

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

## Spam Review Detection to Increase User Reliability on Reviews for Availing Online Service

Piyusha V. Kadam, Prasad S. Halgaonkar

Department of Computer Engineering ZCOER, Pune, Maharashtra, India

**ABSTRACT:** People refer reviews while purchasing product or availing any service from online sites. Spammer leave fake review to promote/harm business demand. Recognizing these spammers and the spam content is a widely debated issue of research. Research is going on to identify feature and increase the accuracy of spam review detection. Various features are available depending on performance features need to selected and used for fake review detection. The existing system proposed Netspam algorithm made use of review-based and user-based features for spam detection and also calculated the weight of each feature to identify performance of individual features. The Proposed System used Sentiment analysis ,Semantic analysis with Netspam Framework to identify fake reviews with increased accuracy. The Proposed System also provide top-k hotels based on user search query and recommend single hotel based on user point of interest.

**KEYWORDS:** Social Media, Social Network, Spammer, Spam Review, Fake Review, Heterogeneous Information Networks, Semantic analysis, Sentiment Analysis.

### I. INTRODUCTION

Online shopping or availing service is in big trend where user mostly refer reviews, rating and features of product/service before purchasing or availing any service, but if reviews are fake it may mislead the user interest and indirectly harm the business demand. Spammer may promote/damage the reputation of business based on contract. The fact that anyone with any identity can leave comments as review provides a tempting opportunity for spammers to write fake reviews designed to mislead users opinion. Reviews are mostly referred and the reliability of user on reviews need to be maintained to increase online availing service . Various analysis are performed on spam review detection to increase the accuracy. Semantic analysis and sentiment analysis are most commonly used for text analysis.Semantic analysis compare one review with remaining other reviews an generate score ,Sentiment analysis analyze each sentence and remark review has Positive, Negative or Neutral .Research is going on to identify features to increase the accuracy.metadata which mean data about data like rating, time, date and content of review are used for analysis. There are also sudden increase of reviews in time periods for a product such areas are called as review bursts. A large number of customers may purchase the product and write reviews for the product in a short period of time. Most reviewers in this kind of bursts are likely to be non-spammers. when a product is under spam attack, a number of spam or fake reviews may be posted.

#### Motivation:

- To identify the spam user using positive and negative reviews in social media.
- To display only trusted reviews to the users.
- User search query (location), it will show top hotels and recommends one hotel by using users point of interest.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

## II. REVIEW OF LITERATURE

Netspam is framework used for detecting fake reviews Saeedreza Shehnepoor [1]used weight calculation to find the best feature among mentioned used for identifying the fake reviews in metapath, but didn't remove the fake reviews.Ch. Xu [2] represented robust pairwise features model along with FraudInformer framework for detecting the fraud groups involved in creating fake reviews but is difficult to automate.

G. Fei [3] found reviewers in different time period and their co-occurrence in particular time period using Markov Random Field and Loopy belief Propagation methods using graph which result in High accuracy., this proposed method is effective to detect review spammers in review bursts and detect spammers automatically. A. j. Minnich [4] developed challenges in decision making, used 142 features for detecting the fraud behavior on multiple site reviews by cross-comparing their reviews. It is difficult to automate. There are various technique to distinguish abnormal behavior from normal. B. Viswanath [5] described technique for detecting unsupervised behavior from normal they detected abnormal behavior on facebook by considering the abnormal likes of ads on the facebook, where attacker try to weak the expenditure of some advertiser also there are model proposed for indentifying the spam review which are language independent like H. Li

[6] proposed model which used Collective Positive unlabeled learning feature which are language independent for spam re- view identification. Using Two different method M. Crawford [7]proposed system for fake review detection they used two feature filter-based and word frequency-based feature but has a drawback where proposed system does not fulfill all approach of spam detection. Fake review can also be indentified by considering the relationship between the user and spammer put forth by H. Xue [8] where effective methods assuming that users mostly believe on reviews written by trust-worthy peoples they are mostly connected with an on other side spammers rarely are connected with normal user.The proposed model overcome the sparsity problem and compute the overall trustworthiness score for every user in the system, which is used as the spamicity indicator.Using limited information and novel consistency model S. Mukherjee [9] overcome the limitation of fine-grained information which is not applicable for long tail items by using consistency feature. Pattern to increase the speed of query online processing and find the similarity between objects is introduced by Y. Sun[ [10] with his colleagues. any of the three feature score between 0.7 to 1 then the review will be considered as spam else Nonspam. Review being detected as Spam will be removed and only trusty reviews will be displayed. Based on user point of interest recommendation of top-k hotel will be displayed to the user. **Advantages:**

- To identify fake reviews.
- Display only trusted reviews to the user.
- Increase user trust on business and product.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

## III. EXISTING SYSTEM

Nowadays people mostly refer written reviews for making correct decision and selecting the best hotels and services among multiple. Social media portals play the leading role in information generation an act as important source for producers for promoting an well as for customers for selecting hotels and services. Thus review play a vital role in success of a business, where positive review can benefit company negative review can cause economic losses. The reviews written to change users perception of how good a product or a service are considered as spam, and are often written in exchange for money. which need to be identified and only trusted reviews to be displayed to the user. The Existing System introduce novel spam detection framework which calculates the importance of each feature using weight calculation method used for spam review detection and defined Review-Behavioral based features, Review-Linguistic based features, User-Behavioral based features and User-Linguistic based features used for spam review detection. The System used real-world datasets from yelp and Amazon websites.

### Disadvantages:

- There is no information filtering concept in online social network.
- Less accuracy.
- More time complexity

## IV. SYSTEM ARCHITECTURE

Text Analysis algorithm like Sentiment analysis and Semantic analysis are most commonly used. Netspam framework along with Semantic analysis and Semantic analysis is applied on Dataset. As shown in Fig. 1 Reviews data is extracted from site using crawler, Data extracted contain time, rating, date, user identity and content of individual review on which 1) Netspam framework features like Review-Behavioral based features, Review-Linguistic based features, User-Behavioral based features and User-Linguistic based features are applied, metadata which mean data about data is used in behavior feature analysis for spam review detection and Content of review are used in Linguistic feature analysis. 2) Sentiment analysis is performed 3) Semantic analysis is applied. Score is generated by each feature in the range  $[0,1]$ . Result yield in range of  $[0,1]$  if

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

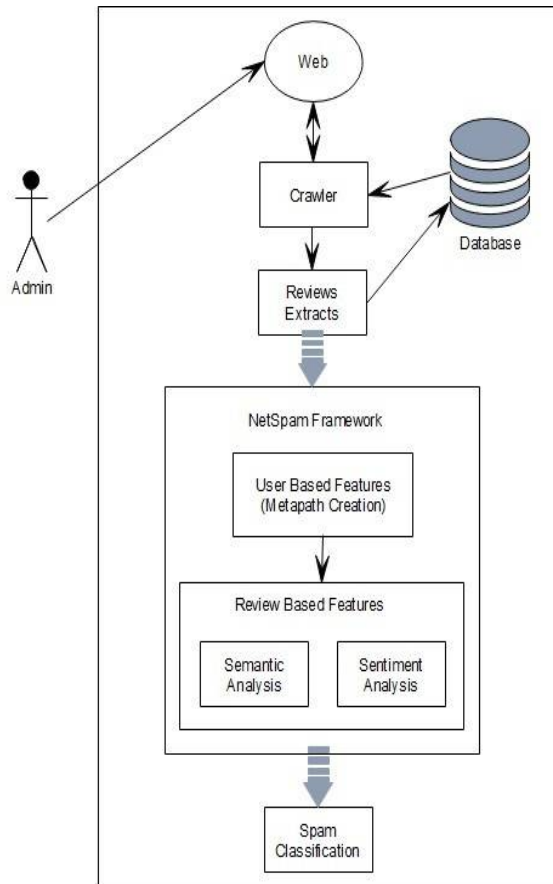


Fig. 1. Proposed System Architecture

## V. MATHEMATICAL MODEL

### A. Spam Features:

#### (1) User-Behavioral (UB) based features:

Burstiness: Spammers, usually write their spam reviews in short period of time for two reasons: first, because they want to impact readers and other users, and second because they are temporal users, they have to write as much as reviews they can in short time.

$$x_{BST}(i) = \begin{cases} 0 & (L_i - F_i) \notin (0, \tau) \\ 1 - \frac{L_t - F_t}{\tau} & (L_i - F_i) \in (0, \tau) \end{cases} \quad (1)$$

Where,

$L_i - F_i$  describes days between last and firstreview for  $\tau = 28$ . Users with calculated value greater than 0.5 take value 1 and others take 0.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

## (2) User-Linguistic (UL) based features:

Average Content Similarity, Maximum Average Content Similarity, Maximum Content Similarity: Spammers, often write their reviews with same template and they prefer not to waste their time to write an original review. In result, they have similar reviews. Users have close calculated values take same values (in [0; 1]).

## (3) Review-Behavioral (RB) based features:

Rate Deviation using threshold:

Spammers, also tend to promote businesses they have contract with, so they rate these businesses with high scores. In result, there is high diversity in their given scores to different businesses which is the reason they have high variance and deviation.

$$x_{DEV}(i) = \begin{cases} 0 & \text{Otherwise} \\ 1 - \frac{r_{ij} - \text{avg}_{e \in E+j} r(e)}{4} > \beta_1 & \end{cases} \quad (2)$$

Where,

$\beta_1$  is some threshold determined by recursive minimal entropy partitioning. Reviews are close to each other based on their calculated value, takesame values (in [0; 1]).

## (4) Review-Linguistic (RL) based features:

Number of first Person Pronouns, Ratio of Exclamation Sentences containing '!': First, studies show that spammers use second personal pronouns much more than first personal pronouns. In addition, spammers put '!' in their sentences as much as they can to increase impression on users and highlight their reviews among other ones. Reviews are close to each other based on their calculated value, take same values (in [0; 1]).

---

## B. NetSpam Algorithm

---

Input: review\_dataset, spam\_feature\_list, pre\_labeled\_reviews

Output: features\_importance (W), spamicity\_probability(Pr)

u, v: review,  $y_u$ : spamicity probability of review u

$f(x_{lu})$ : initial probability of review u being spam

$P_l$  metapath based on feature l, L: features number

n: number of reviews connected to a review

$m_u^{Pl}$ : the level of spam certainty

$m_{u,v}^{Pl}$ : the metapath value

Prior Knowledge

**if** semi-supervised mode

**if**  $u \in \text{pre\_labeled\_reviews}$

$y_u = \text{label}(u)$

**else**

$y_u = 0$

**else** unsupervised mode

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

$$y_u = \frac{1}{L} \sum_{l=1}^L f(x_{lu})$$

Network Schema Definition

schema = defining schema based on spam-feature-list

Metapath Definition and Creation

for  $P_l \in schema$

for  $u, v \in review\_dataset$

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

$$m_v^{Pl} = \frac{|s \times f(x_{lv})|}{s}$$

if  $m_u^{Pl} = m_v^{Pl}$

$$m_{u,v}^{Pl} = m_u^{Pl}$$

else

$$m_{u,v}^{Pl} = 0$$

Classification - Weight Calculation

for  $Pl \in schemes$

$$do W_{Pl} = \frac{\sum_{r=1}^n \sum_{s=1}^n m_{r,s}^{Pl} \times y_r \times y_s}{\sum_{r=1}^n \sum_{s=1}^n m_{r,s}^{Pl}}$$

Classification - Labeling

for  $u, v \in review\_dataset$

$$Pr_{u,v} = 1 - \prod_{Pl=1}^L 1 - m_{u,v}^{Pl} \times W_{Pl}$$

$$Pr_u = avg(Pr_{u,1}, Pr_{u,2}, \dots, Pr_{u,n})$$

return (W, Pr)

---

## C. Latent Semantic Analysis Algorithm

---

Step 1: Documents should be prepared in the following way:

- 1) Exclude trivial words as well as low-frequency terms
- 2) Conflate terms with techniques like stemming or Lemmatization.

Step 2: A term-frequency matrix (A) must be created that includes the occurrences of each term in each document.

Step 3: Singular Value Decomposition (SVD):

- 1) Extract least-square principal components for two

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

- sets of variables: set of terms and set of documents.  
2) SVD products include the term eigenvectors U, the document eigenvectors V, and the diagonal matrix of singular values  $\Sigma$ .

Step 4: From these, factor loading can be produced for terms  $\Sigma U$  and documents  $\Sigma V$

### B. Sentiment Analysis Algorithm:

Input: Text File(comment or review) T, The sentiment lexicon L.

Output:  $S_{mt} = P, Ng$  and N and strength S where P :

Positive, Ng : Negative, N : Neutral Initialization : SumPos = SumNeg = 0, where,

SumPos : accumulates the polarity of positive tokens  $t_i \in \text{smt}$  in T,

-SumNeg : accumulates the polarity of negative tokens  $t_i \in \text{smt}$  in T

**begin**

for each  $t_i \in T$  do Search for  $t_i$  in L

if  $t_i \in \text{Pos-list}$  then SumPos  $\leftarrow$  SumPos +  $t_i$  - smt

else if  $t_i \in \text{Neg list}$  then

SumNeg  $\leftarrow$  SumNeg +  $t_i$  - smt

end if

end for

if SumPos > |SumNeg| then

Smt = P

$$S = \frac{\text{SumPos}}{(\text{SumPos} + \text{SumNeg})}$$

else if SumPos < |SumNeg| then

Smt = Ng

$$S = \frac{\text{SumNeg}}{(\text{SumPos} + \text{SumNeg})}$$

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

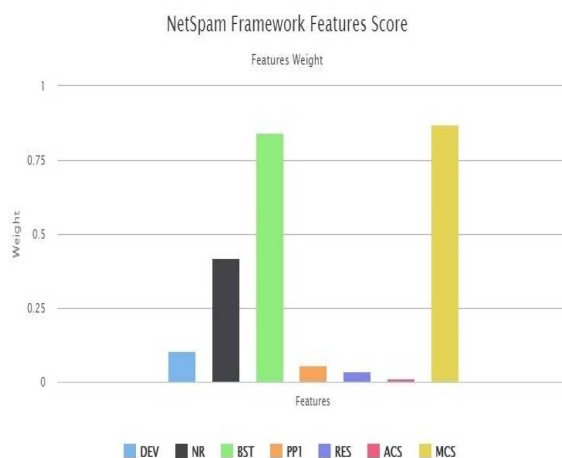


Fig. 2. Feature weights for NetSpam Framework

Fig.3 graph shows the total 435 reviews of single hotel from TripAdvisor website classified that 160 reviews are spam and 275 reviews are non-spam by using NetSpam framework.

```

else
Smt = N
S = (SumPos / (SumPos + SumNeg))
end if
end

```

## V. SYSTEM ANALYSIS

The Proposed system contain Tomcat web server , Java (jdk 1.8) and Mysql Connectivity. The Real-world dataset is used from Tripadvisor site using Crawler on which Sentiment analysis, Semantic analysis and Netspam algorithms are applied. Fig.2 graph show weight of all features calculated using Netspam Framework. Table-I contain Result set of all features.

Table I : Weights of all features

Features	Weight
DEV	0.1034
NR	0.4207
BST	0.8414
RES	0.0345
PPI	0.0575
ACS	0.0115
MCS	0.8713



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

Table II : Classification results of NetSpam Framework for hotel reviews

Review	Count
Spam	160
Non-Spam	275

The Proposed netspam framework time complexity is  $O(e^2n)$ . The netspam framework accuracy is 94.06% which is better than SPaglePlus Algorithm accuracy is 85.14% on using tripadvisor hotel dataset.

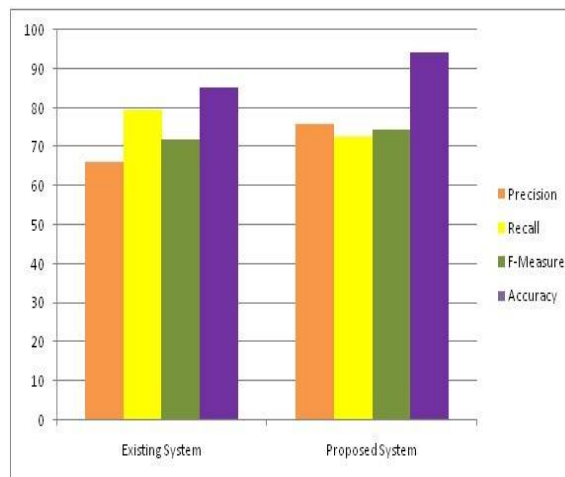


Fig. 4. Performance Analysis between existing and proposed system

## VI. CONCLUSION AND FUTURE WORK

The Netspam framework along with sentiment analysis and semantic analysis increase the accuracy of spam review detection. The performance of the proposed system is evaluated by using dataset of Tripadvisor. Metadata which mean data about data is used in behavior feature analysis for spam review detection and Content of review are used in Linguistic feature analysis. Weight of each Netspam feature is calculated using weight calculation method to analyze the performance of each feature result for better combination of features in spam detection. The system will display only trusted reviews to the user considering the score calculated and will recommend one hotel among top-k hotel based on user point of interest.

Using different combination depending on performance of each feature the accuracy of spam detection can be increased.

## VII. ACKNOWLEDGEMENTS

We would like to thank Trip Advisor for allowing us to crawl all their reviews.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 6, June 2018

## REFERENCES

- [1] SaeedrezaShehnepoor, Mostafa Salehi\*, Reza Farahbakhsh, Noel Crespi. NetSpam: a Network-based Spam Detection Framework for Reviews in Online Social Media. In IEEE TIFS 2016.
- [2] Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 2014.
- [3] G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013
- [4] A. j. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. Trueview: Harnessing the power of multiple review sites. In ACM WWW, 2015.
- [5] B. Viswanath, M. Ahmad Bashir, M. Crovella, S. Guah,
- [6] K. P. Gummadi, B. Krishnamurthy, and A. Mislove. Towards detecting anomalous user behavior in online social networks. In USENIX, 2014.
- [7] H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.
- [8] M. Crawford, T. M. Khoshgoftaar, and J. D. Prusa. Reducing Feature set Explosion to Facilitate Real-World Review Spam Detection. In Proceeding of 29th International Florida Artificial Intelligence Research Society Conference. 2016.
- [9] H. Xue, F. Li, H. Seo, and R. Pluretti. Trust-Aware Review Spam Detection. IEEE Trustcom/ISPA . 2015.
- [10] 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.
- [11] 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.
- [12] C. Luo, R. Guan, Z. Wang, and C. Lin. HetPathMine: A Novel Transductive Classification Algorithm on Heterogeneous Information Networks. In ECIR, 2014.
- [13] C. L. Lai, K. Q. Xu, R. Lau, Y. Li, and L. Jing. Toward a Language Modeling Approach for Consumer Review Spam Detection. In Proceedings of the 7th international conference on e-Business Engineering. 2011.
- [14] S. Feng, R. Banerjee and Y. Choi. Syntactic stylometry for deception detection. Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers; ACL, 2012.
- [15] Y. Sun and J. Han. Rankclus: integrating clustering with ranking for heterogeneous information network analysis. In Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology, 2009.
- [16] 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.