Review Graph based Online Store Review Spammer Detection

Guan Wang, Sihong Xie, Bing Liu, Philip S. Yu
University of Illinois at Chicago
Chicago, USA
gwang26@uic.edu sxie6@uic.edu liub@uic.edu psyu@uic.edu

Abstract— **Online reviews provide valuable information about products and services to consumers. However, spammers are joining the community trying to mislead readers by writing fake reviews. Previous attempts for spammer detection used reviewers' behaviors, text similarity, linguistics features and rating patterns. Those studies are able to identify certain types of spammers, e.g., those who post many similar reviews about one target entity. However, in reality, there are other kinds of spammers who can manipulate their behaviors to act just like genuine reviewers, and thus cannot be detected by the available techniques. In this paper, we propose a novel concept of a heterogeneous review graph to capture the relationships among reviewers, reviews and stores that the reviewers have reviewed. We explore how interactions between nodes in this graph can reveal the cause of spam and propose an iterative model to identify suspicious reviewers. This is the first time such intricate relationships have been identified for review spam detection. We also develop an effective computation method to quantify the trustiness of reviewers, the honesty of reviews, and the reliability of stores. Different from existing approaches, we don't use review text information. Our model is thus complementary to existing approaches and able to find more difficult and subtle spamming activities, which are agreed upon by human judges after they evaluate our results.**[1]

Keywords- *Spammer detection, review graph*

I. INTRODUCTION

Online store reviews are an important resource to help people make wise choices for their purchases. Due to this reason, the review system has become a target of spammers who are usually hired or enticed by companies to write fake reviews to promote their products and services, and/or to distract customers from their competitors. Driven by profits, there are more and more spam reviews in major review websites, such as PriceGrabber.com, Shopzilla.com, or Resellerratings.com. Spammers are starting to corrupt the online review system and confuse the consumers.

A. Challenges

Automatically detecting spammers is an urgent yet under explored task. Unlike other kinds of spam, e.g., Web spam or email spam, review spam is much harder to detect. The main reason is that spammers can easily disguise themselves. It is thus hard for a human user to recognize them, while for Web and email spam, one can tell spam without much difficulty.

[1] This work is supported in part by NSF through grants IIS-0905215, OISE-0968341, DBI-0960443, and IIS-1111092.

The task of detecting fake reviews and reviewers was first proposed by Nitin and Liu in [3], which they call opinion spam detection. They built a classifier using certain types of (near) duplicates reviews as positive training data and the rest as the negative training data. Their approach used features about reviews, reviewers, and products. In [4], they also proposed another algorithm based on mining *unexpected* rules. They found that the top unexpected rules represent abnormal reviews and reviewers. However, their method does not identify true spammers. Using behaviors of reviewers for spammer detection was also studied in [8]. However, this work essentially also relies on duplications, i.e., multiple reviews from the same reviewer targeting the same item or item group. Thus, it is only suited for a special kind of spamming. In [7, 9], review contents were also used for spam detection, which we don't use. Although we also examine reviewers' behaviors, they are only a small part of our solution. The novel review graph concept introduced in this paper captures inter-relationships of reviewers, reviews, and stores, which existing works have not used.

In this work, we are interested in *store* reviews. Even if we can borrow some ideas from previous studies, their clues are not sufficient to tell store review spammers. For example, although it looks suspicious for a person to post multiple reviews to the same product, it may be normal for a person to post more than one review to the same store due to multiple purchasing experiences. Furthermore, as one person has the same writing style on review writing, it may be normal to have similar reviews from one reviewer for multiple stores because unlike different products, different stores basically provide the same types of services. Moreover, many ordinary users only write reviews sporadically. It is reasonable for them to write multiple reviews in a short time period about different shopping experiences. Therefore, reviewer behaviors proposed in the existing approaches for product reviews are insufficient for catching spammers of store reviews. Thus there is a need to look for a more sophisticated and complementary framework. However, the following challenges are major obstacles towards such a framework.

1) There is *no ground truth* whether a review is faked or not. By reading the review text alone, we usually do not have enough clues to tell spam from non-spam.
2) Spammers' behaviors may be hard to capture. For example, in order to successfully mislead customers, spammers can make their writing styles and review habits look very similar to those of genuine reviewers
3) Spammers can also write good and honest reviews,

because they could be real customers of some online stores themselves sometimes. Furthermore, a current genuine reviewer could be a spammer before, and we do not know when a reviewer would write a spam review.

These obstacles are the core challenges that we think make simple behavioral heuristics insufficient. To detect sophisticated spammers, we need to consider more clues.

B. Contribution

Our first contribution is to propose a heterogeneous graph model with three types of nodes to capture spamming clues. We believe that clues for telling if a reviewer is innocent include the reviewer's reviews, all the stores (s)he commented on, and reviews from other reviewers who have shopping experiences in the same stores. Therefore, we propose a novel heterogeneous graph, referred to as the *review graph*, to capture relations among reviewers, reviews, and stores. These are three kinds of nodes in the review graph. A reviewer node has a link to a review if (s)he wrote it. A review node has an edge to a store node if it is about that store. A store is connected to a reviewer via this reviewer's review on the store. Each node is also attached with a set of features. For example, a store node has features about its average rating, its number of reviews, etc. Figure 1 illustrates such a review graph.

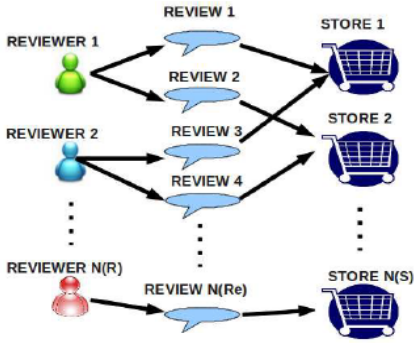

Fig. I Review Graph

Our second contribution is the introduction of three fundamental concepts, i.e. the *trustiness* of reviewers, the *honesty* of reviews, and the *reliability* of stores, and the identification of their interrelationships: a reviewer is more trustworthy if (s)he has written more honesty reviews, a store is more reliable if it has more positive reviews from trustworthy reviewers, and a review is more honest if it is supported by many other honest reviews. Furthermore, if the honesty of a review goes down, it affects the reviewer's trustiness, which has an impact on the store (s)he reviews. And depending on how this reviewer's opinion varies with other reviewers' opinions about the same store, other reviewers' trustiness may change. These intertwined relations are revealed from the review graph.

Our third contribution is the development of an iterative method to compute the three concepts based on the graph model. Starting from the common senses that are uniquely related to the store review system and its spamming scenarios, we derive how one concept impacts another.

To the best of our knowledge, we are the first to use this node reinforcement method to analyze a reviewer's trustiness, a store's reliability, and a review's honesty. Although the general idea of reinforcement based on graph links has been applied in many different scenarios including authority discovery [5] and truth identification from multiple conflicting sources [10], their problem settings and detailed techniques are very different from review spam detection; therefore they are not applicable to our work.

II. SPAM DETECTION MODEL

A. Intuitive Assumptions and Observations

We start by exploring some possible reasons for spamming. Grounded on common sense, we first make the following assumptions:

- Spammers are usually for profits, so they have connections to stores that would benefit spammers to promote their prominence or defame other stores.
- Spammers are usually hired by low quality stores. Such stores have a stronger motivation to hire spammers to write dishonest reviews. Stores with good reputations and stable customer traffic may not hire spammers at all; since they lose much more if they are caught doing so. Even if good stores really entice spammers to say good things about them, it may not be very harmful. Therefore, we assume that less reliable stores are more likely to be involved in review spamming.
- Harmful spam reviews always deviate from the truth. Therefore, they can be either positive reviews about lousy stores, or negative reviews about good stores.
- Not all reviews deviating from mainstream are spam. People may feel differently or have different experiences about the same service.

Based on these assumptions, we have the following observations:

1) We can judge the honesty of a review given the reliability of the store it was posted to, plus the agreement (to be defined later) of the review with its surrounding reviews about the same store.

2) If we have the honesty scores of all the reviews of a reviewer, we can infer his/her trustiness, because one is certainly more trustworthy if one wrote more reviews with high honesty scores.

3) Now we go back to see how to depict a store's reliability. Again from common sense again, a store is more reliable if it is reviewed by more trustworthy reviewers with positive reviews and less reliable if it is reviewed by more trustworthy reviewers with negative reviews.

Figure 2 shows the above influences among a store's *reliability*, a review's *honesty*, and a reviewer's *trustiness*. They are intertwined and affect each other. These influences have some resemblance to the well-known authority and hub relation [5], but the relations in our case are different. First, a reviewer cannot gain more trust by writing more

reviews, nor can we use the mean of review honesty to capture a reviewer's trustiness. Second, reviews have impact on each other. If a review deviates from most of the others, it could be a clue of spamming.

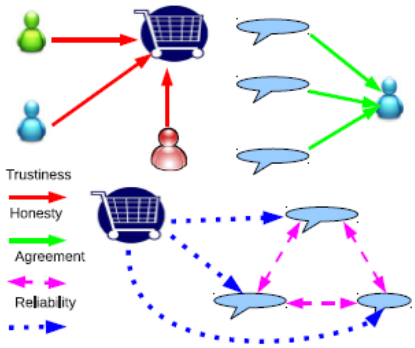

Fig. 2. Influences among different types of nodes

B. Basic Definitions

From the above observations, we define variables that quantify the qualities of reviewers, reviews, and stores.

Definition 2.1 (*Trustiness of reviewers*): The trustiness of a reviewer r (denoted by $T(r)$) is a score of how much we can trust r. For ease of understanding and computation, we limit the range of $T(r)$ to (−1, 1).

Definition 2.2 (*Honesty of reviews*): The honesty of a review re (denoted by $H(v)$) is a score representing how honest the review is, $H(v)$ is within the range (−1, 1).

Definition 2.3 (*Reliability of stores*): The reliability of a store s (denoted by $R(s)$) is a score representing the quality of store s. $R(s)$ is also within the range (−1, 1).

For ease of presentation, we give the notations of node features that are related to the computation of the above definitions in Table I.

Notation	Definition
r	a reviewer
v	a review
s	a store
α^i_r	reviewer r's i^{th} review
β^s_r	reviewer r's review on store s
κ_v	review v's author id
n_r	the number of r's reviews
Hr	honesty summation of all r's reviews
Ψ_v	review v's rating
Γ_v	store id of v's
t_v	review v's posting time
Us	the set of reviews on store s

TABLE I FEATURES ASSOCIATED WITH NODES AND THEIR NOTATIONS

C. Reviewer Trustiness

To model the trustiness, let's see how we can tell if a reviewer is trustworthy or not, based on which we can devise several computational ideas.

1) A reviewer's trustiness doesn't depend on the number of his/her reviews, but on the summation of their honesty scores.

2) A reviewer's trustiness score should be positive or negative if he/she writes more reviews with positive or negative honesty scores.

3) For a given reviewer, his/her trustiness does not grow/drop linearly with the number of high/low honesty reviews that (s)he wrote. It grows/drops faster when the number of such reviews is smaller, and slows down when the number is larger.

We define a reviewer r's trustiness to be dependent on the summation of the honesty scores of all r's reviews,

$$H_r = \sum_{i=1}^{n_r} H(\alpha_r^i) \quad (1)$$

where n_r is the number of reviews from r.

The proposed trustiness score of r should be a function T satisfying the above intuitions, which are formally expressed with the following relations.

$$T(i) < T(j), if\ H_i < H_j \quad (2)$$

$$T(r) < 0\ if\ H_r < 0, T(r) > 0\ if\ H_r > 0 \quad (3)$$

$$\frac{dT(r)}{dH_r} = T(r)(K - T(r)) \quad (4)$$

Relation (2) represents the first idea: One reviewer is more trustworthy than another if one has a larger honesty score. Relation (3) depicts the second idea: We don't trust a reviewer whose reviews tend to be dishonest, and we trust a reviewer whose reviews tend to be honest. Relation (4) represents the growing (or dropping) speed of trustiness is the product of current trustiness level and the room of improvement. Suppose K is the upper bound of the trustiness score, $K-T(r)$ is how much more trustiness one can gain by writing more honest reviews. It becomes smaller as $T(r)$ grows. Solving these elations gives us the general form of trustiness function

$$T(r) = \frac{K}{1 + e^{-KH_r}} \quad (5)$$

Notice this general form is in the range (0, K). As we mentioned before, the upper bound of the trustiness score is +1, and its range should be (−1, 1) to have more practical meanings. We then re-scale the trustiness score T for a reviewer r as

$$T(r) = \frac{2}{1 + e^{-H_r}} - 1 \quad (6)$$

The general shape of the trustiness function is the well-known logistic curve and is shown in Figure 3.

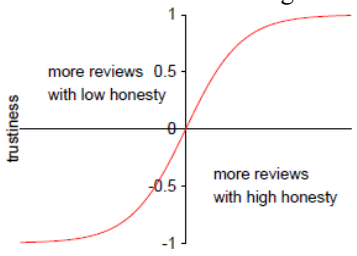

Fig. 3. Relation between *trustiness* and *honesty*

In reality, reviews from the same reviewer can have different degrees of honesty, due to many reasons, e.g.,

subjective bias. A genuine reviewer p may post an unreasonable review i with $H(p(i)) < 0$, while a spammer q may have a honest review j with $H(q(j)) > 0$. However, it is still possible to distinguish malicious reviewers from benign ones, since the trustiness score is controlled by the summation of a reviewer's review honesty, i.e., the general trend of a user's reviews. To calculate $T(r)$, we need the honesty values of r's reviews, which we define next.

D. Review Honesty

How do we interpret a review? When we read a store review, we usually have two things in mind. The first one is the store we are looking at. If the store is a good one such as *Apple.com*, we tend to trust the positive reviews. If we are reading reviews about a store we have never heard of, or we knew it is a bad one, we tend to doubt the high rating reviews. The second factor is the surrounding reviews, which are other reviews about the same store within a certain time window Δt, e.g., 3 months before and after the posting time of the target review. We are likely to trust the mainstream opinions held by most surrounding reviews, rather than outlier opinions. Based on these observations, we model a review's honesty with two factors.

1) The *reliability* of the store it reviews.
2) The *agreement* between this review and other reviews about the same store within a given time window.

We will discuss reliability in the next subsection. We study agreement first by introducing the surrounding set.

Definition 2.4 (*Surrounding Set*): The surrounding set of review v is the set of v's surrounding reviews.

$$S_v=\{i \mid \Gamma_i = \Gamma_v,\ |t_i - t_v| \le \Delta t\}$$

$\forall i, j \in S_v$ agree with each other when their opinions about the same aspects of the store are close. However, opinion mining is too costly to let us assess opinion of a reviewer [2]. Fortunately, in addition to the review text, reviews usually have rating information about a store. Even if two 5-star ratings may mean different aspects of a store, such as customer service and delivery, they are correlated with each other. Therefore, we make some assumptions here.

A review's rating about a store reflects its opinion.
Two reviews with similar rating scores about the same store have similar opinions about the store.

From the assumptions, $\forall i, j \in S_v$ agree with each other if

$$|\Psi_i - \Psi_j| < \delta \quad (7)$$

where δ is a given bound (we use 1 in a 5-star rating system in this paper). Thus, we can partition the surrounding set S_v

$$S_v = S_{v,a} \cup S_{v,d} \quad (8)$$

$$S_{v,a} = \{i \mid |\Psi_i - \Psi_v| < \delta\} \quad (9)$$

$$S_{v,d} = S_v \setminus S_{v,a} \quad (10)$$

We also take the reviewers' trustiness scores into consideration. One review should be good even if it does not agree with any surrounding reviews when it is written by a trustworthy reviewer while the surrounding reviews are posted by untrustworthy reviewers. Similarly, one review may be bad even if it agrees with the surrounding reviews, since they may be all from spammers. Therefore, we define the agreement score of review v within time window Δt as

$$A(v,\Delta t) = \sum_{i \in S_{v,a}} T(\kappa_i) - \sum_{j \in S_{v,d}} T(\kappa_j) \quad (11)$$

Notice that trustiness score T could be either positive or negative. This equation means that if one review agrees with other reviews by trustworthy reviewers, its agreement score increases. On the other hand, if it agrees with untrustworthy reviewers, its score decreases. This equation also promotes a benign user's review agreement score when it is surrounded by spam reviews that it does not agree with, because if spammers' trustiness scores are negative, a benign review's agreement score gets promoted by subtracting the negative numbers. $A(v, \Delta t)$ can be positive or negative. Here we normalize it to (−1, 1) to make later computation easier.

$$A_n(v,\Delta t) = \frac{2}{1+e^{-A(v,\Delta t)}} - 1 \quad (12)$$

Review *re*'s *honesty* $H(v)$ is defined as follows:

$$H(v) = |R(\Gamma_v)| A_n(v,\Delta t) \quad (13)$$

where $R(\Gamma_v)$ is the *reliability* of store Γ_v, which we will define later. Here we just want to note that, by definition, $R(\Gamma_v)$ can be positive (for reliable stores) or negative (for poor quality stores). We take its absolute value as an amplifier of $A_n(v, \Delta t)$. This is consistent with previous discussions. If a store's $|R(\Gamma_v)|$ is large, it is either quite good or quite bad. A review on this store should have a high honesty score if it agrees with many other honest reviews. If $|R(\Gamma_v)|$ is small, the store's reliability is hard to tell. And a review's honesty gets diminished a little by that fact.

E. Store Reliability

To define reliability, we have similar intuitions to trustiness. A store is more reliable if it has more trustworthy reviewers saying good things about it, while it is more unreliable if more trustworthy reviewers complain about it. The increasing/decreasing trend of reliability should also be a logistic curve so as to be consistent with our common sense. Therefore, we define the reliability R of store s as

$$R(s) = \frac{2}{1+e^{-\theta}} - 1 \quad (14)$$

where $\theta = \sum_{v \in Us, T(\kappa_v) > 0} T(\kappa_v)(\Psi_v - \mu)$ (15)

μ is the median value of the entire rating system, e.g., 3-star in a 5-star rating range, and U_s is all reviews of store s.

Therefore, a store's reliability depends on all trustable reviewers who post reviews on it, and their ratings. When considering reliability, we only consider reviewers with positive trustiness score because their ratings really reflect the store's quality. In contrast, whatever a less trustworthy reviewer says about a store, it is less trustable. For example, we don't know the real intention to rate a store as a good

one or bad one, if it comes from a potential spammer.

F. Iterative Computation Framework

Integrating the pieces of information of the review graph together, we have an iterative computation framework to compute reliability, trustiness, and honesty, by exploring the inter-dependencies among them. The algorithm is given in the figure below, which is self-explanatory.

Algorithm 1 Iterative Computation Framework

Input: The set of store $\mathscr{S}$, review $\mathscr{RE}$, and reviewer $\mathscr{R}$
The agreement time window size Δt, review similarity threshold δ, and the $round$ of iterations
Output: The set of $reliability$ R, $honesty$ H, and $trustiness$ T
// Initialization step
Initialize store's reliability and reviewer's trustiness to 1, compute review's agreement using these initial values and Δt, δ according to (11) and (12)
$roundCounter = 0$
while $roundCounter < round$ **do**
  for $re \in \mathscr{RE}$ **do**
    compute $H(re)$ using (13)
  end for
  for $r \in \mathscr{R}$ **do**
    compute $T(r)$ using (6)
  end for
  for $s \in \mathscr{S}$ **do**
    compute $R(s)$ using (14) and (15)
  end for
  for $re \in \mathscr{RE}$ **do**
    update re's agreement using new $\mathscr{T}$ based on (11) and (12)
  end for
  $roundCounter$++
end while

The overall time complexity of the algorithm is

$$O(k(N_r + N_v)),$$

where k is the number of iterations. k is small (usually 4 or 5) as the algorithm converges quite fast in practice.

III. Evaluation

A. Data Set and Data Features

We use the store review data from www.resellerratings.com, a largest host of store reviews, for our experiments. The website provides a unique url for every reviewer's profile, containing meta-data such as reviewer id, all his/her reviews and ratings with posting times about stores, and links to those stores. In each store page, there is information about its average rating score, all reviewers with their reviews.

The data we crawled is a snapshot of all reviews from the website on Oct. 6th, 2010. We clean the data by removing users and stores with no review. After that, we have 343603 reviewers who wrote 408470 reviews on 14561 stores.

B. Spammer Detection Results Evaluation

1) *Evaluation Criteria*: We use IR-based evaluation strategy. First we let our algorithm identify highly suspicious spammer candidates. Then we recruit human judges to make the judgments on the candidates about whether they seem to be real spammers. Therefore, we have precision as our performance measure. Similar evaluation approaches have been used in previous review spam detection research [3, 8]. Therefore this is a well-established way of performance evaluation. The details of our evaluation strategy are as follows.

Human Judgments Consistency Criterion: Since there is no "spammer" label in the review data, human evaluation is necessary. A spammer detection algorithm is effective, if different human evaluators agree with each other about their judgments and concur with the system on the same set of results. Therefore, we use human judgment consistency as another evaluation criterion. Our human evaluators are 3 computer science graduate students who also have extensive online shopping experiences. They work independently on spammer identification.

2) *Human Evaluation Process*: Judging suspicious spammers is a complex task for human and often involves intuition and searching for additional information, especially when we target at more subtle spamming activities. To decide if a candidate is a spammer requires human judges not only to read his/her reviews and ratings, but also to collect evidences from relations with other reviewers, stores, and even the Internet.

To standardize the judgment process, our human evaluators agree upon three conditions as our evidence to claim that a reviewer is a potential spammer.

- A reviewer is suspicious if (s)he has a significant number of reviews giving opposite opinions to others' reviews about the same stores. For example, if a reviewer gives high ratings to all stores he has reviewed, while other reviewers rate these stores low, the reviewer is problematic.
- A reviewer is suspicious if (s)he has a significant number of reviews with opposite opinions about some stores to the ratings from the Better Business Bureaus (BBB)[2] For example, if a reviewer gives high ratings to all the stores (s)he reviewed, but BBB gives them Fs (the rating range is from A to F), the reviewer is clearly suspicious.
- A reviewer is suspicious if (s)he has a significant number of reviews saying opposite opinions about some stores as compared to evidences presented by general web search results. For example, if a reviewer gives high ratings to all stores he/she reviewed, but Google search results about these stores often contain information of them having fake reviews, this reviewer is problematic.

Note that each of the above steps involves human labor effort, intuition, and background knowledge, so the entire evaluation is very hard to be computerized. For example, for the third step, our human judges actually need to go through several search result pages and understand the content, in order to make a judgment.

Although every single condition may not be convincing enough to prove spamming activities, all of them together can be a confident claim of spamming. We evaluate the top 100 suspicious reviewers identified by our model. Note that no existing work focused on such a large scale and subtle individual cases. Our human evaluators gave their independent judgments based only on information from

[2] Better Business Bureau is a well-known corporation that endeavors to a fair and effective marketplace. It gathers reports on business reliability, alert the public to business or consumer scams, and enforce the mutual trustiness between consumers and companies

resellerratings, business honesty information from BBB, and search results from Google, and by reading reviews.

3) *Precision and Consistency:* In our evaluation, if more than one evaluator regards a reviewer as a spammer, we label it as a suspicious spammer. Our evaluators identified 49 out of 100 suspicious candidates to be spammers. The precision is 49%. Although our precision is not very high, we are dealing with much more subtle and complex cases (not simple duplications), which existing studies could not handle. Besides, our precision is meaningful, since human evaluators agree with each other on their judgments.

Table II shows the agreement of human judges. For example, Evaluator 1 identified 49 suspicious reviewers, out of which 33 were recognized by Evaluator 2 and 37 were caught by Evaluator 3. To study their agreement, we used *Fleiss' kappa* [1], which is an inter-evaluator agreement measure for any number of evaluators. The kappa among 3 evaluators is 60.3%, which represents almost substantial agreement [6]. It is also important to note that those reviewers who were not identified as spammers are not necessarily innocent. It is simply because our judges did not found enough evidences to conclude they are spammers.

	Evaluator 1	Evaluator 2	Evaluator 3
Evaluator 1	**49**	33	37
Evaluator 2		**35**	23
Evaluator 3			**40**

TABLE II: HUMAN EVALUATION RESULT

4) *Compare with Baseline*: Since our work is the first to use a review graph and target at subtle spamming activities, there is no existing work to compare. Given the differences between our work and previous studies, we want to demonstrate that suspicious reviewers found by our method can hardly be identified by existing techniques.

We choose to compare with the approach in Lim et al. [8], because it is the state-of-art of behavior based spammer detection techniques. They use some types of duplicate reviews as a strong evidence of spamming. They first look for candidates who have multiple reviews about one target (in our case store), and then compute spamming scores to capture spammers from the candidates.

In our top 100 suspicious candidate list, only 3 candidates can be found based on the duplication criterion. Moreover, only 1 out of these 3 candidates is finally labeled as a spammer by our human evaluators. This result means that our work detects different types of spamming activities from existing researches. They can seldom find the spammer types that we can find. By no means do we claim that the existing methods are not useful. Instead, our method aims to find those that cannot be found by previous methods.

5) *Suspicious Spammer Case Study*: Here we take a close look at reviewers ranked highly suspicious. We pick two candidates: one often promotes stores while the other demotes stores. The first case is the reviewer howcome[3], whose reviews are mostly positive. When we examined these reviews, we found many of them are problematic, such as those highly rated reviews for UBid, OnRebate, ISquared Inc, Batteries.com, and BigCrazyStore.com. For example, UBid is widely complained at different review websites like ConsumerAffairs, epinions, CrimesOfPersuasion, and ResellerRatings. OnRebate was sued for failing to pay rebates to customers. ISquared Inc was rated D by BBB. Batteries.com is generally lowly rated at ResellerRatings. BigCrazyStore has very few records about its quality on the Internet. And this reviewer's review is the only one about BigCrazyStore on ResellerRatings. All these evidences lead us to conclude that this reviewer is suspicious.

The second case is shibbyjk[4]. All reviews of this reviewer are complaints. All stores being complained are high quality stores according to BBB, e.g., 1SaleADay(A), StarMicro(B+), 3B Tech(B), and Accstation(A−). From these unfair ratings, one may argue that this reviewer may be still normal but just picky. However, after reading all his/her negative reviews, we found that all complaints are about transaction problems. It is fishy because the chance of all these quality companies having transaction problems and credit card frauds with this particular customer is very low. Besides, transaction problems are a good excuse to make false complains, since the truth is hard to be verified.

IV. CONCLUTION

This paper proposed a novel review graph model and an iterative method utilizing influences among reviewers, reviews, and stores to detect spammers. The method showed how the information in the review graph indicates the causes for spamming and reveals important clues of different types of spammers. Experimental results show that the proposed method can identify subtle spamming activities with good precision and human evaluator agreement.

REFERENCES

[1] J.L. Fleiss and J. Cohen. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educational and Psychological Measurement, 1973.

[2] M. Hu and B. Liu. Mining and summarizing customer reviews. In KDD, 2004.

[3] N. Jindal and B. Liu. Opinion spam and analysis. In WSDM '08,

[4] N. Jindal, B. Liu, and E.P. Lim. Finding unusual review patterns using unexpected rules. In CIKM, 2010

[5] J. Kleinberg. Authoritative sources in a hyperlinked environment. In Journal of the ACM (JACM), 1999.

[6] J.R. Landis and G.G. Koch. The measurement of observer agreement for categorical data. Biometrics, 1977.

[7] F. Li, M. Huang, Y. Yang, X. Zhu. Learning to identify review Spam. In IJCAI, 2011.

[8] E.P. Lim, V.A. Nguyen, N. Jindal, B. Liu, and H.W. Lauw. Detecting product review spammers using rating behaviors. In CIKM, 2010.

[9] Ott, M., Y. Choi, C. Cardie, and J.T. Hancock. Finding deceptive opinion spam by any stretch of the imagination. In ACL, 2011.

[10] X. Yin, J. Han, and P.S. Yu. Truth discovery with multiple conflicting information providers on the web. In TKDE, 2008.

[3] http://www.resellerratings.com/profile.pl?user=57665

[4] http://www.resellerratings.com/profile.pl?user=565595