Spam Review Detection and Recommendation of Superior Results in NetSpam Framework

Chaitrali kardile
ME Student, Department of Computer Engineering, modern education society's college of engineering, Pune

Prof. Revati M. Wahul
Assistant professor, Department of Computer Engineering, modern education society's college of engineering, Pune

Abstract*: *Use of social media, this marketing strategy can be used to destroy financial state of the product or service provider. Because of customer's trust on social media there are always positive and negative reviews or opinion about the product which affects a service provider. It has a liability that anyone can leave a review and spammers can take off a chance to take sincere survey about a product. There has been considerable amount of studies to categorize this spam reviews as in positive and negative. This paper proposes a use of heterogeneous information network in Netspam framework. Heterogeneous information network contribute in large number of data which will be useful in developing spam detection system. This system will categorize spam reviews on the basis of features that are review on behavioural based, user on behavioural based, review on linguistic based, user on linguistic based, the first type of features performs better than the other categories. The contribution work is when user will search query it will display all top products as well as there is recommendation of the product by using user's point of interest..

Index Terms*: *Social Media, Social Network, Spammer, Spam Review, Fake Review, Heterogeneous Information Networks, Metapath.

I. INTRODUCTION

In today's social activities people always believe in what other people has to say about their interested topic. And this is the reason why the opinion of customer matters while buying products or service providers. Now a day it has become important to take customers opinion about the product. Abundant use of social media has become one of a marketing strategy in past few years and this accommodates the secure financial state of the company. Most of the customers take written reviews in their account while making decision about the product. So that is why written review has become a important parameter in business strategies. Positive reviews can develop good business for product and on the other hand negative reviews can cause economic losses to product and service provider. Using social media anyone can make a fake identity an mislead the survey or the opinion of the customer and when this written review gets shared on social media it leads to wrong image of product and following that economic loss for a company. The reviews which are written to change user's or customer's perception regarding a product or a service provider are classified as a spam which are often written in exchange for money.

1. Background

There is considerable amount of literature available with the different type techniques for spam detection. These techniques classified into linguistic and behavioral. In linguistic pattern in text relay on a unigram and bigram which generate weight for a review and in behavioral pattern it relay on feature extraction using pattern in user's behavior. There are still many techniques and aspect needs more study for spam detection. One of them is calculate importance of features using feature weight in spam detection system. Our spam detection system will model a different dataset using heterogeneous information network and will help to map spam in detection system into classification. In heterogeneous information network, it works with different nodes which are interlinked with each other. To find importance level of each feature our system will use algorithm of weighting. To estimate a solution we will use this weighting which will help to find final labels of the written reviews. Our system will use Yelp and Amazon dataset which will provide linguistic and behavioral views of features. In classified features review behavioral has more weights and can yield better performance in both supervised and unsupervised approaches for detection system.

2. Motivation

In today's social activities, people are always convinced of what others have to say about their subject. And that's why customer opinion matters when you buy products or service providers. Now it is important to take the opinion of the customers about the product. In recent years, the widespread use of social media has become one of the marketing strategies and adapts to the company's secure financial state. This marketing strategy can be used to destroy the financial condition of the product or service provider using the social media. Due to the customer's confidence in social media, positive and negative reviews or opinions about the product affecting a service provider are always available. It is the responsibility of everyone to leave a review and spammers can take the opportunity to conduct a sincere product survey.

II. LITERATURE SURVEY

The paper [1] represents the pair wise features are first explicitly utilized to detect group colluders in online product review spam campaigns, which can reveal collusions in spam campaigns from a more fine-grained perspective. A novel detecting framework named Fraud Informer is proposed to

cooperate with the pair wise features which are intuitive and unsupervised. Advantages are: Pairwise features can be more robust model for correlating colluders. Manipulate perceived reputations of the targets for his or her best interests.
To rank all the reviewers in the website globally so that top-ranked ones are more likely to be colluders. Disadvantages are: Difficult problem to automate.

The paper [2] builds a network of reviewers appearing in different bursts and model reviewers and their co-occurrence in bursts as a Markov Random Field (MRF), and employ the Loopy Belief Propagation (LBP) methodology to infer whether or not a reviewer could be a transmitter or not within the graph. A novel analysis methodology to judge the detected spammers mechanically victimization supervised classifica tion of their reviews. Advantages are: High accuracy. The proposed method is effective. To detect review spammers in review bursts. Detect spammers automatically. Disadvantages are: a generic framework is not used for detect spammers.

In paper [3], the challenges are: The detection of fraudulent behaviors, assessing the trustworthiness of review sites, since some may have policies that enable misbehavior, and creating effective review aggregation solutions. The TrueView score, in three different variants, as a proof of concept that the synthesis of multi-site reviews can provide important and usable information to the end user. Advantages are: Develop novel features capable of identifying cross-site discrepancies effectively. A hotel identity-matching method has 93% accuracy. Enable the site owner to detect misbehaving hotels. Enable the end user to trusted reviews. Disadvantages are: Difficult problem to automate.

The paper [4] describes unsupervised anomaly detection techniques over user behavior to distinguish potentially bad behavior from normal behavior. To detect diverse attacker strategies fake, compromised, and colluding Facebook identities with no a priori labeling while maintaining low false positive rates. Advantages are: Anomaly detection technique to effectively identify anomalous likes on Facebook ads. Achieves a detection rate of over 66% (covering more than 94% of misbehavior) with less than 0.3% false positives. Disadvantages are: The attacker is trying to drain the budget of some advertiser by clicking on ads of that advertiser.

In [5] paper, a collective classification algorithm called Multi-typed Heterogeneous Collective Classification (MHCC) and then extends it to Collective Positive and Unlabeled learning (CPU). The proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings. Advantages are: Proposed models can markedly improve the F1 scores of strong baselines in both PU and non-PU learning settings. Models just utilize language independent features; they can be effectively summed up to different languages. Detect a large number of potential fake reviews hidden in the unlabeled set. Disadvantages are: Fake reviews hiding in the unlabelled reviews that Dianping's algorithm did not capture. The specially appointed names of clients and IPs utilized as a part of MHCC may not be exceptionally exact as they are figured from names of neighboring audits.

The paper [6] elaborates two distinct methods of reducing feature subset size in the review spam domain. The ways embrace filter-based feature rankers and word frequency primarily based feature choice. Advantages are: The first method is to simply select the words which appear most often in the text. Second method can use filter based feature rankers (i.e. Chi-Squared) to rank options so choose the highest stratified options. Disadvantages are: There is not a one size fits all approach that is always better.

In [7] paper, providing an efficient and effective method to identify review spammers by incorporating social relations based on two assumptions that people are more likely to consider reviews from those connected with them as trustworthy, and review spammers area unit less probably to take care of an oversized relationship network with traditional users. Advantages are: The proposed trust-based prediction achieves a higher accuracy than standard CF method. To overcome the scarcity problem and compute the overall trustworthiness score for every user in the system, which is used as the spam city indicator. Disadvantages are: Review dataset required.

III. EXISTING SYSTEM APPROACH

In today's social activities people always believe in what other people has to say about their interested topic. And this is the reason why the opinion of customer matters while buying products or service providers. Now a day it has become important to take customers opinion about the product. Abundant use of social media has become one of a marketing strategy in past few years and this accommodates the secure financial state of the company. Most of the customers take written reviews in their account while making decision about the product. So that is why written review has become a important parameter in business strategies. Positive reviews can develop good business for product and on the other hand negative reviews can cause economic losses to product and service provider. Using social media anyone can make a fake identity and mislead the survey or the opinion of the customer and when this written review get shared on social media it leads to wrong image of product and following that economic loss for a company. There are some disadvantages in existing system one o them is that there is no concept for filtering a information on social media. And anyone can create fake account or registration and leave a comment or review. Because of this issues system has less accuracy regarding information filtration also has more time complexity.

IV. PROPOSED SYSTEM ARCHITECTURE

The proposed system will model the data set for the review as a heterogeneous information network and map the spam detection problem into a HIN classification issue. In particular, we model the data set for review as a HIN in which reviews are connected through different node types(such as features and use). A weighting algorithm is then used to calculate the importance of each feature(or weight). These weights are used to calculate the final labels for examinations using both unattended and monitored approaches. Based on our observations, which define two views for features(review- user and behavioral- linguistic), the classified features as review behavioral have more weights and better spotting performance in both semi- supervised and unsupervised approaches The feature weights can be added or removed for labeling and therefore the time complexity for a specific level of accuracy can be scaled. Categorizing features in four main categories(review- behavioral, user- behavioral, language- review, user- language) helps us to understand how much spam detection is helped by each category of features.

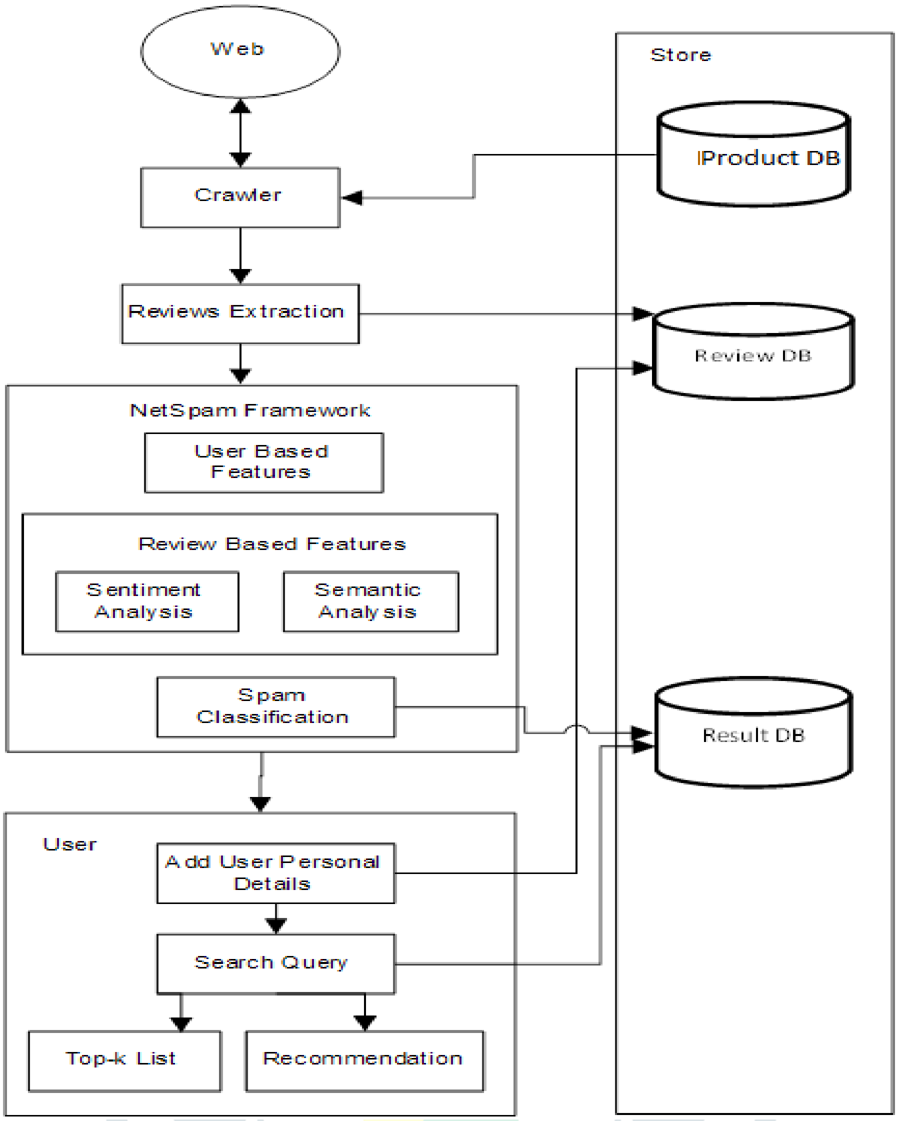

Fig 1: Proposed System Architecture

V. MATHEMATICAL MODEL

Spam Features:

User-Behavioural (UB) based features:

Burstiness: to impact readers and other users spammers have to write as much as reviews they can in short time.

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

Where,
$L_i - F_i$ Describes days between last and first review for $\tau = 28$.
Users with calculated value greater than 0.5 take value 1 and others take 0.

User-Linguistic (UL) based features:

Average Content Similarity, Maximum Content Similarity: spammers usually have templates for reviews so that they won't waste any time and it leads to content similarity. Users have close calculated values take same values (in [0; 1]).

Review-Behavioural (RB) based features:

Early Time Frame: spammers write their review in short period of time so that users can see their review as soon as possible.

$$x_{ETF}(i) = \{0 \quad (L_i - F_i) \notin (0,\delta)\ 1 - \frac{L_t - F_t}{\delta} \quad (L_i - F_i) \in (0,\delta) \qquad (2)$$

Where,
$L_i - F_i$ denotes days specified written review and first written review for a specific business. We have also δ = 7. Users with calculated value greater than 0.5 takes value 1 and others take 0.

Rate Deviation using threshold: Spammers give high score to the business or a product they are in contract with so that particular business get profits. In result, because of high diversity in different business leads to high rate deviation.

$$x_{DEV}(i) = \{0 \qquad Otherwise\ 1 - \frac{r_{ij} - avg_{e \in E*j} r(e)}{4} > \beta_1 \qquad (3)$$

Where,

β1 is some threshold determined by recursive minimal entropy partitioning. Reviews are close to each other based on their calculated value, take same values (in [0; 1)).

Review-Linguistic (RL) based features:

Number of first Person Pronouns, Ratio of Exclamation Sentences containing '!': studies shows that spammers mostly uses a second personal pronoun than first. And to get more users attention they add multiple exclamation marks in review and it increases a impression on users that product or service provider is better than others. Reviews are close to each other based on their calculated value, take same values (in [0; 1]).

VI. CONCLUSION

We use novel spam detection framework named NetSpam based on a metapath creation as well as new graph-based method for labeling reviews relying on a rank-based labeling approach. The calculated weights by utilizing this metapath concept can be very impressive in identifying spam reviews and spammers leads to a better performance. In extension, we found that even without a train set, NetSpam can calculate the consequence of each feature and it yields better performance in the features' addition process, and performs better than existing works, with only a small number of features. Moreover, after defining four main categories for features our conclusions show that the reviews behavioral category performs better than other categories.

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.
6. M. Crawford, T. M. Khoshgoftaar, and J. D. Prusa. Reducing Feature set Explosion to Faciliate Real-World Review Spam Detection. In Proceeding of 29th International Florida Artificial Intelligence Research Society Conference. 2016.
7. H. Xue, F. Li, H. Seo, and R. Pluretti. Trust-Aware Review Spam Detection. IEEE Trustcom/ISPA. 2015.
8. E. D. Wahyuni and A. Djunaidy. Fake Review Detection From a Product Review Using Modified Method of Iterative Computation Framework. In Proceeding MATEC Web of Conferences. 2016.
9. R. Hassanzadeh. Anomaly Detection in Online Social Networks: Using Datamining Techniques and Fuzzy Logic. Queensland University of Technology, Nov. 2014.
10. R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networks and metadata. In ACM KDD, 2015.
11. L. Akoglu, R. Chandy, and C. Faloutsos. Opinion fraud detection in online reviews bynetwork effects. In ICWSM, 2013.
12. R. Shebuti and L. Akoglu. Collective opinion spam detection: bridging review networksand metadata. In ACM KDD, 2015.
13. 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.
14. N. Jindal, B. Liu, and E.-P. Lim. Finding unusual review patterns using unexpected rules. In ACM CIKM, 2012.
15. E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw. Detecting product review spammers using rating behaviors. In ACM CIKM, 2010.
16. [16]. 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.
17. S. Xie, G. Wang, S. Lin, and P. S. Yu. Review spam detection via temporal pattern discovery. In ACM KDD, 2012.
18. G. Wang, S. Xie, B. Liu, and P. S. Yu. Review graph based online store review spammer detection. IEEE ICDM, 2011.
19. Y. Sun and J. Han. Mining Heterogeneous Information Networks; Principles and Methodologies, In ICCCE, 2012. [20]. A. Mukerjee, V. Venkataraman, B. Liu, and N. Glance. What Yelp Fake Review Filter Might Be Doing?, In ICWSM, 2013.