Detection of Spammer Groups from Product Reviews

Jatin Tonape, Aniket Kamble, Rajesh Bhad, Omkar Ugale, Ass. Prof.Bhagwan Kurhe

Jatin Tonape, Student, Department of Computer Engineering, SKNSITS Lonavala, Maharashtra, India

Aniket Kamble , Student, Department of Computer Engineering, SKNSITS Lonavala, Maharashtra, India

Rajesh Bhad , Student, Department of Computer Engineering, SKNSITS Lonavala, Maharashtra, India

Omkar Ugale , Student, Department of Computer Engineering, SKNSITS Lonavala, Maharashtra, India

Ass.Prof.Bhagwan Kurhe, Professor, Department of Computer Engineering, SKNSITS Lonavala, Maharashtra, India

---

Abstract - *Digital product reviews are now playing a crucial role in consumer purchasing decisions. A high proportion of positive reviews will bring significant growth in revenue, while negative reviews will result in loss of sales. Driven by the huge financial gain, by posting false and biased online reviews, most spammers try to promote their products or demote the products of their competitors. Through registering accounts numbers or releasing tasks on crowd sourcing sites, most individual spammers could be coordinated as spammer teams to exploit the product reviews together and could be more harmful. Existing work on spammer group identification extracts spammer group candidates from review data and uses unsupervised spam rating methods to classify actual spammer groups. Actually, according to previous research, it is easier than one expects to mark a small number of spammer classes, but few methods try to make good use of these essential marked information. In this task, we propose to detect spammer groups a partially supervised learning system (PSGD). PSGD uses Positive Unlabeled Learning (PU-Learning) to test a classifier as a positive spammer group detector (labeled spammer groups) and unlabeled instances (unlabeled groups) by labeling certain spammer groups as neutral instances. In particular, in terms of positive instances and distinguishing features, we extract consistent negative collection. By combining the positive instances, extracted negative instances, and unlabeled instances, we transform the PU-Learning problem into the well-known semi-supervised learning problem, and then use the Naive Bayesian model and EM algorithm to train a spammer group detection classifier. Amazon.cn real-life experiments show that the proposed PSGD is successful and outperforms state-of - the-art methods for team identification of spammers.*

1.INTRODUCTION

Online product reviews are becoming increasingly important in e-commerce sites, as consumers ' purchasing decisions are strongly influenced by these reviews. Because of financial incentives, most impotents seek to manipulate the processes and customers by publishing skewed ratings and reviews to promote their products or demote their rival products. These imposters, also known as Review Spammers or Opinion Spammers, are becoming more and more harmful as crowd sourcing tasks may coordinate them. Because there are many accounts, with little abnormal behavior, the coordinated spammers, called Spammer Group, may take full control of the feeling on their target items. Although there has been a lot of effort to monitor spam and identify individual spammers, the spammer community identification has received limited attention. Generally speaking, since there are typically no classified instances (groups), most of the existing work will first find spammer group candidates and then use unsupervised ranking methods to distinguish real spammer classes from these candidates. Nevertheless, according to the research in, to obtain

such classified instances (i.e., labeled spammer groups or non-spam groups) we could easily tag certain groups manually. It is apparent that it will greatly improve the accuracy of spammer group identification by combining these classified instances with other unlabeled groups. A typical problem of partially supervised learning is the simultaneous use of marked and unlabeled data. Strictly speaking, according to the constitution of marked data, there are two forms of partially supervised learning.

Another category is that the marked information includes instances of all classes (e.g. this project contains both spammer and non-spammer groups), commonly referred to as semi-supervised learning. The second type is that only positive instances (e.g. spammer groups) are found in the marked information and we need to learn from positive and unlabeled instances. In this venture, we call Positive Unlabeled Learning the second form of partially supervised learning (PU-Learning for short, where P and U stand for positive and unlabeled instances, respectively). Because labeling spammer groups is much simpler than labeling non-spammer groups, we follow PU-Learning as the key technique for detecting spammer groups without any non-spammer groups being marked. We are proposing a Partially Supervised Learning Group Detection (PSGD) framework in this project. Like most existing methods of spammer group detection, we use frequent item mining (FIM) to extract candidates from the spammer community and then use PU-Learning to identify actual spammer groups from these candidates.

Specifically, we tag as positive instances those spammer groups from the identified group candidates manually. Instead, supervised by these positive instances, an algorithm is programmed to automatically extract consistent negative set (RN) consisting of only non-spammer classes. Combine the positive and negative cases, transform the PU-Learning problem into the well-known semi-supervised learning problem, then use the Naive Bayesian method and the Expectation Maximization (EM) algorithm to train a classifier to identify spammer classes. Our previous work has proposed a semi-supervised method of learning to detect spammer groups using Naive Bayesian model and EM algorithm, involving both marked spammer groups and non-spammer groups.

2. Existing System

E-commerce has become a significant review market. There are sites dedicated to the test. Yelp, to name a few, amazon.com. The problem of fake reviews also came into the picture along with the value of the feedback. In the comments, incredible trends can be noted. (i) In the review material, most reviewers share the common word. (ii) The majority of views are contrary to the government. (iii) The time of publication of the summary is also zero. We all seem sincere individually, but the study makes it clear that they are not. Those comments obviously took full control of the product's emotions. There are many instances in which people post incorrect comments about certain items whether individually or in a group. These people are sometimes paid very well. Certain forms of spam review can be commonly categorized as positive and negative feedback. Positive spam reviews tend to support the business with goods they advertise. This also increases the brand company's profit.

The latter seek to malign the image of the product in order to limit the selling of the item. The result is worse if for a good product a negative spam review is written and for a bad product a positive spam review is written. Both of these negative reviews are detrimental in both respects because they fully capture consumer emotions, impacting

the product's sale. To provide consumers with true business experience, these forms of feedback should be identified and removed.

DISADVANTAGES OF EXISTING SYSTEM

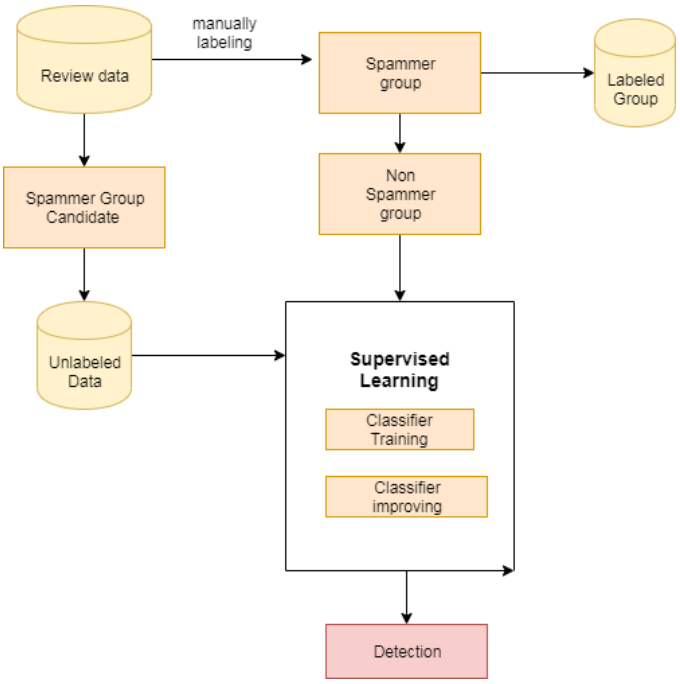
- Imposters try to game systems and customers with skewed ratings and reviews to promote their products or demote the products of their rivals.
- A high proportion of positive reviews will

bring significant growth in revenue, while negative reviews will result in loss of sales.

3. PROPOSED SYSTEM

A model based on spammer group detection (PSGD) based on partially supervised learning. Like most existing methods of spammer group detection, we use frequent item mining (FIM) to extract candidates from the spammer community and then use PU-Learning to identify actual spammer groups from these candidates. Specifically, we mark as positive instances those spammer groups from the identified community candidates manually. Instead, supervised by these positive instances, an algorithm is programmed to automatically extract consistent negative set (RN) consisting of only non-spammer classes. Combine the positive and negative cases, transform the PU-Learning problem into the well-known semi-supervised learning problem, then use the Naive Bayesian model and Expectation Maximization (EM) algorithm to train a classifier to detect spammer classes. Our previous work has proposed a semi-supervised method of learning to identify spammer groups using Naive Bayesian model and EM algorithm, involving both marked spammer groups and non-spammer groups. In comparison to previous work, in this project, we concentrate on detecting only marked spammer classes. Our key contributions are summarized as follows:

1) We are proposing PSGD, a partly supervised method of learning to detect spammer feedback. Specifically, from the marked and unlabeled instances, we just tag certain spammer classes as positive instances and learn a classifier. This is the first time that PU-Learning is applied to spammer group identification to the best of our knowledge.
2) We build a robust negative set (RN) extraction algorithm that defines a characteristic strength function to calculate the discriminative power of group characteristics, and then iteratively extracts instances with high discriminative characteristics from unlabeled instances set to get RN. The PU-Learning problem can be translated into the well-known semi-supervised learning problem by adding the positive instances and the extracted negative instances, so many mature methods like the Naive Bayesian model and the EM algorithm can be applied to construct the classifier.
3) We are performing detailed tests on a real-life data set obtained from Amazon.cn. We are proposing two new group features and checking their effect to improve detection efficiency. In view of PSGD's overall performance, we also evaluate the influence of unlabeled data's weighting factor and assess the effectiveness of our proposed RN extraction algorithm. The experimental results show that PSGD can detect spammer groups effectively and outperform the state-of - the-art methods of group detection.



ADVANTAGES OF PROPOSED SYSTEM

- Highly secured.
- Easy to handle.
- Detect spammer group from product review.

ALGORITHM OF PROPOSED SYSTEM

- Expectation–maximization.
- Naive Bayesian.
- RN extraction algorithm.
- Clustering algorithm

4. LITERATURE SURVEY

[1] Somayeh Shojaee, Masrah Azrifah Azmi Muradt, Azreen Bin Azman, Nurfadhlina Mohd Sharefl and Samaneh Nadali.

Use lexical and syntactic features to spot misleading feedback. Due to the rapid growth of social media users, misleading opinion identification has attracted a lot of interest in science. Given the availability of a large number of features of opinion and techniques of classification, classification of analysis remains a challenging task. We used stylometric features in this study, i.e. lexical and syntactic, using supervised classifiers for machine learning, i.e. Help for misleading opinion by Vector Machine (SVM) with Sequential Minimal Optimization (SMO) and Naive Bayes.

[2]Rinki Patel, Priyanka Thakkar. Opinion Spam Detection Using Feature Selection.

In modern times, ecommerce companies have become very critical in encouraging their end customers to write reviews of the products they have used. These reviews provide vital information sources for these products or services. When deciding to buy new products or services, this data will be used by potential future buyers. Marketers often exploit these thoughts or feedback to find out the drawbacks of their own products or services and, respectively, to find vital information about the products or services of their rival. This in effect allows brand flaws or strengths to be established. Sadly, this significant usefulness of opinions also increased the spam question, which involves fake negative or positive opinions. This paper focuses on identifying spam with false opinion. A recently proposed method of opinion spam detection based on n-gram techniques is expanded through the collection of features and specific representation of opinions.

[3] Gabriel Pui Cheong Fung, Jeffrey Xu Yu, Hongjun Lu, Philip S. Y.

Text Classification without Labeled Negative Documents

This paper presents a new approach to create a text classifier with a small set of marked positive (P) documents and a large set of unlabeled (U) documents. Here, both positive and negative documents were combined with the unlabeled documents. No file, in other words, is marked as negative. This makes it challenging to construct a robust text classifier. Generally speaking, the current approaches to solving this type of problem use a two-step approach: I extracting the negative documents (N) from U; and ii) constructing a P and N classifier. Nonetheless, none of the published studies attempt to extract any meaningful (P) documents from U. Intuitively speaking, Extracting P from U would make the classifier more reliable. It's hard to

remove P from U, though. A U report with some of the features shown in P doesn't necessarily mean it's a good document, and vice versa. Extracting positive documentation is very sensitive because the positive specimens obtained may become noises. Perhaps adding to the challenge of obtaining any relevant documentation is the very large size of U and the very high diversity shown there. We suggest a partition-based heuristic in this paper that aims to extract both positive and negative documents in U. Extensive experiments based on three benchmarks are conducted. The favorable results indicated that our proposed heuristic outperforms all of the existing approaches significantly, especially in the case where the size of P is extremely small. Extensive studies were performed on the basis of three criteria. The favorable results indicated that all existing strategies were significantly outperformed by our proposed heuristic, especially in cases where P is extremely small.

[4] C.L. Lai, K.Q. Xu, Raymond Y.K. Lau

Toward A Language Modeling Approach for Consumer Review Spam Detection

The frequency of fake reviews (i.e., spam) posted on various e-commerce or opinion sharing websites has been mentioned in numerous reports. Nonetheless, due to the lack of an appropriate analytical approach, very few studies have been conducted to evaluate the trustworthiness of online consumer reviews. Unlike other types of Web spam, untruthful reviews might just look like other legitimate reviews (i.e., ham), so it's hard to use any features to distinguish between the two classes. The development of a new mathematical approach to counter spam online review is a major contribution of our research work. Our experimental results confirm the efficacy of the KL divergence and the computational model based on probabilistic language modeling to detect untruthful feedback. Empowered by the proposed computational methods, our empirical study found that spam is about 2% of consumer reviews published on a major e-commerce page.

5. CONCLUSIONS

This project proposes a partially supervised learning based model PSGD to detect spammer groups from product reviews. First, the PSGD model uses frequent item mining (A. Mukherjee, 2013)learning problem, and employ Naive Bayesian model and EM algorithm to construct a classifier as spammer group detector. Experiments on Amazon.cn demonstrate that the proposed PSGD model outperforms both supervised and unsupervised learning methods on spammer group detection. Our future work in the area will focus on the improvementof the PSGD model. Beyond the Naive Bayesian model used in PSGD, we will investigate and incorporate more classification models such as neural network, Semi-Supervised SVM (S3VM) and even ensemble methods. On the positive instances acquisition and RN extraction.

ACKNOWLEDGEMENT

This research was guided by Assistant Proffessor Bhagwan Kurhe. We thank our teaching staff from our college SKNSITS Lonavala who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper.

REFERENCES

[1] Somayeh Shojaee, Masrah Azrifah Azmi Muradt, Azreen Bin Azman, Nurfadhlina Mohd Sharefl and Samaneh Nadali., “Use lexical and syntactic features to spot misleading feedback,”

Published in: 2013 13th International Conference on Intelligent Systems Design and Applications

Date of Conference: 8-10 Dec. 2013

Publisher: IEEE

[2] Rinki Patel and Priyanka Thakkar, "Opinion Spam Detection Using Feature Selection."

Published in: 2014 International Conference on Computational Intelligence and Communication Networks

Date of Conference: 14-16 Nov. 2014

Publisher: IEEE

[3] Gabriel Pui Cheong Fung, Jeffrey Xu Yu, Hongjun Lu and Philip S. Y.

"Text Classification without Labeled Negative Documents,"

Published in: 21st International Conference on Data Engineering (ICDE'05)

Date of Conference: 5-8 April 2005

Publisher: IEEE

[4] C.L. Lai, K.Q. Xu and Raymond Y.K. Lau, "Toward a Language Modeling Approach for Consumer Review Spam Detection,"

Published in: 2010 IEEE 7th International Conference on E-Business Engineering

Date of Conference: 10-12 Nov. 2010

Publisher: IEEE