject or items. The obligation that anyone can leave spam audits about the item and administrations for various interests. Recognizing these spammers and spam content is broadly bantered about issue of research and regardless of the way that an impressive number of works have been done starting late towards this end yet so far systems set forward still scarcely recognize spam audits, and none of them demonstrate the significance of each removed element sort. A novel scheme is proposed named as NetSpam, which uses spam characteristics for representing audit datasets as heterogeneous data systems to outline spam recognition conspire into a characterization issue in such network. Utilizing the significance of spam highlights assist us with acquiring superior outcomes regarding different metrics on review datasets.*

***Keywords*****—** *Social Media, Social Network, Spammers, Spam Review, Fake Review, Heterogeneous Information Networks*.

## I. Introduction

Online social media portals play a powerful part in data spread which is assumed a vital hotspot for makers in their publicizing efforts as well with respect to clients in choosing items and administrations. In previous years, human depends on the composed audits in their basic leadership process and constructive/contrary audit engaging or weakening them in their choice of items and administrations. These audits in this way have transformed into primary concern in an advance of business while positive audits can shows profits for an organization, negative audits causes financial losses. The fact that anybody can leave remarks as audit, gives an attractive chance for spammers to compose fraud audits outline to misguide clients conclusion. These misguiding audits are then duplicated by the sharing capacity of internet based life and proliferation over the network. The audits composed to change customer's impression of how extraordinary a thing or an organization are supposed as spam [5], and are consistently formed in kind for cash. As appeared in [1], 20% of the audits in the howl site are all things considered spam audits. In another way, a lot of writing has been distributed on the systems used to recognize spam and spammers and additionally extraordinary kind of investigation on this subject [8], [9]. These methods can be classified into various classifications; some utilizing linguistic pattern [2-4], which depend on bigram and unigram, some are in light of behavioural examples that depend on highlights separated from designs in client conduct which are for the most part metadata-based [6-9]. Regardless of this incredible arrangement of endeavours, numerous angles have been missed or remained unsolved. One of them is a classifier that can ascertain element's level of significance in deciding spam audits. The common idea of proposed structure is to represent a given audit dataset as Heterogeneous data arrange [13], and to outline issue of spam identification into a HIN categorization issue. A weighting algorithm is at that point used to figure every segment weight. These weights are used to compute last names for audits utilizing both unsupervised and managed approaches. To appraise the proposed framework, we utilized two specimen audit datasets from howl and Amazon sites. As the consequence of the weighting step, we can utilize some highlights with more weights to get well precision with less time complexity. Features can be classified in four categories (audit-behavioural, client-behavioural, audit-linguistic, client linguistic), assists us to know how much each classification of features is provided to spam identification.

1. A latest weighting techniques for spam features is proposed to decide the respective significance of each feature and shows from ordinary reviews.
2. It enhances the exactness of time multifaceted nature which depends to the amount of highlights used to perceive spam audits.

## II. Preliminaries

As introduced before, we demonstrate the issue as a heterogeneous structure where hubs are either genuine elements in dataset like audits or items. To well recognize the proposed structure we exhibit an ideas of some portion of the concepts in heterogeneous data systems [14], [15].

Spam characteristic: In specific, the characteristic for clients and audits fall into four classes as follows.

### *1. Review-Linguistic (RL) Based Characteristic*

Attributes in this class depend on the audit itself and express straightforwardly from wording of the audit [6].

### *2. Review-Behavioural (RB) Based Characteristic*

The attributes depend on metadata and not simply the audit wording. The survey behavioural arrangement contains two attributes early time allotment edge rating deviation of audit [11].

### *3. User-Linguistic (UL) Based Characteristic*

Average content similarity, Maximum content similarity: spammers usually record their audits with indistinguishable layout and they support not to squander their opportunity to record a unique audit [11].

*4. User-Behavioural (UB) Based Characteristic*

Burstiness: Spammers normally record their spam audits in small duration of time for two purposes. First, because they need to effect readers and other clients, and second because they have to record more audits in small duration.

## III. System Architecture

There are two ways to detect spam audits i.e. spam review detection 1 process and spam review detection 2 process.

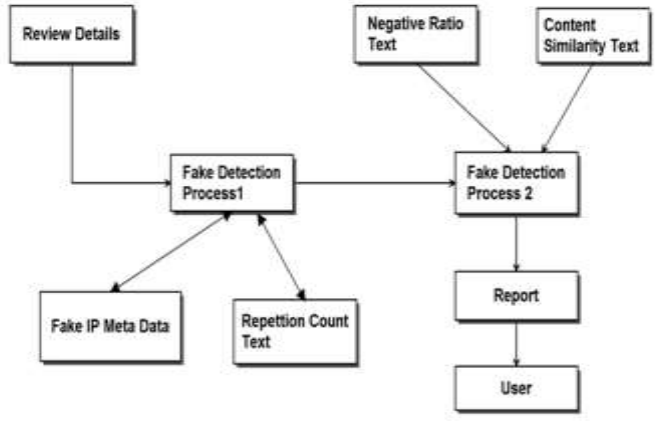

Fig. 1. System architecture.

### A. Fake Review Detection Process 1

In fake review detection 1 process, data will be read from database and checks whether IP address and User id fake or real based on metadata table and insert fake reviews into the fake review table and also checks whether the number of reviews from IP address are exceeding the threshold limit within the threshold limit, if any IP address exceeds threshold limit then the reviews will be inserted to fake review table and IP address will be inserted meta fake IP address table and reset the reviews will be inserted to real reviews table.

### B. Fake Review Detection Process 2

In fake review detection 2 process, reviews will be read from the real reviews table, Considering each reviews in the first level unnecessary words and special characters will be removed, In second level categorize each and every word is noun or adjective, In third level paring the noun and adjacent adjective, in the fourth level checks whether the adjective which is paired with the noun is negative or positive, in fifth level checks whether the maximum number of pairs are positive or negative, based on the maximum count of positive or negative, assign the review value as positive or negative, in the sixth level calculate and insert the two gram and three gram pairs into the database, in the seventh level calculate the count percentage, positive percentage and n-gram percentage of each user and add all the percentages and get total percentage threshold, if any user exceeds total percentage threshold, consider that user is fake and insert that user into the meta fake user table.

## IV. Related Work

In this section, we survey the related work of effort including those behaviour-based methods, linguistic based methods and graph based methods.

*1. Behaviour based methods*: Perceptive in this class nearly utilize audits metadata to express highlights those which are regular example of auditor behaviours [11]. Using different classification perceptive need various number of features to attain required implementation.

*2. Linguistic based method*: This perspective express linguistic based feature to find spam audits [13] and use unigram, bigram and their structure.

*3. Graph-based methods*: An abstract data type representing relations or connections. Concentrates in this classification centers to influence a chart between customers, to audit and things and utilize relations in the diagram and furthermore some system based calculations to group audits and client [11].

## V. Design Process

The following algorithm shows sentiment analysis and pre-processing algorithm.

```
Algorithm 1: Sentiment Analysis
Input: Data set with audits
Output: sentiment analysis with positive, negative and neutral count
Step 1: Let n be the audits
Step 2: for i=0 to n
Step 3: Count positive=0, negative=0, neutral=0;
Step 4: count=count+1;
Step 5: based on positive, negative and neutral count
Step 6: give sentiment process
Step 7: End for loop
Step 8: Stop
```

```
Algorithm 2: For Pre-processing
Input: audit - dataset, spam-feature - list, pre- labeled -audits
Output: features – importance(W)
Spamicity – probability(pr)
Step 1: Let n be the sentence
Step 2: for i=0 to n
Step 3: Remove unnecessary words from sentence
Step 4: Remove hyperlinks from sentence
Step 5: Remove special characters from sentence
Step 6: End for loop
Step7: Stop
```

## VI. Results and Discussion

When the file is loaded the first page is index.jsp. The home page is loaded initially because of iframe concept.

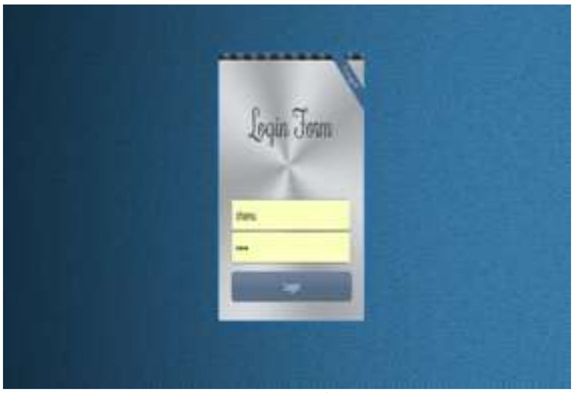

Fig. 2. Login form.

Figure 2 indicates login form. Initially user needs to enter user name and password in login form and then it will login.

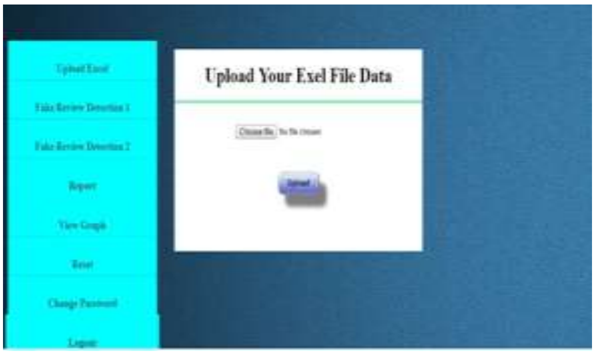

Fig. 3. Upload excel file.

Figure 3 indicates excel file data. Excel file contains file name, IP address, date, time and userid and user uploads this exel file.

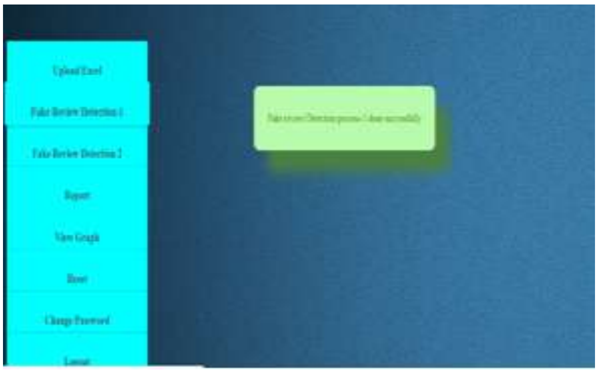

Fig. 4. Fake review detection 1.

Figure 4 indicates fake review detection process 1. Read Excel file which contains reviews, file path, IP address, date, time and userid and check Meta fake table for existing fake Userid and fake IP address. If the IP address or the userid exists in Meta fake table, insert those reviews in the fake review table and insert unchecked reviews into the real reviews table.

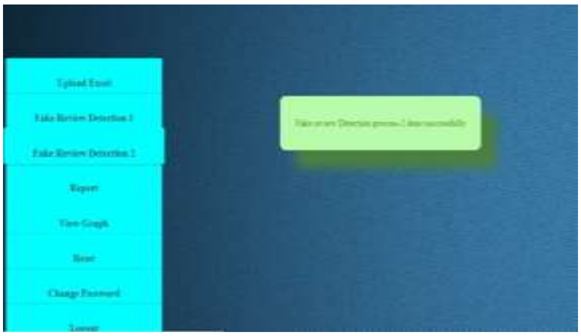

Fig. 5. Fake review detection 2.

Figure 5 shows fake review detection process 2. Read audits from the real audits table with respect to the userid and remove unnecessary words from the audits. Segregate the noun, adjective, positive, the total negative and neutral words and calculate count of positive, negative and neutral words of each audit.

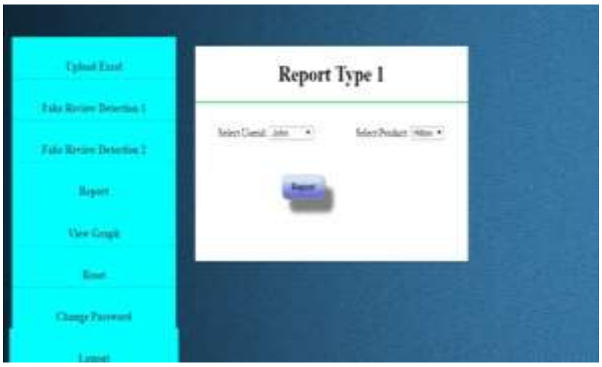

Fig. 6. Report type.

Figure 6 indicates report type 1. List of reports contains five options first is one user- one product, second is one product- all user, third is one user- all product, fourth is fake reviews by all users and fifth is meta data fake review.

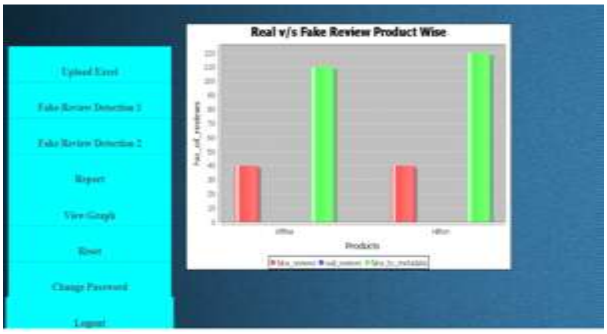

Fig. 7. Real v/s fake review graph.

Figure 7 shows the real v/s fake review wise and it shows the ratio of products and number of reviews.

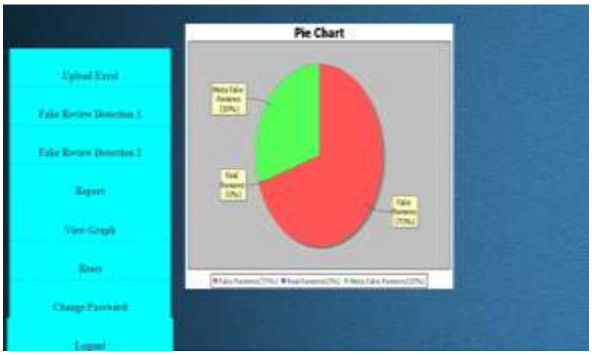

Fig. 8. Pie chart.

Figure 8 indicates pie chart graph that shows the percentage of fake reviews, real reviews and Meta fake reviews.

## VII. Conclusion

This survey presents a novel spam detection system in particular project in perspective of a metapath scheme and another graph based strategy to name reviews depending group-based naming method. The implementation of the proposed structure is assessed by using review datasets. Our perception determines that ascertained weights by using this metapath idea exceptionally strong in recognizing spam reviews and prompts a superior execution. In future metapath idea can be connected to other issue.

## References

[1] S. Mukherjee. S. Dutta, and G. Weikum. Credible Review Detection with Limited Information using Consistency Features, in book: Machine learning and Knowledge Discovery in Databases, 2016.

[2] M. Ott, C. Cardie, and J. T. Hancock. Estimating the prevalence of deception in online review communities. In ACM WWW, 2012.

[3] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock. Finding deceptive opinion spam by any stretch of the imagination. In ACL, 2011.

[4] Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 2014.

[5] L. Akoglu, R. Chandy, and C. Faloutsos. Opinion fraud detection in online reviews bynetwork effects. In ICWSM, 2013.

[6] 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.

[7] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013.

[8] A. Heydari, M. A. Tavakoli, N. Salim, and Z. Heydari. Detection of review spam: A survey. Expert Systems with Applicants, Elsevier, 2014.

[9] M. Crawford, T. D. Khoshgoftar, J. N. Prusa, A. Al. Ritcher, and H. Najada. Survey of Review Spam Detection Using Machine Learning Techniques. Journal of Big Data. 2015.

[10] H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.

[11] A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, M. Castellanos, and R. Ghosh. Spotting opinion spammers using behavioral footprints. In ACM KDD, 2013.

[12] R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networks and metadata. In ACM KDD, 2015.

[13] Y.Sun and J. Han. Mining Heterogeneous Information Networks; Principles and Methodologies, In ICCCE, 2012.

[14] S. Feng, L. Xing, A. Gogar, and Y. Choi. Distributional footprints of deceptive product reviews. In ICWSM, 2012.

[15] Y. Sun, J. Han, X. Yan, P. S. Yu, and T. Wu. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks. In VLDB, 2011.