SPAM REVIEW DETECTION USING LINGUISTIC AND NETSPAM FRAMEWORK

Shrikant Bhosale[1], Swapnil Gaikwad[2], Vishwajeet Dongare[3], Sushant Bhingare[4]

1,2,3,4Department of Computer Engineering, JSPM's Imperial College of Engineering & Research, Wagholi, Pune.

---***---

Abstract - Reviews are play an influential role in advertising campaigns as well as for customers in selecting products and services. In the past years, people rely a lot on the written reviews in their decision-making processes, and positive/negative reviews encouraging/discouraging them in their selection of products and services. Spam reviews are written to promote/demote targeted products or services, this practice is known as review spamming. By using spam review detection system users will stay away from fake/spam reviews.

Key Words: Spam, Online Reviews, Fake Reviews, Spam Detection, Fake Review Detection.

1. INTRODUCTION

Nowadays, the World Wide Web (WWW) has become a major source of self-expression. People can easily share their ideas on any product or service through commerce sites, forums and blogs. Everyone on the web now acknowledges the value of these online updates for both customers and sellers. Most people read reviews about products and services before buying them. Vendors can also design their own future production or marketing strategies based on these updates. For example, if various customers purchase a specific model of laptop, send updates on problems related to its screen design, the manufacturer can also solve the problem to increase customer satisfaction. Recently, the trend of spam attacks has increased because anyone can simply write a review of spam and post it online without hindrance. Anyone can hire people to write non-specific updates for their products and services, such people are called spam.

This paper describes two methods of checking a review is spam or not. Following are these two methods:

1) Spam Test Detection using the Linguistic Method (SRD-LM) uses linguistic model of machine learning and calculate the spam review used to identify spam and spam reviews.

2) NetSpam Framework (SRD-NS) applies to content and uses modification, feature selection and editing to receive spam updates.

Specifically, SRD-NS achieved 91.1% accuracy while SRD-LM achieved 86.2% accuracy in identifying spam reviews.

2. LITERATURE SURVEY

Existing research has explored various ways to get spam reviews to get spam reviews. This study reviewed the literature in two ways: (1) SRD using the language method and (2) SRD using the NetSpam framework. The aim is to determine new contributions for the proposed project by comparing it with previous studies.

(A) SPAM REVIEW DETECTION USING LINGUISTIC METHOD (SRD-LM):

According to the literature review, it has been observed that most of the existing studies did not include many important linguistic features while designing SRD models based on languages and used a single separator to train their proposed models. The current study, therefore, extends the SRD domain to design a language model using several features, including stem and N-gram techniques. In addition, the proposed model uses and compares the accuracy of four different categories, including Naïve Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM) and Random Forest (RF) to advance and improve. accurate prediction of spam reviews.

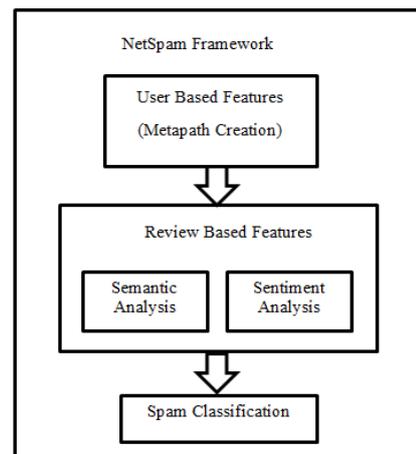**(B) SPAM REVIEW DETECTION USING NETSPAM FRAMEWORK (SRD-NS):**

Fig -1: NetSpam Framework

A large number of literature has been published on the techniques used to identify spam and a different type of analysis on this topic. These methods can be divided into different categories; some use language patterns in text, mainly based on bigram, and unigram, others are based on behavioral patterns based on patterns derived from patterns in user behavior based primarily on metadata, and other graphs and algorithms based on graphs and classifiers. Despite such great efforts, many factors have been missed or remain unresolved. One of them is a classier capable of calculating feature elements that indicate the level of each element of importance in determining spam reviews. A common idea of NetSpam framework is to model the review data provided by the Heterogeneous Information Network (HIN) and to include the detection of spam in the HIN filtering problem. In particular, we emulate database updates such as HIN in

updates linked to various node types (such as features and users).

3. PROPOSED SYSTEM

The proposed system is made up of using Linguistic and NetSpam Framework. We calculate the average spam score given by these two models and shows it to user.
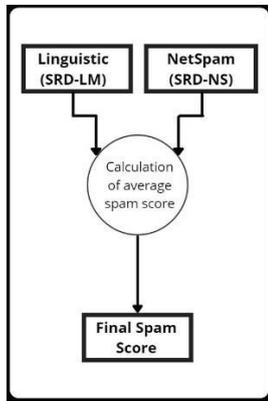

Fig -2: Obtaining final spam score

There are two parts of our proposed system- User Side and Admin Side. Let's discuss about these parts in brief.

(A) User Side:

- User first register account in web application.
- After activate account, user should login.
- User search query.
- Display top list of filtered results (products) and after get one recommendation.
- User may submit review.
- View store's offline location.

(B) Admin Side:

- Admin login to system.
- Admin authenticate the user.
- Create meta-path of user and sentiment analysis on reviews.
- Classify spammer and spam reviews.
- Admin publish/deactivate users' reviews.
- Display Offline shop's location.

The accuracy derived using Linguistic and NetSpam Framework is 92.1%. Our proposed system also provide location of local/offline store where product is available for checking or buying.

4. CONCLUSIONS

Upon all the research & observations above, we move towards following conclusions.

1. Based on our observations, defining two views for features (user-netspam and behavioral-linguistic), the classified features as netspam have more weights and yield better performance on spotting spam reviews in both semi-supervised and unsupervised approaches.
2. Categorizing features in four major categories (review-netspam, user-netspam, review-linguistic, user-linguistic), helps us to understand how much each category of features is contributed to spam detection.
3. Using both Linguistic Method and NetSpam Framework, we achieved 92.1% accuracy in identifying spam reviews.

5. REFERENCES

1. Ch. Xu and J. Zhang. Combating product review spam campaigns via multiple heterogeneous pairwise features. In SIAM International Conference on Data Mining, 2014.
2. G. Fei, A. Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh. Exploiting burstiness in reviews for review spammer detection. In ICWSM, 2013.
3. A. J. Minnich, N. Chavoshi, A. Mueen, S. Luan, and M. Faloutsos. Trueview: Harnessing the power of multiple review sites. In ACM WWW, 2015.
4. 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.
5. H. Li, Z. Chen, B. Liu, X. Wei, and J. Shao. Spotting fake reviews via collective PU learning. In ICDM, 2014.
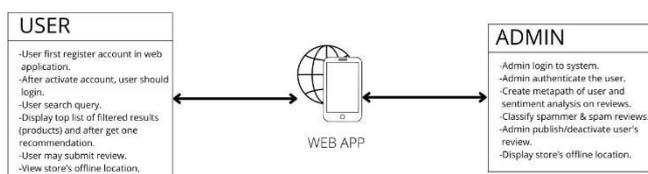
Fig -3: Parts of proposed system