Detecting Deceptive Reviews Utilizing Review Group Model

Yuejun Li[1,2,3], Fangxin Wang[1,2], Shuwu Zhang[1,2] , Xiaofei Niu[3]
{liyuejun2014@ia.ac.cn, wangfangxin2014@ia.ac.cn, shuwu.zhang@ia.ac.cn, xiaofein@sdjzu.edu.cn}
Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China[1]
University of Chinese Academy of Sciences, Beijing, 100049, China[2]
School of Computer Science and Technology, Shandong Jianzhu University, Jinan, Shandong, 250101, China[3]

Abstract. Online product and store reviews play an important role in product and service recommendation for new customers. However, due to economic or fame reasons, dishonest people are employed to write fake reviews which is also called "opinion spamming" to promote or demote target products and services. Previous research has used text similarity, linguistics, rating patterns, graph relation and other behavior for spammer detection. It is difficult to find fake reviews by a glance of product reviews in time-descending order while It's more easy to identify fraudulent reviews by checking the list of reviews of reviewers. We propose sieries of novel review grouping models to identify both positive and negative deceptive reviews. The review grouping algorithm can effectively split reviews of reviewer into groups which participate in building new model of review spamming detection. Several new features which are language independent based on group model are constructed. Additionally, we explore the collusion behavior between reviewers to build group collusion model. Experiments and evaluations show that the review group method and relevant models can effectivly improve the precision of 4%-7% in deceptive reviews detection task especially those posted by professional review spammers.

Keywords: Deceptive review detection, Opinion spamming, Review group detection, Reviewer collusion.

1 Introduction

People have been active for years in registering in online platform like amazon.com, eBay.com, taobao.com to buy products and many of them are willing to share opinions about the product they buy. Meanwhile, many professional review website emerged to let reviewers to review products or services even when they have not bought or experienced ever. Hence, some dishonest people write fake reviews just for economic reasons or getting credit from the website. Some other people who we called "pseudo professional reviewer" may gather together via offline interaction or instant messaging software to complete a task assigned by store employers.

It's a difficult thing to identify a review to be fake or not just by browse the review list of online product or store because fraudulent reviews are always mixed together when presenting by the website. Through the labeling process, we find that it's easier to identify fake reviews by browsing the review list of a reviewer in time-ascending order. Reviews of reviewers can

MOBIMEDIA 2019, June 29-30, Weihai, People's Republic of China

DOI 10.4108/eai.29-6-2019.2282593

be grouped based on the assumption that reviews in a group are likely to have similar credibility which express the possibility to be fake or not.

We propose an algorithm to segment reviews of reviewers to groups and build a series of models based on group. These group models are used to classify reviews to be fake or not or used to predict the credibility of each review in regression process. Mukherjee [6] has used the concept "group" to spot reviewer group which is different from ours in that we consider the review group but not the reviewer group. Due to the fact that some "pseudo professional reviewers" cooperate with each other to spam the target store or product, we propose collusion detection method and a model to compute the collusion index of each reviewer. The main contribution of the paper includes：

Proposed a review group detection algorithm to split reviews of reviewer into groups utilizing all the reviews of each reviewer.

Built a series of models based on group to classify reviews as fake or not, or to predict the credibility of a review using regression.

Built models which are used for classification or regression based on reviewer collusion behavior.

Experiment and evaluations show that the proposed algorithm and models are effective both in fake review detection and review credibility prediction.

2 Related Work

Researches have been done since Jindal [1] first proposed the concept "opinion spam". On the whole, the review spam (opinion spam) detecting method can be classified to two types: supervised learning and unsupervised methods. Supervised learning method [1, 3, 4] is often used when sufficient labelled data are provided. It has relatively good performance given the right feature and labelled training data. The advantage of the suppervised learning is that it can utilize the latent character of labelled data and make use of apriori knowledge to conduct the classification task. The disadvantage of the method is it requires rather quantity of labelled data which may require more effort of labor in labelling expecially when human labelling is needed. Research in Jindal [1] builds a classifier utilizing duplicate or near-duplicate reviews as fake reviews and the rest as non-fake reviews. The authors of Ott[4] and Law[5] utilize review content part-of-speech(POS) , LIWC text features and language model to find deceptive opinion spam, but did not consider the user behavior which is very useful to detect fake review and reviewer. Unsupervised method has been used to detect group spammers [6, 8] and review burstiness [7] and behavioral footprints [2]. Mukherjee [6] Proposed a method to find candidate reviewer groups and build relation models based on relationships among groups. Our work still uses the concept "group" but considering the relation between reviews instead of the reviewers which is quite distinct from their work. Fei et.al. [7] exploit the burstiness nature of reviews to identify review spammers.

In a wide field, recent studies in opinion spamming are done in social networks [9, 10]. Opinion spamming detection of social networks is quite different from that in product and store review spamming. Lack of user relation links like follow-friend relation makes detection of review spamming a challenging work.

Before we propose the review group detection method, we first introduce the previous work on review and reviewer spamming detection to be as benchmark.

Content similarity based features are commonly used content feature in [1][6][7]. Researchers have been using the n-gram[4, 20, 21, 22] and POS features to model the content of reviews. Similar contents of reviews have bigger probability to be spam. Here we incorporate three different levels of content similarity which include: (a) self-content similarity in review (b) content similarity between pairwise reviews of reviewer and (c) content similarity between reviewer and other reviewers. The three levels of content feature covers inner similarity of a review, inner similarity of a reviewer, and similarity between reviewers about content. Additionally, we add review length as another content based feature.

Behavior features are important indicators of reviewer spamming. Research have been done based on behavior which include rating behavior [2][6], user profile characteristics and so on. Sentiment analysis has also been incorporated in the detection process [16][26]. Review sentiment polarity is computed to be utilized as import feature to identify fake reviews. Comparison of exsited method and our proposed method are listed in table 1.

Table 1. Comparison of exsited method and our proposed method.

Authors	Key Concept	Features	Learners	Result
Jindal and Liu [1]	Text duplication	review, reviewer and product related	Logestic regression	AUC:78%
Ott et al. [4]	Text Categorization	LIWC(Linguistic Inquiry and Word Count), Bigram	SVM, Naïve Bayes	Precision:89%
Mukherjee et al. [2]	Behavioral footprints	Behavioral feature	SVM, clustering	Precision:74.6%
Mukherjee et al. [6]	Spammer group	Spammer group feature	SVM rank	User agreement result
Shojaee et al. [27]	Stylometric	Lexical and Sytactical	SVM	F1:84%
Rout et al. [26]	Content similarity and sentiment polarity	Sentiment score, linguistic features and unigram	SVM Naïve Bayes Decision Tree	Accuracy:88% Accuracy:91% Accuracy:92%
Proposed approach	Review group detection and collusion	Review group and collusion related	SVM Naïve Bayes Random Forest	Precision:92% Precision:87% Precision:96%

Graph based review spammer group detection techniques are studied in recent years. Wang[12] first proposed heterogeneous review graph to capture the relationship among reviewers, reviews and stores that the reviewers have reviewed. Wang Z. [25] exploited the topological structure of the underlying reviewer graph which reveals that co-review collusiveness and modeled spammer groups as bi-connected graphs. Graph based method can sufficiently utilize the relation of reviewers and mining the relying information, but the disadvantage is that graph can be very big and build big network or graph can be very time costing.

3. Review Group Detection

Browsing the review list of a reviewer is more helpful to fake review detection than browsing the review list of a product or store. The review list of reviewer contains more information of reviewer about review habits, spamming abnormal behavior and regularities.

Previous works in spamming detection mainly focus in fake reviewer detection and considers all the reviews of fake reviewers to be fraudulent which is too arbitrary and coarse-grained. In fact, even a fake reviewer or "pseudo professional reviewer" sometimes posts real reviews and at some other time post fake reviews. Because the fake reviewer may post real reviews to cover up his spamming behavior or post reviews to product he really bought and experienced for personal need. Hence, spamming behavior and normal behavior often occurs intensively. Adjacent reviews often have similar credibility when they are similar in behavior and content. We group similar reviews together from the first review of the reviewer. Then distance formula between a review and group can be defined as follows:

$$dis\ tan\ c\ e(g, re) = \sqrt{\sum_{i=1}^{|v|} \omega_i (v_{gi} - v_i^{re})^2} \quad (1)$$

ω is the weight of review posting time, review rank, store category and review content emotion etc. Initially the weight of time and rank is given a relatively high value(0.35) because of its importance and category and emotion weight is given a relatively low value(0.15). v_{gi} is the vector of group i and v_i^{re} is the vector representation of review re.

The distance computation implies that the larger the distance between a review and group is, the lower the possibility of the review being clustered into this group. The distance measure considers several factors which include time distance, rank distance, category distance and emotional distance. Time distance refers to the post date similarity between a review and group. Rank distance denotes the rank distance between a review and group. Category distance check if the category of the store associated with the review is the same with that of the group. Emotion factor takes the emotion balance of each review into consideration. Emotion can be positive, negative or neutral by means of computing the emotional tendency. Then each review has an index of postive, negative and neutral emotion.

We designed an algorithm (**Algorithm 1**) to detect the groups of reviews of each reviewer. All reviews are parted into many groups with different size. The group detection algorithm is listed below:

Algorithm 1 Sequential Group Detection

Input: R: The reviewer set
RE: set of reviews of each reviewer.
α: max threshold of date distance
δ: distance threshold
Output: group $\mathbf{G_{i,j}}$
Initialization: m=1;
$G_{1,j}$ = {1st review of reviewer j};
currentGroup cg = $G_{1,1}$;

```
1: for each reviewer j in R
2:    for each review RE_i in RE
3:       if cg is null
4:          cg = {RE_i}
5:       if ( date_between ( RE_i , cg ) > α))
6:          output cg
7:          create new group cg = G_{|Gj|+1, j} = {}
```

```
8:              continue
9:          Compute distance β utilizing Eq. (1)
10:         If ( β < δ )
11:             cg = cg∪{ RE_i };
12:             Update centroid of cg;
13:             continue
14:         else
15:             output cg
16:             cg = G_{|Gj|+1}= { RE_i }
17:             continue
18:     end for
19: end for
```

$\mathbf{G_{i,j}}$ denotes the i^{th} group which belongs to reviewer j. RE_i denotes review i in review set RE of a reviewer. In algorithm 1, line 3 and 4 initialize the current group with the current review when current group is empty. Line 5 to 8 consider the review date difference between the current group and current review, if the difference is big enough (than the parameter α), then the review would be put in another new group. In line 9 to 17, first the distance between review and current group is computed, then if the distance surpasses the threshold β, the review is added to the current group, meanwhile, centroid of current group would be updated; Otherwise, new group is constructed.

4 Group Collusion Model

For profit reasons, some people may be employed to cooperate with each other to collude in posting fake reviews to target store or product. If a group of reviewers work together only once to promote or to demote a product of store, it's hard to detect them based on their collective behavior [6]. Since many reviewers collaborate to review the same product and store, the data mining technique frequent itemset mining (FIM) [15] is used to detect groups of reviewers who have ever reviewed at least two (support) stores or products with other reviewers. We use the FPGrowth [14] algorithm which is generally the fastest and most memory efficient algorithm compared with the early Apriori, AprioriTID [15] algorithm. We model the collusion index of each review collusion(r) utilizing the FIM results.

Collusion(r) is the index of reviewer r who have collusion behavior with other reviewers targeting the same stores.

$$collusion(r) = \frac{\sum_t |itemset_{r,t}| \times |support_{r,t}|}{\sum_r \sum_t |itemset_{r,t}| \times |support_{r,t}|} \quad (2)$$

$itemset_{r,t}$ is set of reviewer ids who have reviewed the same two or more stores. $support_{r,t}$ is the support number of the itemset which means the reviewers in itemset have a number of $|support|$ common reviewed stores.

We create user collusion features using different support numbers which include 4,6,8,10,12 to express different degrees of collusion behavior. When the support number

becomes higher, the reviewers cooperate to review more stores which means the probability of collusion is high too.

5 Group Model Constructions

5.1 Reviewer Group Size Related Model

The training set annotation process reveals that big group is less reliable than small groups. Dishonest Reviewers often post many reviews that are similar in posting time, rank preferences, category distribution, content emotional distribution etc. to increase their credibility or to promote the stores that employ them or for other illegal reasons. Bigger group usually means more fake reviews posted with the same intension and target which would save their time and enhance efficiency. We demonstrate three group size related features which include the average reviewer group size (ARGZ), max reviewer group size (MRGZ), and big group size ratio (BGSR) to cover the group size domain. We normalize ARGZ and BGSR to [0,1].

$$ARGZ(r) = \frac{|RE_r|}{|G_r|} \tag{3}$$

$|G_r|$ is the number of groups of reviewer r, $|RE_r|$ is the number of reviews of reviewer r.

$$BGSR(r) = \frac{|BG_r^{\lambda}|}{|G_r|} \tag{4}$$

$\left|BG_r^{\lambda}\right|$ is the number of big group of reviewer r based on the parameter λ which defines the number of reviews per group as a big group. In the experiment, we set λ to five which is medium in size.

5.2 Inner Group Behaviors Model

(1) Inner Group Content Similarity

The content of reviews in the same group is similar when spammers copy their reviews among themselves. So when a review is largely similar to the content of the group, it is more possible to be faked or spammed. The content similarity between review re and group g is defined as:

$$IGCS(re, g) = \max_{re \in g}(cosin(re, re_g)) \tag{5}$$

where g is the group which contains reviews re. re_g is the review set in g except re itself, cosin() is the cosine similarity function.

(2) Inner Group Time Density

If a group contains several reviews with the same posting date which means a "high time density", the group is more likely to be spammed. We model inner group time density as follows:

$$IGTD(re, g) = \frac{|t_{re,g}|}{|g|} \tag{6}$$

$$IGTD(g) = \frac{|t_g|}{|g|} \tag{7}$$

where $|t_{re,g}|$ is the number of reviews in group g that has the same posting time with review re. $|t_g|$ is the number of different post time in group g. $|g|$ denotes the size of group g.

(3) Inner Group Category Density

Professional spammers post reviews with the same category when they are given an order to promote the same series of stores like hotels. Reviews on same category with high frequency are abnormal and the reviewers are suspected to be employees of the company. Lower inner group category density reflects review preferences with diversity that normal reviewers have. Inner group category density between review re and group g is defined as:

$$IGCD(re, g) = \frac{|RE^c_{re,g}|}{|g|} \tag{8}$$

where $|RE^c_{re,g}|$ is number of reviews which has the same category c with review re in group g. $|g|$ denotes the size of group g.

(4) Inner Group Store Similarity

Similar to inner group category density, multiple reviews targeting the same or very similar store is suspicious when they are in the same group because this situation usually does not happen by accident but on purpose. Sometimes spammers post reviews to the store of the subbranch of the same brand. We model the similarity of stores in group as follows:

$$IGSS(re, g) = \max_{re \in g, i}(sim(s_{re}, s_{i,g})) \tag{9}$$

where s_{re} is the store which review re is target on, and $s_{i,g}$ is store i of group g. We omit the situation of similarity of two same stores which is in vain.

(5) Inner Group Rank Diversity

The rank diversity of groups of reviews can also indicate spamming. At an extreme case, the spammer ranks all 5 star to the store he/she reviewed with no rank diversity. The standard deviation is a measure that is used to quantify the amount of variation or dispersion of a set of data values, thus variance of review ratings of group can be modeled as:

$$IGRD(g) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(pr_g - \overline{pr}_g)^2} \quad (10)$$

pr_g denotes the proportion of every star rating. Here, we use the five star rating mechanism and n is set to five. $\overline{pr_g}$ is the average of every pr_g. The inner group rank diversity value attains 0 when all the ratings of reviews are the same like all 5-star rating or all 4-star rating.

Based on the group model, nine features are constructed for the fake reviews detection task as is shown in table 2.

Table 2. Feature Construction Based On Review Group Model

Feature number	Group related feature	Description
F1	Group contents similarity	Max contents similarity in group
F2	Group Time Density	Number of different Date in group / group size
F3	Group Category Density	Number of Category in group / group size
F4	Group Store Similarity	Max Store Similarity in group
F5	Group Rating Diversity	Mean rating variance in group
F6	Group Size	Group size / max group size
F7	Average Group size	Number of reviews in group / number of groups of current reviewer
F8	Max Group Size	Max number of reviews in group
F9	Big Group Ratio	Number of big groups / number of groups of current reviewer

5.3 User Collusion Feature Construction

As is mentioned in section 4, a series of user collusion features are constructed based on the group collusion model using different support number of 4,6,8,10,12 which means that reviewers review at least the same four or more products (stores). We compute the user collusion value using formula (2). As is shown in table 3, five features based on user collusion behavior are constructed to cover different level of collusion behavior. Some other support level of features can be designed like features of support of 3, 5, 7, 9 etc. We select the even support number for simplicity which are described in table 3.

Table 3. Construction of Reviewer Collusion feature of different support level

Feature number	Collusion related feature	Description
F10	Collusion Support 4	Collusion model with support number of 4
F11	Collusion Support 6	Collusion model with support number of 6
F12	Collusion Support 8	Collusion model with support number of 8
F13	Collusion Support 10	Collusion model with support number of 10
F14	Collusion Support 12	Collusion model with support number of 12

6 Experiments and Evaluation

6.1 Datasets and Evaluation Metrics

One of the major challenge in this area of research is availability of gold standard datasets[23]. Very few datasets with spamming labeled are open to researchers. Ott.[4] published a dataset with gold standard deceptive opinions gathered using Amazon Mechanical Turk (AMT) in hotel domain. But the dataset is too small which contains only 1600 opinions total and employed AMTs are not genuine reviewers in post reviews. In addition, paid crowdsourcing like AMT can only imitate the content of genuine reviews but they do not reflect actual behavioral and psychological state of mind of fake reviewers[23] which is rather important in fake review detection. Other datasets of product ratings seldom have the spamming labels and are lack of reviewer behavior information. So, we crawled a dataset of reviews and reviewers from dianping.com which is a popular opinion sharing website as it provides us a real time scenario for data analysis. First, we search the stores of express inn in a given city which is prone to be spammed for economic reasons, and collect all the reviews of reviewers who have reviewed hotels. Due to the large number of reviewers and corresponding reviews in our data, it would have taken too much time for human judges to assess all the reviewers in a short time. We thus only randomly selected 20 reviewers which contain 4189 reviews to label. Two researchers who are familiar with online shopping and rating are involved in the labelling process. They are given the same dataset of 20 reviewers with all of their reviews and metadata of reviewers such as date of register, number of fans, number of flowers got from other reviewers etc. Table 4 shows the statistical information of labeled dataset. Annotators label the credibility of each review with a float number varying from 0 to 1 instead of binary labeling of spam and non-spam since people are often uncertain about the truth of review. Finally, average rating of the two annotators are computed as final annotation of reviews. The credibility of reviews can be used in classification or regression process. We split the dataset into two part: spam and non-spam with threshold of 0.5 of credibility in classification.

For the classification evaluation metrics, we use the precision (P in Table 5), recall (R in Table 5), F1 measure (F1 in Table 5) for classification evaluation and use correlation coefficient and MAE for regression process. Precision, Recall and F1 measure are defined as follows:

$P = TP/(TP+FP)$, $R = TP/(TP+FN)$, $F1 = 2P*R/(P+R)$ which TP means number of instances that are true positive, FP means number of instances that are false positive, FN means number of instances that are false negative.

Table 4. labeled dataset of reviews

dataset	#reviews	#reviewers	Average #review per reviewer	#store	#review (spam)	#review (non-spam)	#store category	#group per reviewer	#review per group
dianping.com	4189	20	209	3658	1974	2215	18	8.6	24.3

6.2 Evaluation Using Classification

Since we have labeled the training data with credibility value between 0 and 1, we consider each review as spam when the credibility is less than 0.5 and take each review as non-spam when the credibility is labeled as 0.5 and above. We select five popular machine learning classifiers in weka [11] which include Ada Boost, Bagging, Logistic Regression, Naïve Bayes and Random Forest and perform 10-fold cross validation.

Table 5 describes reviewer behavior related feature used in previous research[16,24].

Table 5. Behavior feature commonly used in previous research

Feature number	Behavior related feature	Description
F15	Good ratio	Good rating number / number of reviews
F16	Bad ratio	Number of bad rating / number of reviews
F17	User review number ratio	Number of user review / max user review number
F18	Review length [24]	Words count per review
F19	Rating variance [16]	Mean variance delta of user reviews
F20	Review number per day	Average review number / #active date
F21	Favorite review ratio	Number of favorite / number of reviews of current reviewer
F22	Focus review ratio	Number of focus / number of reviews of current reviewer
F23	Fans Review ratio	Number of fans / number of reviews of current reviewer
F24	Category density [16]	Number of reviews in current category of review / number of reviews of reviewer
F25	Review density [16]	Number of reviews in current store / number of reviews of reviewer
F26	Time density [16]	Number of different review date / number of reviews of reviewer

In the classification experiment as is shown in Table 6, five different classifiers are used to evaluate the performance of the four feature sets described in table 2, table 3, table 5 and including N-gram and part of speech (POS) features. N-gram and part of speech (POS) features are commonly used in fake review detection task[4,5,23,24]. Since the reviews in dataset of dianping.com are written in Chinese and there isn't any space between words, we first run the word segmentation tool IKAnalyser and perform the POS tag analysis. Results in Table 6 show that N-gram and POS feature set perform worst in spam detection task which is similar to the results in [24]. This illuminates that content-based method like unigram or bigram and POS feature perform poor in fake review detection task. Spammers will imitate real reviewers and write reviews similar to them with different ratings (high rate for promotion purpose and low rate for defame).

Table 6. Comparison of Classification between Different Feature Composition using Classifier. P, R, F1 denote precision, recall and F1 measure respectively

Classifier / **Feature**	**SVM**			**Bagging**			**Logistic**			**Naïve Bayes**			**Random Forest**		
	P	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**
N-gram+POS[23,24]	0.6	0.6	0.6	0.611	0.611	0.611	0.550	0.553	0.551	0.548	0.548	0.548	0.577	0.579	0.578
Behavior +reviewer meta info(BM) [24]	0.868	0.86	0.864	0.89	0.89	0.89	0.877	0.876	0.876	0.837	0.836	0.836	0.909	0.909	0.909
N-gram+POS +BM+Similarity (NPBS)	0.857	0.850	0.853	0.911	0.910	0.910	0.888	0.886	0.887	0.826	0.825	0.825	0.917	0.917	0.917
Group+ Collusion (BC)	0.924	0.924	0.924	0.955	0.955	0.955	0.897	0.907	0.902	0.876	0.851	0.863	0.964	0.964	0.964
Improvement of (BC)than (NPBS)(%)	6.7	7.4	7.1	4.4	4.5	4.5	1.0	2.1	1.5	5.0	2.6	3.8	4.7	4.7	4.7

The BM feature set includes features F15-F26 which include reviewer basic history meta information and behavior related feature like review density related feature in our previous work[16]. Results of BM feature set shows that behavior and reviewer basic meta information especially the behavior feature can boost the efficiency up to more than 20% than N-Gram and POS feature set. This illuminated that behavior related features contain more information than content of the reviews in fake review detection task. The Group and Collusion feature set (BC) contains the feature F1-F14 described in table 2 and table 3 which are constructed by means of group model in section 5. We take the "NPBS" feature set as the standard comparing to the **BC** feature sets. Results in table 6 show improvement of BC than NPBS up to about 7% in precision, recall and F1 measure when SVM classifier is used. BC features achieve an increasement of performance at about 4.5%-4.7% in F1 measure by utilizing the Bagging and random forest classifier and get 3.8% improvement by the Naive Bayes classifier. Classifier of Bagging and Random Forest do the best compared to other classifiers in almost all the feature sets. Even using the group feature only, it can achieve a rather good result compared to the NPBS feature. Obviously, group related features captured nature of the fake review and reviewer.

Figure 1. shows the performance of four main feature sets test when using five classifers. The group and collusion model can be effectively applied in supervised learning and get a rather good performance, in fact the group and collusion model can also be used in an unsupervised learning way such as review clustering.

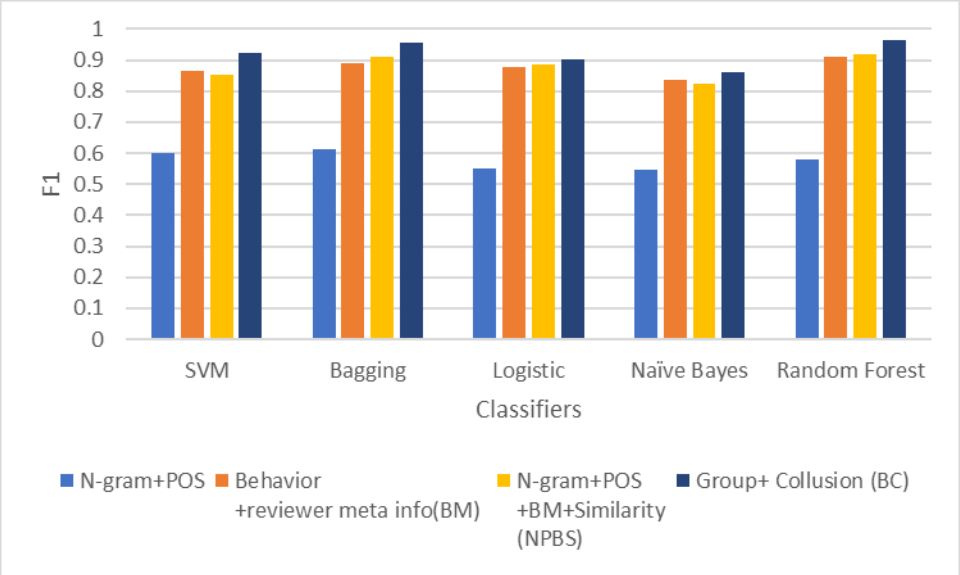

Fig. 1. F1 of different classifier with four types of feature sets

6.3 Regression Evaluation

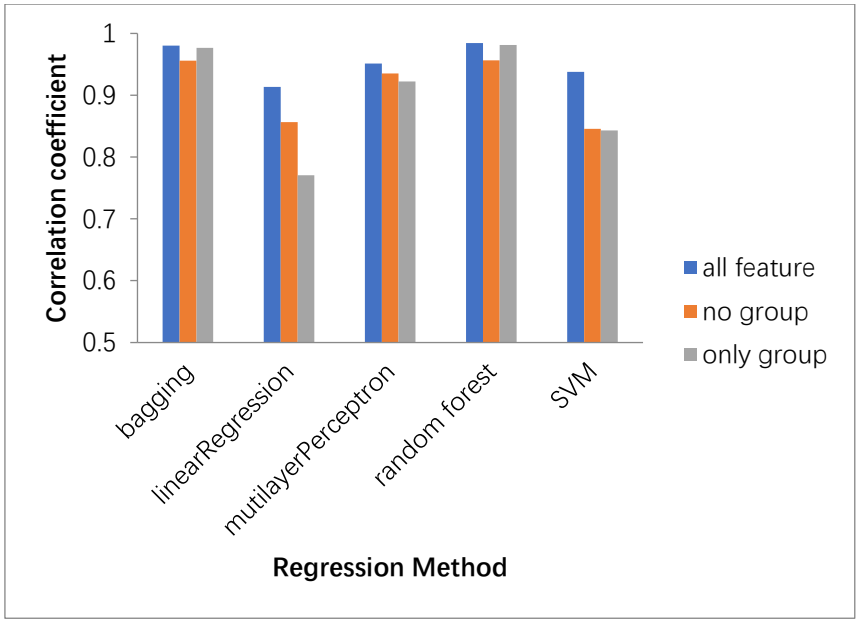

Fig. 2. Correlation Coefficient of five different Regression Method with different Feature Sets

Due to the reason that people sometimes are uncertain on deciding whether the review is fake or not, we label the review with float number ranging from 0 to 1 to represent the credibility of each review. Thus, the newly proposed group model and collusion behavior can be evaluated by regression method. The "all feature" set contains all the feature of N-gram, POS, behavior and reviewer meta information and recommended group features. The "no group" feature set contains the features in "all feature" except the group related feature. "only group" feature set contains only the group related feature of F1 to F14. From the results of regression in Figure. 2, we can see that by using the group model, the correlation coefficient of "all feature" with group model reach high to more than 0.95 which is higher than the situation when no group feature is imported (Figure. 2). The "only group" feature set achieves better results with less features than the "no group" feature set when two regression method(bagging and random forest) are used while the same result occurred when SVM is applied. The "only group" feature set performs not very well utilizing the linear regression because of its simplicity and less features. The mean absolute error (MAE in Figure. 3) of regression process is rather low when bagging and random forest with group feature (all feature) are used. Linear regression performs worst for its simplicity.

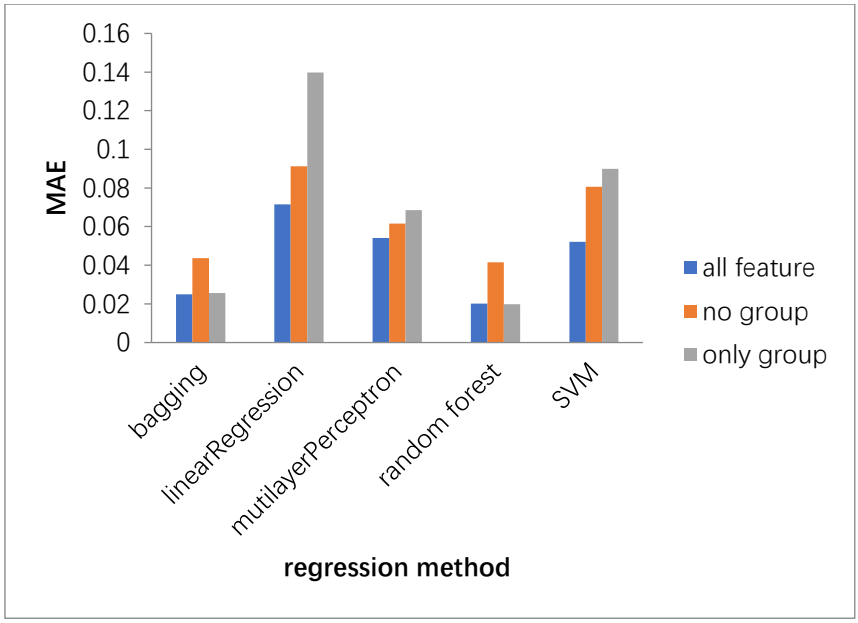

Fig. 3. MAE of five different Regression Method with different Feature Sets

6.4 Group parameter evaluation

Some parameters are to be configured to build review group, one of which is the date threshold α of building groups. The value of parameter α defines the number of days a group has to be closed and need to open a new group. We test different value of date threshold α of building groups on five classifiers. It is shown in the Figure. 4-5 that bagging and random forest perform the best in precision (likewise in recall and F1 measure) whatever parameter value of threshold α is. But the precision, recall and F measure is decreasing when threshold α changes from 0 to 30 or more for bagging, random forest and SVM. This means when the threshold is 0, it performs the best in evaluation. The reason may be that our method can effectively identify useful review groups without manually specify any threshold of dividing groups. The performance of Logistic and Naïve Bayes method rises when date threshold α of building groups changes from zero to seven and then drop when the threshold α becomes bigger than 14. It can be concluded that reviews of reviewer tend to have the same credibility when they are date related. We can also see that performance changed not much with different date threshold using the same classifier which imply that the performance does not only depend on date threshold but the group distance computation model.

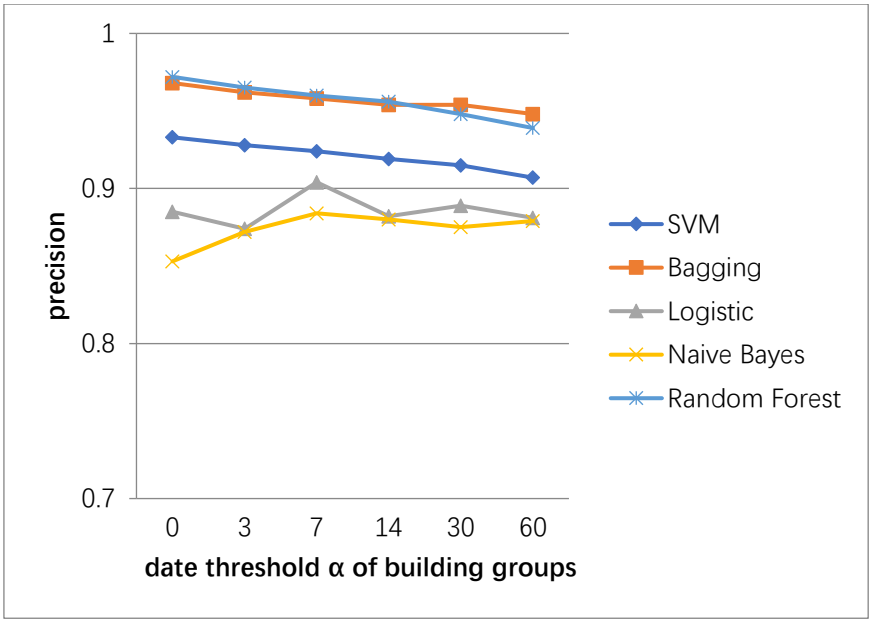

Fig. 4. Precisionof five different Regression Method with different Feature Sets for Group Parameter Evaluation

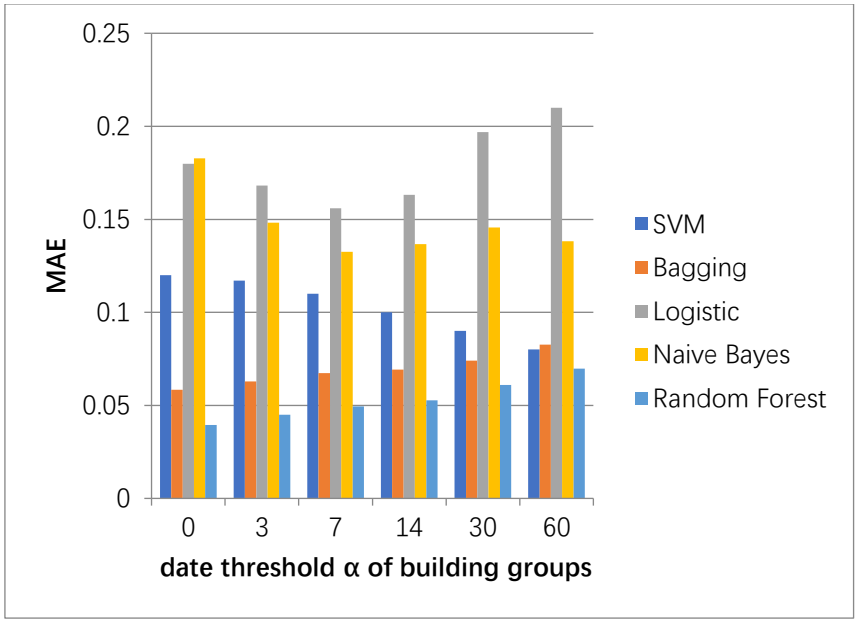

Fig. 5. MAE of five different Regression Method with different Feature Sets for Group Parameter Evaluation

Figure 5 demonstrates the MAE (mean absolute error) of each classifier given different date threshold α in building groups. Still bagging and random forest perform the best and got the lowest MAE. For logistic and naïve Bayes, the MAE level is lower down to bottom when the threshold is 7 and goes up again when the threshold goes up. Generally, date threshold α in building groups gets a rather good performance at the date 7. But when the parameter α is bigger than 14, the performance drops down which means that reviews that posted two weeks ago are not likely to be in the same group which is consistent with our instincts. So in the main experiment shown in table 6., we select 7 as the alpha threshold to get the best result.

7 Conclusion

Review spamming is becoming a serious problem to the development of e-commerce and constantly increasing fake reviews make normal users difficult to find the right product or store to deal with. According the group characteristic of reviews of reviewer, we propose novel review grouping method and models to identify latent fake reviews. The review grouping algorithm can effectively split reviews of reviewer into groups which participate in building new model of review spamming detection. Additionally we explore the collusion behavior between reviewers to build group collusion model. Experiments and evaluations show that the review group method and relevant models is easy to implement and can effectively identify fake reviews especially those posted by professional review spammers. Future works include application of review group model in a unsupervised method and do more work in reviewer collusion behavior via community detection technology.

Acknowledgments. This work has been supported by the National Key R&D Program of China under Grant NO.2017YFB1401000 and the Key Laboratory of Digital Rights Services, which is one of the National Science and Standardization Key Labs for Press and Publication Industry, National Natural Science Foundation of China(NO.61672328) .

References

[1] Jindal, N., Liu, B.: Opinion spam and analysis. In: Proceedings of the 2008 International Conference on Web Search and Data Mining. pp. 219-230(2008)

[2] Mukherjee, A., Kumar, A., Liu, B., Wang, J., Hsu, M., Castellanos, M., Ghosh, R.: Spotting opinion spammers using behavioral footprints. In: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 632-640(2013)

[3] Feng, S., Banerjee, R., Choi, Y.: Syntactic stylometry for deception detection. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers. Vol 2. pp. 171-175(2011)

[4] Ott, M., Choi, Y., Cardie, C., Hancock, J.T.: Finding deceptive opinion spam by any stretch of the imagination. In: Proceedings of the 49th Annual Meeting of the Association for Compu-tational Linguistics: Human Language Technologies. Vol 1. pp. 309-319(2011)

[5] Lau, R.Y., Liao, S., Kwok, R.C.-W., Xu, K., Xia, Y., Li, Y.: Text mining and probabilistic language modeling for online review spam detection. ACM Transactions on Management Information Systems (TMIS) (2011)

[6] Mukherjee, A., Liu, B., Glance, N.: Spotting fake reviewer groups in consumer reviews. In: Proceedings of the 21st international conference on World Wide Web. pp. 191-200(2012)

[7] Fei, G., Mukherjee, A., Liu, B., Hsu, M., Castellanos, M., Ghosh, R.: Exploiting Burstiness in Reviews for Review Spammer Detection. In Proceedings of ICWSM-13. pp. 175-184(2013)

[8] Mukherjee, A., Liu, B., Wang, J., Glance, N., Jindal, N.: Detecting group review spam. In: Proceedings of the 20th international conference companion on World wide web. pp. 93-94(2011)
[9] Stringhini, G., Kruegel, C., Vigna, G.: Detecting spammers on social networks. In: Proceedings of the 26th annual computer security applications conference. pp. 1-9(2010)
[10] Zheng, X., Zeng, Z., Chen, Z., Yu, Y., Rong, C.: Detecting spammers on social networks. Neurocomputing. Vol.159, pp.27-34(2015)
[11] Witten, I.H., Frank, E., Hall, M.A., Pal, C.J.: Data Mining: Practical machine learning tools and techniques. (2016)
[12] Wang, G., Xie, S., Liu, B., Philip, S.Y.: Review graph based online store review spammer detection. In: Data mining (icdm), 2011 ieee 11th international conference on. pp. 1242-1247(2011)
[13] Qin, M., Ke, Y.: Overview of web spammer detection. J. Softw. vol.25, pp.1505-1526(2014)
[14] Han, J., Pei, J., Yin, Y., Mao, R.: Mining frequent patterns without candidate generation: A frequent-pattern tree approach. Data mining and knowledge discovery. vol.8, pp. 53-87(2004)
[15] Agrawal, R., Srikant, R.: Fast algorithms for mining association rules. In: Proc. 20th int. conf. very large data bases, VLDB. pp. 487-499(1994)
[16] Li, Y., Feng, X., Zhang, S.: Detecting Fake Reviews Utilizing Semantic and Emotion Model. In: 2016 3rd International Conference on Information Science and Control Engineering (ICISCE) . pp. 317-320(2016)
[17] Amleshwaram, A.A., Reddy, N., Yadav, S., Gu, G., Yang, C.: Cats: Characterizing automa-tion of twitter spammers. In: Communication Systems and Networks (COMSNETS), 2013 Fifth International Conference on. pp. 1-10(2013)
[18] Lin, C., He, J., Zhou, Y., Yang, X., Chen, K., Song, L.: Analysis and identification of spamming behaviors in sina weibo microblog. In: Proceedings of the 7th Workshop on Social Network Mining and Analysis, pp. 5(2013)
[19] Xu, C.: Detecting collusive spammers in online review communities. In: Proceedings of the sixth workshop on Ph. D. students in information and knowledge management. pp. 33-40(2013)
[20] Li, J., Ott, M., Cardie, C., Hovy, E.: Towards a General Rule for Identifying Deceptive Opinion Spam. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics(2014)
[21] Ott, M., Cardie, C., Hancock, J.: NegativeDeceptiveOpinionSpam. In: Proceedings of NAACL-HLT 2013. pp. 497-501(2013)
[22] Li F , Huang M , Yang Y , et al.: Learning to Identify Review Spam[C]// International Joint Conference on Artificial Intelligence.(2011)
[23] Wang F, Qi S, Gao G, Zhao S, Wang X. : Logo information recognition in large-scale social media data. Multimed Syst. Vol 22(1), pp. 63–73(2016)
[24] Mukherjee A, Venkataraman V, Liu B, Glance NS.: What yelp fake review filter might be doing?. In: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media. (2013)
[25] Wang Z , Gu S , Zhao X , et al.: Graph-based review spammer group detection. Knowledge & Information Systems. Vol. 55(3), pp. 571-597(2018)
[26] Rout J K , Singh S , Jena S K , et al. : Deceptive review detection using labeled and unlabeled data. Multimedia Tools & Applications. Vol. 76(3), pp. 1-25(2017)
[27] Shojaee S, Murad MAA, Bin Azman A, Sharef NM, Nadali S: Detecting deceptive reviews using lexical and syntactic features. In: Proceedings of 13th International Conference on Intelligent Systems Design and Applications. pp 53–58(2013)
[28] Peng Q, Zhong M.: Detecting spam review through sentiment analysis. J Softw. Vol. 9(8), pp.2065–2072(2014)
[29] Qi S, Wang F, Wang X, Wei J, Zhao H. : Live multimedia brand-related data identification in microblog. Neurocomputing. Vol.158, pp. 225–233(2015)