

# Social Trust Analysis: How Your Behavior on the Web Determines Reliability of the Information You Generate?



Rhea Sanjay Sukthanker and K. Saravanakumar

**Abstract** Can I trust a review? A very common question for someone accustomed to online shopping. The Internet hosts a large number of reviews. Many e-commerce Web sites like Amazon, eBay, Flipkart ask their customers for their reviews once the product is bought. There is an important aspect of trust in an online context. Often reviews diverge widely on their star ratings from 1–5 which clearly show bias for a brand or product. What actually guarantees the reliability of a review? Some of the effective ways to ensure the trustworthiness of a review are to use the reviewer profile information and his previous reviews. Opinions of others have a greater impact on consumers rather than verified information provided by the product’s producer, thus, ensuring that misleading reviews do not creep in is a necessity. The goal of this work is to develop a trustworthy reviews model by taking into consideration all the factors which make a review reliable.

**Keywords** Social networks · Feature extraction · Unsupervised clustering  
Trustworthiness · E-commerce

## 1 Introduction

With the evolution of Internet technology at a very fast pace, information about any product we wish to buy or any movie we are planning to watch is right at our fingertips. Social media platforms like Facebook have more traffic than search engines like Google. Thus, it is evident that we are moving toward a world which revolves around social interactions and information exchange. But what guarantees that this information is correct, unbiased and trustworthy? The economist Kenneth Arrow defines trust as a “lubricant of the social system.” Social trust is the essence of

---

R. S. Sukthanker (✉) · K. Saravanakumar  
VIT University, Vellore, India  
e-mail: sukthankerrhea.sanjay2014@vit.ac.in

K. Saravanakumar  
e-mail: ksaravanakumar@vit.ac.in

efficient functioning of communities which facilitates collaborative growth and self-enhancement. The field of trust analysis in social networking sites like Facebook, Twitter has been studied extensively in [1]. But as we evolve and let online vendors replace local stores, what actually matters now is the trustworthiness of product reviews. Online reviewing works like a breeding ground for public opinion and can influence business trends in a disguised manner [2]. We tend to trust the “word of mouth” phenomenon and the product reviews more than the information provided by vendors. After the age of bloggers [3], we have reached the age of reviewers; thus, people easily trust reviews of widely acknowledgeable reviewers. What further makes this trust issue more convoluted are the myriad factors which play a role in determining the trust. Some approaches to tackle this problem use graph-based approaches [4] to detect review spam. Another significant work in this area which uses the correlation between reviewer honesty and trustworthiness is [5]. Another factor which often goes unnoticed is that the text of the review also plays a critical role in the review helpfulness and trustworthiness. This paper will attempt to cluster similar reviewers by extrapolating from reviewer history, reviewer buying/reviewing behavior and count of useful words used in their reviews.

## 2 Related Work

The intricacies involved in the field of social trust analysis surfaced less than a decade ago when the web became a strong platform for opinion expression. There have been many significant approaches to tackle this issue which vary widely depending on the task or context in which they are used. One such example is given in [6] which extract social trust relationships between users on Twitter using factors such as influence, cohesion and the valence (sentiment) of the user in an unsupervised manner. Recently there has been an upsurge in the use of personalization in user recommendation which closely relates to our work. One of the pioneers in the area of social-context aware trust influence [7] considers both the participant’s personal characteristics as well as mutual relations for improved recommendation. Some other significant works are [8–10] which further establish that a person prefers recommendations from trusted friends. Sinha and Swearingen [11] and Bedi et al. [12] also demonstrated that given a choice between the recommendation from trusted friends and a recommendation system, the former is more preferred. These provide a firm background to establish the necessity and need for trust-based review recommendation.

One of the dataset widely used to mine trust relationships is the Epinions [13] web of trust dataset which is crawled from [14] and is available in many data repositories. Many researchers [15] use this dataset for mining trust relationships. Epinions dataset has been widely exploited using Bayesian analysis in [16] for trust aware recommendation and for rating prediction in [17]. The task of trustworthiness prediction becomes even more complex when a dataset like Amazon reviews [5] is used which is meant to function as a staging ground for businesses rather than opinions. So broadly speaking there are two types of datasets which are widely exploited in

this regards one which uses user–user and user–product rating like Epinions or Slash Dot and others like Amazon datasets which uses only user–product ratings in trust prediction. Both types of datasets have their own idiosyncrasies. Prediction accuracy in the former type of datasets can be tested, and the negative or positive links can be predicted [17]. The major issue faced in evaluating the trust in the latter is that the accuracy of predictions cannot be adequately supported. Some examples of earlier work in trust and helpfulness detection are discussed in [18].

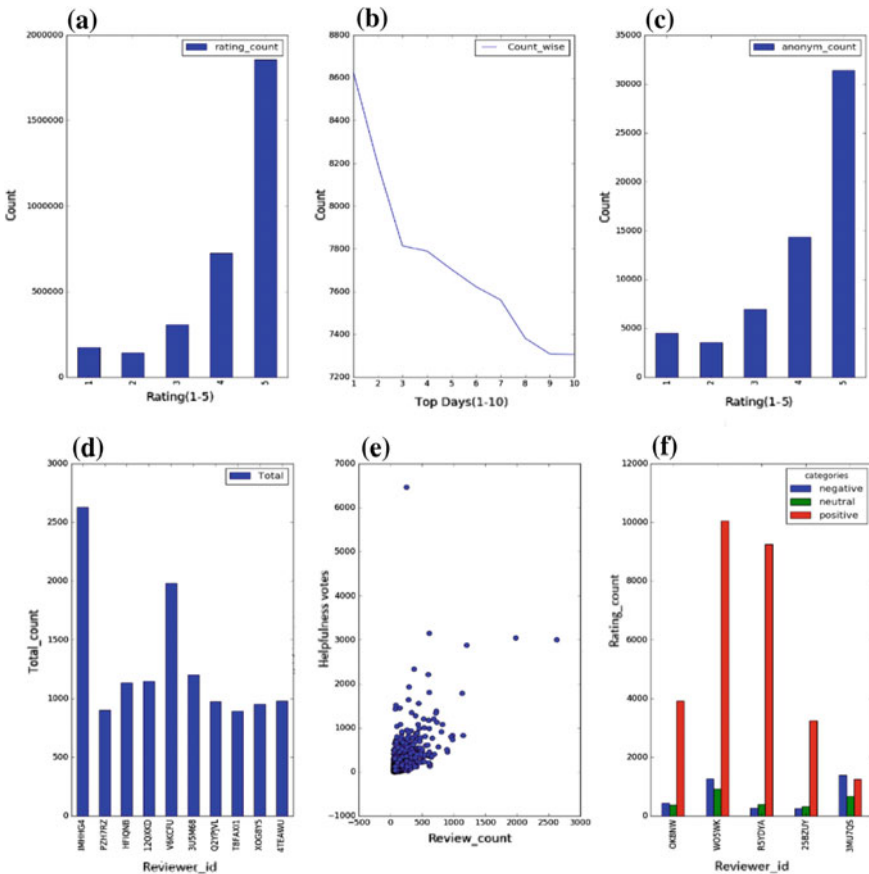
The idea for this project emerged from the analysis and discrimination of the earlier research work conducted in this area the most significant ones being [2] which explore how online review forums provide customers with powerful platforms to express opinions and influence business trends. This paper contributed toward creating a trustworthy co-created recommendation model. The foremost pioneers of this idea were Prahalad and Ramaswamy [18] who defined the concept of co-creation in customer communities. The first functionality of the review is derived from the aforementioned paper. Here the reviewer’s profile is unboxed to evaluate the trust metrics. The parameters used will be reviewer profile history, past helpful votes, reviewer rank, percentage of posts made, account activity, etc. Another paper which significantly inspired our work was the recent work of Wu et al. [19] on how credibility of an advisor is actually perceived in a marketplace. This motivated the second functionality of our project—the advisor segmentation phase.

### 3 Problem Statement

E-commerce Web sites have bought the world closer and have enabled businesses to make huge strides in their profit. What we often tend to ignore is another and perhaps the most significant influence of e-commerce Web sites. It has developed trust relationships among the customers and between the buyer and seller. E-commerce Web sites have heralded the era of mutual trust. Now the question which arises next is whom shall we trust and who can be trusted? Does the reviewer rank alone suffice to trust his/her review? The answer is no. Though reviewer’s rank does play a role in review helpfulness, it does not guarantee a frank and unbiased review. A number of attempts have been made toward the goal of developing a trusted e-commerce network. Most customers prefer to buy expensive electronics and delicate objects in person rather than buying them through e-commerce Web sites. This is because a certain factor of trust does influence buying decision making. Most of us do not trust the social web enough to risk critical buying decisions. This can be changed if we are able to identify and eliminate spammers and provide the vulnerable customer with only trusted reviews.

### 4 Exploratory Analysis on the SNAP Amazon Kindle Dataset: Getting to Know the Data

Figure 1 shows the exploratory analysis of the dataset. Each graph and the insights derived from it are discussed below. The plot (a) displays the number of reviews grouped by their rating. As clearly shown in Table 1, the web is overall positive. A customer does not give a negative review unless he has had a very bad experience with the product. Also the distribution is highly skewed. This means that the number of positives is much greater than the number of negatives. Notice that more than 50% of the reviews are extremely positive. This can mean either they are extremely



**Fig. 1** a Ratings from 1 to 5 and the count in the Kindle dataset. b The top 10 days when the maximum number of reviews were recorded. c Ratings with the anonymity count. d A subset of reviewers and their total review count. e The reviewer's review count and his total helpfulness. f Product id and their positive, negative and neutral review counts

**Table 1** Distribution of ratings in the dataset

Rating	Nature	Rating percentage
1	Extremely negative	5.3974
2	Negative	4.468179
3	Neutral	9.573706
4	Positive	22.647745
5	Extremely positive	57.912966

biased or the data crawled is for books that were highly acknowledgeable. Another insight is the number of reviews per day in Fig. 1b which shows a sudden hike in the number of reviewers given on the Web site during specific days. In the plot below, only the top 10 days with the highest number of reviews are considered. A huge hike in number of reviews on September 6, 2012 is noticed, i.e., 6619 recorded reviews could be possibly attributed to the release of Kindle Fire on the same day in Europe.

Another thing which is quickly revealed on exploring the dataset is that there are a significant proportion of reviews which are anonymous. As can be seen clearly from Fig. 1c, the numbers of anonymous reviewers giving a rating of 5 are large in number. Can anonymous reviews be trusted? The answer is mainly a no as no background information about the reviewer can be obtained. It can be possible that these anonymous reviewers are extremely biased friends of the book’s author or publisher. For the sake of the project, we will not be considering anonymous reviewers though they can be said to show unique patterns and deviations.

The graph in Fig. 1e shows the count of the number of votes per reviewer and the total helpfulness votes that particular reviewer received. There are some visible outliers and also a majority of reviewers with no helpfulness votes. On one hand, there exist some reviewers with relatively few reviews and abnormally large number of helpfulness votes and on the other hand there exist reviewers with a large number of reviews but very little helpfulness votes.

The graph in Fig. 1f shows the number of positives, neutral and negative reviews (y-axis) for a product (x-axis), and the following conditions are considered. A small subset of the products is displayed in the graph. The rating is discretized to categories negative, if ratings are 1 and 2, neutral, if rating is 3, and positive, if ratings are 4 and 5. Another factor considered for dataset exploration is the review count per user. Only some users have counts above 1200, and only two of them show deviations from others in the category showing exorbitantly high number of review count. These could be either spammers or extremely popular and active reviewers.

## 5 The Proposed Model

### 5.1 A Brief Overview

Figure 2 shows the concise model of the proposed approach. Each and every phase is described in detail below.

**Phase 1: Data Collection.** The data sources used are available at SNAP [20] repository, which is the data repository for many large social network datasets. We use the data crawled by jmcauley [21] from the Amazon Web site, which consists of review dataset of a number of product categories. As shown in Fig. 2b, we have considered Amazon Kindle dataset as our major dataset and others like baby products, pets, sports, beauty, health and cell phones as our assistive datasets. The assistive datasets will be majorly used for comparison of reviewer helpfulness across myriad categories. These categories are used to measure the competence of the reviewer and his expertise in a particular category. The main dataset is the Kindle dataset on which trustworthiness is to be predicted.

**Phase 2: Handling Massive Dataset.** The dataset of kindle reviews is split into multiple parts and stored in a compressed format called pickles in python for faster processing. The JSON files are converted to Python pandas dataframe for the analysis purpose.

**Phase 3: Data Cleaning and Transformation.** The next phase is the phase of data preprocessing. In this phase, the following two major subtasks are performed:

- (a) *Data Cleaning.* In this phase, the attributes irrelevant to the study like the reviewTime, reviewerName are eliminated from the dataframe. The dataframe is checked to ensure that any of the attributes considered for the study are not missing, and any noisy attributes are eliminated. The review text is cleaned

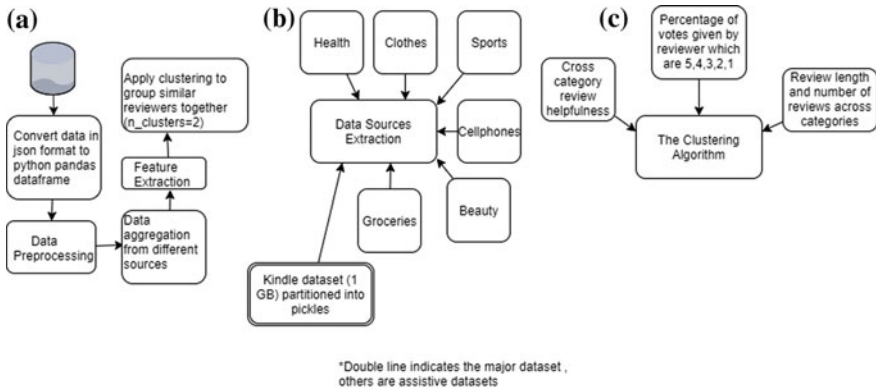


Fig. 2 a Overall architecture. b Data aggregation phase. c The clustering phase

by tokenizing it, eliminating punctuations and eliminating stopwords. Only the feature of interest, i.e., the word count of the review is retained.

- (b) *Data Transformation*. In this stage, a number of transformations are applied on the dataframe attributes. Data transformation stage is crucial for extracting useful features from the dataframe. Details of the feature extraction phase are given in Sect. 5.2.

**Phase 4: Clustering**. In this phase, the reviewers with similar online behavior and reputation are grouped into clusters using K-means algorithm. The algorithm is discussed further in Sect. 5.3.

## 5.2 Feature Extraction

Real-world datasets like the one used in our study are composed of redundant attributes, all of which may not be particularly useful toward achieving the goal. Feature selection is a phase in which the attributes providing obsolete or redundant information are eliminated from the dataset. Feature extraction on the other hand is related to dimensionality reduction; it starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant. The Amazon review dataset has a very high dimensionality. Thus, the use of feature extraction to extract most relevant features is both inevitable and necessary.

**Feature 1: Weighted Helpfulness in All Categories**. The formula below calculates the weighted helpfulness of reviewer over different categories. It is to assign weight depending on the total number of helpfulness votes.

$$WH_i = \alpha * \sum_{j=1}^n \frac{H_j}{T_j} \quad (1)$$

where  $WH_i$  is the weighted helpfulness for reviewer  $i$  and  $H_j$  is the total helpfulness votes for review  $j$  and  $T_j$  is the total views for review  $j$  including both instances when it was voted and not voted.  $\alpha$  is the weight associated with the reviewer. This scheme of weighting is used to ensure that uninformative data, e.g., 1/1 helpfulness which corresponds to 100% helpfulness is handled appropriately (Table 2).

**Feature 2: Count of Useful Words Used**. Average of the review word count of all the reviews for a particular reviewer is taken and is exploited as a useful feature. Tokenize function tokenizes the text. Eliminate punctuations removes the punctuation marks. Eliminate stopwords removes stopwords like and, or, to.

$$AWC_i = \sum_{j=1}^n wci / length(RTi) \quad (2)$$

**Table 2** Deciding value of  $\alpha$  for reviewer  $i$

$T_j$	$\alpha$
$T_j > 1000$	1
$500 < T_j \leq 1000$	0.75
$100 < T_j \leq 500$	0.5
$50 < T_j \leq 100$	0.25
$T_j > 10$	0.1

where  $AWC_i$  is the average word count of reviewer  $i$ ,  $w_c$  is the total number of useful words in review  $j$  and  $RT_i$  is the total number of reviews by reviewer  $i$ .

**Feature 3: Percentage of Ratings.** Percentage of ratings 1 through 5 for each reviewer’s comments is calculated as follows:

$$PR_{ij} = \frac{RAT_i}{T_j} \tag{3}$$

where  $PR_{ij}$  is the percentage of rating  $i$  for reviewer  $j$ ,  $RAT_i$  is the total number of ratings for rating  $i$  ( $i = 1-5$ ), and  $T_j$  is the total number of reviews for reviewer  $j$ .

We classify the type of reviewers in the social e-commerce network as follows:

- (1) *The Popular ones (feature: Popularity).* These are the reviewers who are highly appreciated by other customers. Their typical characteristics are high number of helpful votes in a particular category as compared to other categories. Popularity is calculated as follows:

$$Popularity_i = \sum_i^{Num\_cat} \alpha * WH_i / Num\_cat \tag{4}$$

where  $WH_i$  is weighted helpfulness in category  $i$  (whose weighted helpfulness is non zero) and  $Num\_cat$  is the number of categories in which  $WH_i$  is not zero,  $\alpha$  the weight metric is 1 when the category is books and  $\alpha$  is 0.5 otherwise.

- (2) *The Unbiased (fair) ones (feature: Fairness).*

$$Fairness_i = -0.75 * P5_i - 0.5 * P4_i + 0.5 * P3_i - 0.5 * P2_i - 0.5 * P1_i \tag{5}$$

where  $P5_i, P4_i, P3_i, P2_i, P1_i$  are percentage ratings for reviewer  $i$  calculated earlier. This weighting scheme is used to assign a highly negative value to reviewers with a large percentage of highly negative or positive rating. Here we use modified version of the method proposed in [2] for universal applicability. Above equation assigns a high negative weight to percentage of rating 5 as reviewers with high percentage of highly positive reviews are not fair and a slightly lower negative weight to percentage of rating 4. Percentage of neutral



is added unchanged, and weighing scheme similar to 5 and 4 is followed for percentage of 1 and 2 ratings, respectively.

Five rating is extremely positive; thus, reviewers with a very high percentage of reviews as positive are not trustworthy. Reviewers with a high percentage of 4 ratings are moderately trustworthy, while those with a high percentage of neutral reviews are considerably trustworthy. Also extremely negative reviewers with percentage of negative ratings very high are highly not trustworthy, and similar behavior corresponds to reviewers with high number of 2 star ratings.

- (3) *The Experts (feature: Expertise)*. These are typically the domain experts of the category (e.g., voracious readers in the kindle books categories). They typically write long reviews and are devoted to reviewing a particular category only. This particular feature uses cross-category comparison. According to the Merriam dictionary [21], an expert is a person having, involving or displaying special skill or knowledge derived from training or experience. Thus, experience of a reviewer reflects how much he is trusted in the community. Some reviewers are as the saying goes “jacks of many and masters of none.” Thus, it is necessary to extract the expertise as a feature (Table 3).

$$E_i = [(bc_i) - [(bbc_i) + (byc_i) + (spc_i) + (cpc_i) + (htc_i) + (ptc_i)]] * (awc_i) \quad (6)$$

where  $E_i$  specifies expertise of particular reviewer. In the above equation, the difference between the total review count of book category and sum of total review count of other categories is taken mainly to detect the distribution of the reviewer’s reviews over different amazon categories. For an expert reviewer of the Kindle category, this difference will be significantly high. This value is later multiplied with the average word count of the reviewer. Thus, for an expert reviewer, this equation yields a very high value. All the features above are normalized for better visualization of clusters during clustering.

**Table 3** Abbreviations used

Feature	Abbreviation used
Book category review count	bc
Baby category review count	bbc
Beauty category review count	byc
Sports category review count	spc
Health category review count	htc
Pets category review count	ptc
Cell phones category review count	cpc
Average word count	awc

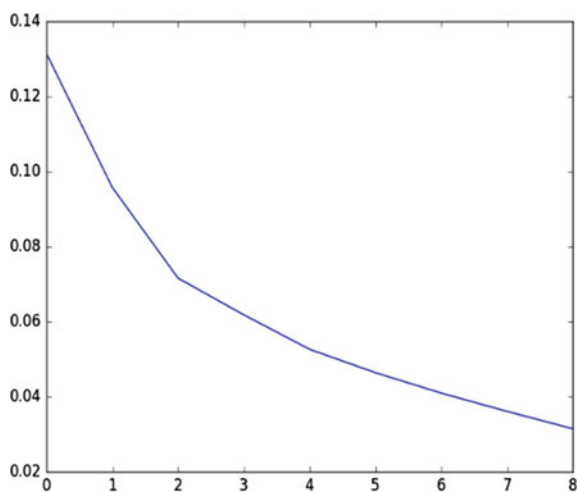
### 5.3 Clustering Phase and Results

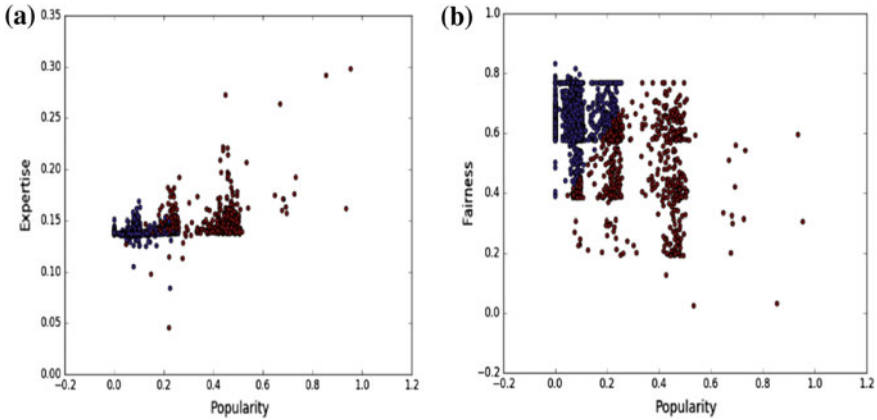
As it is obvious that the detection of trustworthy reviews on an e-commerce platform like Amazon wherein trust relations are not explicitly known, is an unsupervised type of problem. K-Means clustering is the type of clustering in which the items in the dataset which are closely related or whose behavior is similar are said to belong to a particular cluster. K-medoids is also widely used as a clustering metric but here we refrain from using K-medoids mainly because the time complexity is  $O(n^2 * k * i)$ , whereas K-Means runs in  $O(n * k * i)$ . Our dataset spans over a large number of reviewers; thus, in this work, we use K-Means [22] to find the reviewers whose behavior closely resembles each other. The type of clustering performed is 3-D clustering based on the three types of personality traits, i.e., “Popularity”, “Fairness”, “Expertise”, while deciding the number of clusters, the elbow method is used as shown in the graph below. Therefore, the number of clusters to be selected is 2 (Figs. 3, 4).

In K-means, the labels of the clusters are not known. Thus, explicit segregation of trustworthy and not trustworthy reviewers is not possible. This is an issue faced by majority of the unsupervised learning algorithms. Domain-specific knowledge can be incorporated in this study to make the identification of trustworthy reviewers easier (Table 4).

After applying domain knowledge to cluster, the reviewers based on the 3 parameters are categorized into two clusters: the cluster with label as 1 is found to be the group of trustworthy reviewers. The ranks of the labeled trustworthy reviewers were used to determine the prediction accuracy of the algorithm. The ranks were obtained by manually inspecting the rank the reviewers in the sample are assigned by Amazon. A small sample of 67 reviewers shows the results as tabulated above. Thus, analyzing the results obtained the following insights into reviewer trustworthiness:

**Fig. 3** Elbow plot to decide optimum number of clusters





**Fig. 4** **a** Clustering on basis of popularity and fairness. **b** Clustering on basis of popularity and fairness

**Table 4** Percentage of trustworthy reviewers per category in a given sample

Reviewer ranks (sample)	% of trustworthy sample they compose (%)
Rank < 1000	17.91
1000 ≤ Rank < 2000	25.37
2000 ≤ Rank < 3000	11.94
3000 ≤ Rank < 4000	14.92
4000 ≤ Rank < 5000	7.46
5000 ≤ Rank < 6000	7.46
6000 ≤ Rank < 7000	1.49
7000 ≤ Rank < 8000	1.49
8000 ≤ Rank < 9000	2.98
9000 ≤ Rank ≤ 10,000	1.49
Rank > 10,000	7.46

- (1) The majority of the trustworthy reviewers are neither the ones ranked at the top nor the unranked ones but those who lie somewhere between these 2 extremities.
- (2) Though the distribution of trustworthy reviewers does not show a very even pattern in general, trust decreases as the reviewer rank increases, i.e., the probability that a reviewer with rank  $x$  greater than  $y$  is generally less trustworthy.

## 6 Future Work and Conclusion

An in-depth study in the field of social trust analysis defines several possible future extensions to our work. This study does not take into consideration many of the crucial factors which can to a large extent determine the trustworthiness of the reviewer. The

first area of work can be in studying the reviewing behavior of the reviewer. A customer who reviews related products is much more reliable than a customer who reviews products in random categories and significantly unrelated products. Another potential work can be a more detailed analysis of the review text, which tells us a lot about reviewer's personality traits and reliability. First phase of this particular work can be to extract the category relevant product aspect extraction [23], e.g., in books category story, characters, suspense, motivation. The polarity of a review also need to be evaluated carefully to test whether a review is positive or negative [24]. Further the synset may need to be generated to identify similar aspects, e.g., story, plot. Thus, if a customer's review speaks about large number of aspects of the product, the reviewer is considered more trustworthy. Also there are other domains like loyalty in social media [25] which provide scope for a comparative analysis of trust and loyalty.

Thus, this work makes an attempt to identify the trustworthy reviewers in large social networks like Amazon, where users do not explicitly have any trust relations. This is a typical problem of unsupervised pattern recognition. From studying a sample of reviewers clustered on basis of similarity based on their trustworthiness level, it is evident that though highly ranked reviewers are generally trustworthy, this assumption may not always be true. Thus, discovering more accurate and better approaches for social trust analysis is definitely the need of the hour.

## References

1. Chen, L., Tsoi, H.K.: Privacy concern and trust in using social network sites: a comparison between french and chinese users. In: IFIP Conference on Human-Computer Interaction, pp. 234–241. Springer, Berlin, Heidelberg (2011)
2. Li, S.T., Pham, T.T., Chuang, H.C., Wang, Z.W.: Does reliable information matter? towards a trustworthy co-created recommendation model by mining unboxing reviews. *ISEB* **14**(1), 71–99 (2016)
3. Sunny, S., Divya, M., Rachna, M., Soorea, L., Revathi, C., Saravanakumar, K.: Recommendation of blogs in E-learning. *Int. J. Eng. Technol.* **5**(3), 2515–2518
4. Wang, G., Xie, S., Liu, B., Philip, S.Y.: Review graph based online store review spammer detection. In: 2011 IEEE 11th International Conference on Data Mining (icdm), pp. 1242–1247. IEEE (2011)
5. Shinzaki, D., Stuckman, K., Yates, R.: Trust and Helpfulness in Amazon Reviews: Final Report (2013)
6. Vedula, N., Parthasarathy, S., Shalin, V.L.: Predicting Trust Relations Among Users in a Social Network: On the Role of Influence, Cohesion and Valence
7. Wang, Y., Li, L., Liu, G.: Social context-aware trust inference for trust enhancement in social network based recommendations on service providers. *World Wide Web* **18**(1), 159–184 (2015)
8. Berscheid, E., Reis, H.T.: *Attraction and Close Relationships* (1998)
9. Fiske, S.T.: *Social Beings: Core Motives in Social Psychology*. Wiley (2009)
10. Yaniv, I.: Receiving other people's advice: influence and benefit. *Organ. Behav. Hum. Decis. Process.* **93**(1), 1–13 (2004)
11. Sinha, R.R., Swearingen, K.: Comparing recommendations made by online systems and friends. In: DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries, vol. 106 (2001)
12. Bedi, P., Kaur, H., Marwaha, S.: Trust based recommender system for semantic web. *IJCAI* **7**, 2677–2682 (2007)

13. Unbiased reviews by real people. [www.Epinions.com](http://www.Epinions.com)
14. Zhang, Y., Yu, T.: Mining trust relationships from online social networks. *J. Comput. Sci. Technol.* **27**(3), 492–505 (2012)
15. Guo, L., Ma, J., Jiang, H.R., Chen, Z.M., Xing, C.M.: Social trust aware item recommendation for implicit feedback. *J. Comput. Sci. Technol.* **30**(5), 1039–1053 (2015)
16. Yuan, W., Guan, D., Lee, Y.K.: The small-world trust network. *Appl. Intell.* **35**(3), 399–410 (2010)
17. Leskovec, J., Huttenlocher, D., Kleinberg, J.: Predicting positive and negative links in online social networks. In: *Proceedings of the 19th International Conference on World Wide Web*, pp. 641–650. ACM (2010)
18. Prahalad, C.K., Ramaswamy, V.: Co-creating unique value with customers. *Strateg. Leadersh.* **32**(3), 4–9 (2004)
19. Wu, K., Noorian, Z., Vassileva, J., Adaji, I.: How buyers perceive the credibility of advisors in online marketplace: review balance, review count and misattribution. *J. Trust Manag.* **2**(1), 2 (2015)
20. Leskovec, J., Krevl, A.: SNAP Datasets: Stanford Large Network Dataset Collection, 2011. <http://snap.stanford.edu/data/index.html> (2014)
21. He, R., McAuley, J.: Ups and downs: modeling the visual evolution of fashion trends with one-class collaborative filtering. In: *Proceedings of the 25th International Conference on World Wide Web*, pp. 507–517. IW3C Steering Committee (2016)
22. Kanungo, T., et al.: An efficient k-means clustering algorithm: analysis and implementation. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 881–892 (2002)
23. Poria, S., Cambria, E., Ku, L.W., Gui, C., Gelbukh, A.: A rule-based approach to aspect extraction from product reviews. In: *Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP)*, pp. 28–37 (2014)
24. Banan, T., Sekar, S., Mohan, J.N., Shanthakumar, P., Saravanakumar, K.: Analysis of student feedback by ranking the polarities. In: *Proceedings of the Second International Conference on Computer and Communication Technologies*, pp. 203–214 (2016)
25. Hamilton, W., Zhang, J., Danescu-Niculescu-Mizil, C., Jurafsky, D., Leskovec, J.: Loyalty in online communities. In: *AAAI International Conference on Weblogs and Social Media (ICWSM)* (2017)