

## SRC Model to Identify Beguiling Reviews

TANYA GERA, DEEPAK THAKUR AND JAITEG SINGH

Chitkara University, Punjab, India

**E-mail: tanya.gera@chitkara.edu.in**

Received: December 18, 2014| Revised: February 17, 2015| Accepted: May 12, 2015

Published Online: June 29, 2015

The Author(s) 2015. This article is published with open access at [www.chitkara.edu.in/publications](http://www.chitkara.edu.in/publications)

**Abstract:** Today, E-trade sites are giving colossal number of platform to clients in which they can express their perspectives, their suppositions and post their audits about the items on the web. Such substance helped by clients is accessible for different clients and makers as a significant wellspring of data. This data is useful in taking imperative business choices. Despite the fact that this data impact the purchasing choice of a client, however quality control on this client created information is not guaranteed, as audit area is an open stage accessible to all. Anybody can compose anything on web which may incorporate surveys which are not true. As the prevalence of Ecommerce destinations are hugely expanding, nature of the surveys is deteriorating step by step subsequently influencing clients' purchasing choices. This has turned into an enormous social issue. From numerous years, email spam and web spam were the two primary highlighted social issues. At the same time these days, because of notoriety of clients' enthusiasm toward internet shopping and their reliance on the online audits, it turned into a real focus for audit spammers to delude clients by composing sham surveys for target items. To the best of our insight, very little study is accounted for in regards to this issue reliability of online reviews. To begin with paper was distributed in 2007 by NITIN JINDAL & BING LIU in regards to review Spam detection. In the past few years, variety of techniques has been recommended by researchers to accord with this trouble. This paper intends to introduce Suspicious Review Classifier model (SRC) for identifying suspicious review, review spammers and their group.

**Keywords:** Rule based classification; Rule Matrix; Suspicious Review Classifier (SRC); Review Matrix.

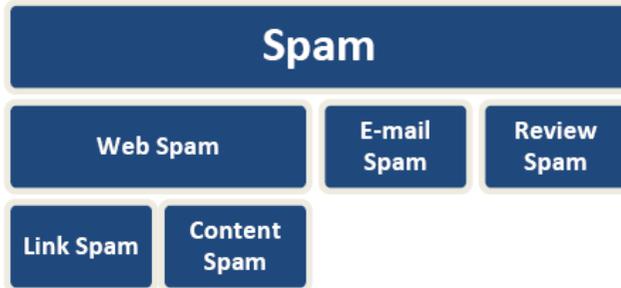
Journal on Today's Ideas –  
Tomorrow's Technologies,  
Vol. 3, No. 1,  
June 2015  
pp. 41–51

## 1. INTRODUCTION

Web has been continually giving important wellspring of opinions on items, administrations, occasions, people, and so forth. Numerous scientists have helped in the field of sentiments extraction on item audit sites, gathering posts, and online journals. Nonetheless, a large portion of the work has been centered on order and outline of presumptions utilizing regular dialect preparing and information mining methods. An imperative and genuine issue that has been disregarded so far is assessment spam or constancy of those online assessments. These exercises are performed by some online characters that cheat others and the term utilized for them is sock puppeting. These days, it is peaceful basic that before shopping, clients like to peruse other's encounters on item survey locales. More number of positive surveys sways them to purchase the item and then again, by discovering negative audits clients get disheartened. In the event that these online surveys on which your choices, your buys, your cash consumption transfer are discovered to be NOT Genuine then You will without a doubt feel like being duped, its similar to as though someone has harmed your feelings, your significant cash is squandered. Also, the above all, it is the greatest danger to the general public. Bewilder remarks guarantee profit for associations and people. This, shockingly, has turned into a huge wellspring of wage for some notion spammers at the expense of deliberately selling out the clients. The terms like survey spammers, sentiment spammers, fake analysts are utilized conversely. Conclusion spam can exist in assortment of diverse structures e.g., fake surveys, delusive remarks, un-trustful online journals, fanciful informal community postings, double dealings, deceptive messages, counterfeit audits.

By and large, there are three sorts of spam- Web spam, email spam and survey spam. The point of Web spam is to charm the individuals to visit some target pages and therefore raising the rank of those. An alternate type of spam is email spam, which is additionally very not the same as audit spam. Email spam (additionally called garbage messages) includes acquiring the uninvited business commercials.

Untruthful opinion spam is much harder to manage on the grounds that; one can't say that a specific audit is not a bona fide one by simply manually understanding them. Spammers deliberately compose a fake survey which is by all accounts genuine and authentic knowledge of some client. Obtaining conduct of a large portion of the clients is affected by the nature of data gave to them through destinations. Presumption spammer lives up to expectations for some target items, administrations, associations, people, and even without uncovering their actual propositions. Such illicit exercises are stain for our

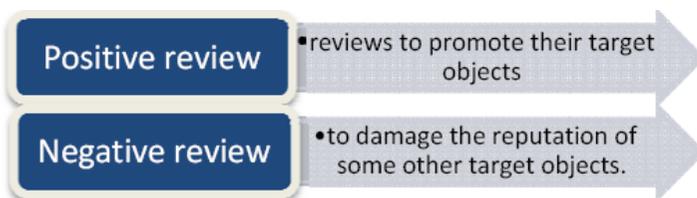


**Figure 1:** Various Types of Spam.

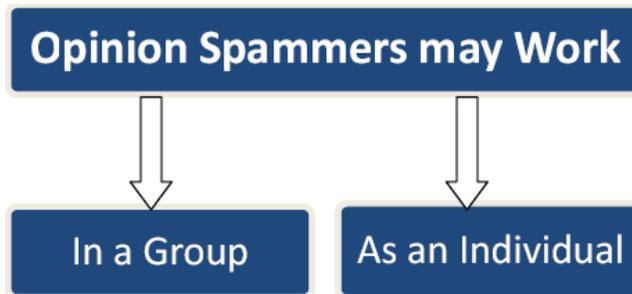
general public. Along these lines, fake survey identification calculations ought to concentrate on recognizing spam audits.

#### **A. Criteria Followed by Spammers to Promote/Demote a Target product**

Spammers may work in a gathering and in addition a single person. On account of the spammer working exclusively focusing on an item, he/she simply composes fake audit him/herself utilizing a solitary client id. These sorts of spammers don't like to work with anybody. On the other side, gathering of spammers works in arrangement to push a target element and/or to harm the notoriety of an alternate. The individual spammers might know one another. Furthermore a solitary individual can likewise carry on much the same as a gathering of spammers by enrolling various client ids and perform spam exercises utilizing these client ids. In any case, a fake commentator bunch (a gathering of commentators who work collectively to compose fake audits) is much all the more harming as they can take aggregate control of the notion on the target item because of its sheer size. Note that by a gathering of analysts, we mean number of analyst ids. The genuine commentators behind the ids could be a solitary individual with numerous ids, various persons, or a blend of both.



**Figure2:** Types of Review Posted on Ecommerce Sites by Spammers.



**Figure 3:** How Spammers work.

## 2. RELATED WORK

In 1997, shingle technique [1] was presented by A.Z. BRODER for figuring out similarity and control in the two documents by computing similarity score. At that point that idea was connected on two sentences to check closeness between them. In 2007[2], NITIN JINDAL & BING LIU proposed Review Spam Detection Mechanism, in which copy audits were concentrated utilizing shingle technique [1]. At that point on rest of the surveys, logistic relapse was connected to group them as a spam or not spam. In 2008[3], NITIN JINDAL & BING LIU Introduced new strategy for ordering. In 1997, shingle framework [1] was introduced by A.z. BRODER for evaluating similarity and regulation in the two reports by enlisting comparability score. By then that thought was associated on two sentences to check closeness between them. In 2007[2], NITIN JINDAL & BING LIU proposed Audit Spam Identification Component, in which duplicate reviews were concentrated using shingle method [1]. By then on rest of the reviews, logistic backslide was associated with orchestrate them as a spam or not spam. In 2008[3], NITIN JINDAL & BING LIU Presented new framework for requesting reviews into three classes, for instance, type1, type2, type3. Type 1 as untruthful notions, Type2 as reviews which talks of brand simply and Type3 as non-overviews which join request, answers, self-assertive substance et cetera. They amassed a classifier using certain sorts of (close) duplicates reviews as positive planning data and the rest as the negative get ready data. Their approach used eccentricities about reviews, examiners, and things however manual stamping and long logistic backslide were considered as delimits. In 2010[4], Ee-Peng Lim, Viet-A Nguyen, Nitin Jindal, Bing Liu, Hady W. Lauw focused on distinguishment of overview spammers using rating behavior. They perceive a couple of trademark practices of review spammers and model these practices to recognize the spammers. In 2011 [7], Arjun Mukherjee, Bing Liu, Junhui Wang, Natalie Look, Nitin Jindal proposed a capable technique to recognize a social

---

occasion of experts who simply collaborated once to lift or to minimization a lone thing or various other target things. In 2011[8], Fangtao Li, Minlie Huang, Yi Yang and Xiaoyan Zhu worked for Figuring out how to Distinguish Audit Spam in which they use oversaw learning methods and break down the effect of various contrivances in study spam unmistakable confirmation. In 2011 [9], Wang, Guan, Sihong Xie, Bing Liu, and Philip S. Yu helped towards review spam area using overview diagrams. Their technique showed how the information in the review graph exhibits the establishments for spamming and reveals vital indications of different sorts of spammers. Trial happens moreover exhibited that the procedure can perceive inconspicuous spamming activities with incredible precision and human evaluator understanding. In 2012[11], Arjun Mukherjee, Bing Liu and Natalie Look helped towards ID of fake review social affairs using Incessant thing set mining method and a couple of behavioral models. In 2013 [13], Arjun Mukharjee, Abhinav Kumar, Bing Liu made a principled method to experience viewed studying practices to get suspicion spammers (fake examiners) in an unsupervised Bayesian construing schema.g audits into three classes, for example, type1, type2, type3. Type1 as untruthful sentiments, Type2 as audits which talks of brand just and Type3 as non-surveys which incorporate inquiry, answers, arbitrary content and so forth. They assembled a classifier utilizing certain sorts of (close) copies audits as positive preparing information and the rest as the negative preparing information. Their methodology utilized peculiarities about audits, commentators, and items yet manual naming and extensive logistic relapse were considered as delimits. In 2010[4], Ee-Peng Lim, Viet-A Nguyen, Nitin Jindal, Bing Liu, Hady W. Lauw concentrated on location of survey spammers utilizing rating conduct. They recognize a few trademark practices of survey spammers and model these practices to catch the spammers. In 2011 [7], Arjun Mukherjee, Bing Liu, Junhui Wang, Natalie Glance, Nitin Jindal proposed a successful procedure to identify a gathering of analysts who just cooperated once to elevate or to downgrade a solitary item or numerous other target items. In 2011[8], Fangtao Li, Minlie Huang, Yi Yang and Xiaoyan Zhu worked for Learning to Identify Review Spam in which they utilize regulated learning systems and examine the impact of diverse peculiarities in survey spam ID. In 2011 [9], Wang, Guan, Sihong Xie, Bing Liu, and Philip S. Yu helped towards survey spam discovery utilizing audit charts. Their strategy indicated how the data in the survey diagram shows the foundations for spamming and uncovers critical hints of diverse sorts of spammers. Test comes about likewise demonstrated that the technique can recognize unpretentious spamming exercises with great accuracy and human evaluator understanding. In 2012[11], Arjun Mukherjee, Bing Liu and Natalie Glance helped towards location of fake audit gatherings utilizing Frequent thing set mining technique and a few behavioral models. In 2013[13], Arjun Mukharjee,

---

Gera, T  
Thakur, D  
Singh, J

Abhinav Kumar, Bing Liu constructed a principled strategy to adventure watched inspecting practices to distinguish notion spammers (fake commentators) in an unsupervised Bayesian deduction schema.

### 3. PROBLEM DEFINITION

---

These days, client's purchasing choices are subject to online opinions given by different clients. With respect to them part of profitable data about items and administrations are given to buyers on web. Then again, from late numerous years, this social group has been joined by spammers whose destination is to mislead pursuers by posting fake surveys on item audit sites. Spammers are vigorously paid by some association to diversion the entire arrangement of data by notion spamming (e.g., composing fake surveys). Their target is to elevate or to downgrade the notoriety of some other skillful target items. This circumstance now requests, recognizable proof and identification of fake surveys and fake analysts; as this has turned into a huge social stain. Past endeavors for spam discovery incorporate conduct of analysts, content deception, phonetics peculiarities and rating examples. Those studies have the capacity recognize certain sorts of spammers, e.g., the individuals who post numerous comparable audits around one target element. Be that as it may, in actuality, there are different sorts of spammers who can control their practices to act much the same as authentic analysts.

### 4. PROPOSED WORK

Review Spam Detection Mechanism [2] was the building block of spam detection scheme which include detection based on duplicate finding and classification. This work intends to introduce a new rule based classification method for identifying suspicious reviews on product review websites.

### 5. METHODOLOGY

This work includes **rule based classification detection** in which certain number of rules are applied on review dataset crawled from product review websites. The rule contains some pre-defined characteristics about reviews which can be untruthful. Reviews in review dataset are classified in resultant separate categories if it entails characteristics of rules. This will lead to produce separate lists of identified suspicious reviews. The number of output lists is one more than number of number of rules. Rules which are applied over the set of reviews are explained in the algorithm given below:

Algorithm: Rule Based Classification

**Input:** Review Dataset.

**Algorithm:**

**Step 1:** Search review repository for exactly duplicates.

**Step 2:** For Product M: Extract all data entries for which  
CONTENT [REVIEW X] == CONTENT [REVIEW Y];

**Step 3:** Maintain two separate classes:

    If (USER [REVIEW X] == USER [REVIEW Y])

        Insert in class I.

    Else

        Insert in class II.

**Step 4:** Repeat Step 2 & Step 3 for each product.

**Step 5:** For all pairs of product M and N: Extract all data  
entries for which CONTENT [REVIEW X] == CONTENT  
[REVIEW Y];

**Step6:** Maintain two separate classes:

    If (USER [REVIEW X] == USER [REVIEW Y])

        Insert in class III.

    Else

        Insert in class IV.

**Step 7:** For Product M: Extract all data entries for which  
CONTENT [REVIEW X] != CONTENT [REVIEW Y];

**Step8:** Maintain two separate classes:

    If (USER [REVIEW X] == USER [REVIEW Y])

        Insert in class V.

    Else

        Insert in class VI

**Step 9:** STOP.

Here, REVIEW X denotes review id (unique identification for review in database); CONTENT [REVIEW X] is comment in a particular review; USER [REVIEW] denotes the review posted by which user id.

The methodology explained above results in one more than number of rules applied i.e. total number of output lists would be number of rules applied. *For example:* Six rules are expected to produce six output lists which contains suspicious reviews, classified on the basis of characteristics matched with rules. Rests of non-suspicious reviews are entailed under another new list. In this manner, output will contain some suspicious lists of reviews and one non-suspicious list of reviews.

---

---

Gera, T  
Thakur, D  
Singh, J

---

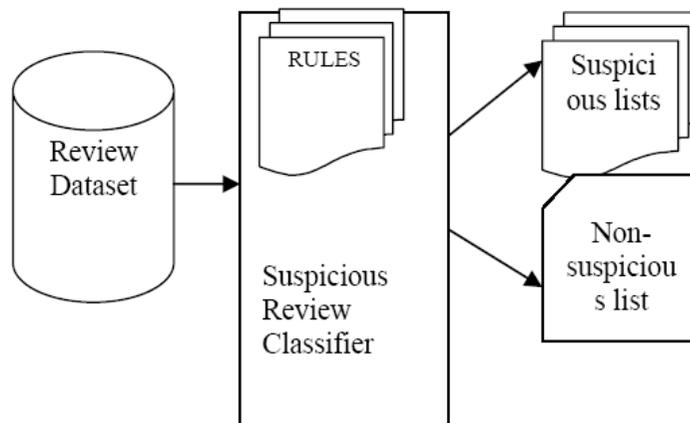
This work makes use of six numbers of rules which are responsible for uncovering six lists of suspicious reviews. The possible number of scenarios which are covered by the rules is stated below:

1. Duplicate review posted by same user on same product.
2. Duplicate review posted by same user on different products.
3. Duplicate review posted by different users on same product.
4. Duplicate review posted by different users on different product.
5. Non-duplicate reviews posted by same user on same product.
6. Non-duplicate reviews posted by same user on different product.

All the cases described above form the basis for suspicion. As it can be observed that exactly same review posted by same/ different user-ids multiples times under same product category arise suspicion. Similarly duplicates from same/ different user-ids multiples times under different product categories point towards suspicion. On the other hand, different reviews posted by same user-id under same/different products give a point for suspicion. All these possible cases have been covered by SRC model. Consider all the lists of suspicious and non-suspicious reviews as individual classes ( $c_1, c_2, \dots, c_n$ ). Baye's rule is helpful in predicting the probability that a review will fall in which class.

$$P(C_i / R) = \frac{(P(R / C_i) * P(C_i))}{(P(R))} \quad (1)$$

Probability of reviews that it belongs to class  $C_i$  is multiplication of probability of  $R$  given  $C$  and probability of class, divided by probability of review. Here denominator doesn't matter because denominator will remain



**Figure 4:** Rule based classification of reviews among separate lists.

---

same for all competing classes. It is assumed that at earlier there is equally likely chance for review to fall in each of classes. So probability of every class is same i.e.  $P(CI) = 1/7$ . Now the whole probability relies on term in numerator. A review is supposed to be a member of that class for which value of numerator is highest.

For quick suspicious reviews identification, a *Review Matrix* is generated. A Review Matrix contains reviews as rows and rules applied as columns. In the matrix value 1 specifies that review is identified as suspicious based on column number rule and 0 otherwise. A sample value matrix is shown below:

This matrix results in the observation that more the number of 1s in tuple, more is the chances of its been a suspicious review. In the rule based classification, accuracy and coverage can be found as:

$$\text{Coverage of rule} = \frac{N_{\text{covers}}}{|N|} \quad (2)$$

$$\text{Accuracy of Rule} = \frac{N_{\text{correct}}}{N_{\text{covers}}} \quad (3)$$

Here, coverage of each rule can be computed as a ratio of numbers of reviews covered by that rule and total number of reviews.  $N_{\text{covers}}$  specifies number of reviews covered by rule.  $|N|$  is total number of tuples.

Accuracy of rule can be defined in terms of ratio of numbers of correct tuples covered by that rule and total number of tuples covered.

## 6. CONCLUSION

In the past few years, the major highlighted field of research was sentiment analysis. These studies assume all the reviews to be genuine. However, due to increase in demand of online shopping, spam has become a big social issue. It is important to identify and detect those review spam. This works intends to introduce rule based classification method to identify suspicious reviews on shopping websites. The terms like *coverage and accuracy* can also be computed which help in producing how accurate the results would be.

## 7. FUTURE SCOPE

This paper proposes a novel approach to identify the suspicion over each review posted on product review sites. The proposed SRC Model includes rules based classification of reviews and identifying accuracy of rules. In our

**Table 1:** Sample Values for Review Matrix.

Sample	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6
Review 1	1	0	0	0	0	0
Review 2	1	0	1	0	0	0
Review 3	0	0	0	1	0	1

---

future work, we would implement the proposed SRC Model that will definitely help to produce results with acceptable accuracy. It is ascertain that some more rules and parameters would be inducted to this approach so as to provide better rating quality for a product.

## REFERENCES

- [1] A. Mukherjee, A. Kumar, B. Liu, J. Wang, M. Hsu, Malu Castellanos, and R. Ghosh. "Spotting Opinion Spammers using Behavioral Footprints," Proceedings of SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2013), August 11-14 2013 in Chicago, USA.
- [2] A. Mukherjee, B. Liu, and N. Glance, "Spotting Fake Reviewer Groups" <http://dx.doi.org/10.1145/2187836.2187863>
- [3] A. Mukherjee, B. Liu, J. Wang, N. Glance, N. Jindal. "Detecting Group Review Spam," WWW-2011 poster paper, 2011.
- [4] A. Mukherjee, V. Venkataraman, B. Liu, and N. Glance. "What Yelp Fake Review Filter Might Be Doing," Proceedings of The International AAAI Conference on Weblogs and Social Media (ICWSM-2013), July 8-10, 2013, Boston, USA
- [5] A. Z. Broder, "On the resemblance and containment of documents," Proceedings of Compression and Complexity of Sequences 1997, IEEE Computer Society, 1997.
- [6] Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, Bing Liu and H. Lauw. "Detecting Product Review Spammers using Rating Behaviors," Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM-2010, full paper), Toronto, Canada, Oct 26 - 30, 2010.
- [7] G. Fei, A. Mukherjee, B. Liu, M. Hsu, Malu Castellanos, and R. Ghosh., "Exploiting Burstiness in Reviews for Review Spammer Detection," Proceedings of The International AAAI Conference on Weblogs and Social Media (ICWSM-2013), July 8-10, 2013, Boston, USA.
- [8] G. Wang, S. Xie, B. Liu, Philip S. Yu. "Identify Online Store Review Spammers via Social Review Graph," ACM Transactions on Intelligent Systems and Technology, accepted for publication, 2011
- [9] G. Wang, S. Xie, B. Liu, Philip S. Yu. "Review Graph based Online Store Review Spammer Detection," ICDM-2011, 2011.
- [10] <http://consumerist.com/2010/04/14/how-you-spot-fake-online-reviews/> last accessed on 29.05.2014 at 3:00pm.
- [11] <https://support.google.com/places/answer/2622994?hl=en/> last accessed on 27.08.2014 at 4:00pm.

- [12] Li, Fangtao, M. Huang, Y. Yang, and X. Zhu. "Learning to Identify Review Spam," Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-2011). 2011.
- [13] N. Jindal and B. Liu, "Opinion Spam and Analysis," Proceedings of <http://dx.doi.org/10.1145/1341531.1341560>
- [14] N. Jindal and B. Liu. "Review Spam Detection" Proceedings of WWW-2007 (poster paper), May 8-12, Banff, Canada.
- [15] N. Jindal, B. Liu and E. Lim. "Finding Unusual Review Patterns Using <http://dx.doi.org/10.1145/1871437.1871669>
- [16] Ott, M., Y. Choi, C. Cardie, and J.T. Hancock. "Finding deceptive opinion spam by any stretch of the imagination," ACL, 2011.
- [17] S. DIXIT & A.J.AGRawal "SURVEY ON REVIEW SPAM DETECTION," International Journal of Computer & Communication Technology ISSN (PRINT): 0975 - 7449, Volume-4, Issue-2, 2013.
- [18] S. Feng, L. Xing, A. Gogar, and Y. Choi. "Distributional Footprints of Deceptive Product Reviews," Proceeding of ICWSM. 2012
- [19] T. Qian, B. Liu., "Identifying Multiple Userids of the Same Author," Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP-2013), October 18-21, 2013, Seattle, USA