

A survey of the correlation between trust and reputation in electronic commerce B2C using Alexa rank

Issa NAJAFI

Computer Engineering Department, Quchan University of Advanced Technology, Quchan, Iran

najafy@qiet.ac.ir, issa.najafi@gmail.com

Abstract

In general approaches of electronic trust in B2C e-commerce, to facilitate purchase decision process, the customers select valid brands based on the authenticity of web and regulatory mechanisms. This approach is not suitable for purchase from small companies without active brands in e-commerce. To purchase from these companies, we should follow other e-trust approaches. One of the approaches applied by information literate people is referring to valid ranking sites like Alexa rank. This study evaluates the correlation between Alexa rank and formal measures of trust in electronic space and extracts a new method for e-trust based on electronic reputation extracted from the results of existing ranking in ranking sites. To do this, 100 store websites are selected based on regulatory lists and the formal criteria of trust is measured and the correlation of these criteria with Alexa rank is compared.

Keywords: Trust, Reputation, Correlation, B2C E-Commerce, E-Transaction, Trust Publishing, Alexa rank

1. Introduction

The increase of online trade volume makes the exact measurement on this field as necessary. Different types of frauds are occurred in this field. In customer-business trade form, as the customer has no access to the physical location of business, he cannot be sure that such business exists. In different environments, it is attempted to create regulatory processes to avoid fraud. For example, in Iran, the sign of e-trust affiliated to the ministry of industry, mine and trade is used to determine the e-trust to a website. This process has some drawbacks as:

- The companies receiving the mentioned sign and provide services to smaller companies, take the risk of fraud and they can use payment services for small business by receiving a part of transaction price.
- In an open global space, non-local businesses cannot be forced to observe local regulatory rules.
- In this approach, only the fraud is considered but often, the dissatisfaction is mostly due to the difference between the customer expectation and the purchased goods. In other

words, after being assured of non-fraud, the main barrier to an electronic purchase is assurance of the quality of product.

All these factors define that by similar mechanisms to e-trust sign, we cannot create a good space for all businesses. In other words, the challenge of the significance of important brands and the discriminatory nature of this space is problematic.

From another view except B2C, B2B interactions exist. In these interactions, the totally mechanized interactions are considered. In other words, the software should be integrated to do financial exchanges. Under such condition, there should be fully mechanized methods by which we can measure the reliability. This study aims to study the first set of trade and it measures the correlation between reputation and trust metrics.

2. Review of literature

The investigations based on resources, field online studies in business space, extraction of information from interview with experts show that reliability in electronic trade is a basic challenge as sometimes this leads to the lack of intention of citizens to e-purchase. In recent decades, various studies have been performed to increase trust in online environment but the development of using e-commerce in the target community of citizens and companies in comparison to other target communities as online social networks is not universal. This study attempts to take another step but not the last step to create online trust by the ranking results in internet ranking sites namely Alexa.

3. Review of literature, conceptual model, theoretical basics of measuring e-trust

Various measures are proposed to measure trust. A review of criteria is performed in this section. These measures are mostly based on the opinion of users regarding different sites. In other words, trust of a website is measured by the view of other users. Various considerations including local view are used to measure trust. In the general model, we can consider the following view for trust calculation models:

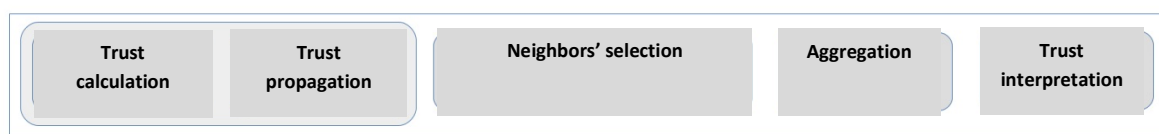


Figure 1: The process of trust calculation based on the users opinion

To achieve neighbors, we can make this division based on the interest of users. In other words, the users with similar interests are considered as neighbors. To avoid attack in these systems, we can use the combination of similarity of interests and the trust of a user to another user [1,2,3]. In other words, each user determines a degree of trust for another one as applied to compute the neighbors of the same user.

After the selection of neighbors, we deal with aggregation. To do this, we should select a good aggregation function. One of the most common aggregation functions is Resnick aggregation function [4] as follows.

$$\hat{r}_{Resnick}(u, i) = \bar{r}_u + \frac{\sum_{v \in Neighbours\ of\ u} (r(v, i) - \bar{r}_v) W(u, v)}{\sum |W(u, v)|}$$

Where, \hat{r} is predicted rate and r is actual rate to an entity. Also, \bar{r} is the average previous rate of user regarding different entities and finally, $W(u, v)$ is the view of user u to v . In the final step, we should make interpretation for the calculated values regarding the trust of a user to different entities. In most cases, the values are not used directly and should be interpreted. This function is implemented as repeating. In other words, in different methods, these calculations are continued until all trust values are converged. One of the first methods is Advogato [5]. This method determines the parameter of importance of users to each other view. In other words, this method determines a reliable value for each user. To achieve this value, at first graph of users relation is made. Then, by an initial set of trustworthy users, the rest is rated based on the shortest path to one of the initial sets. For those with equal distance to the initial set, one's output degree in graph is involved in the trust metrics. One of the presented methods is AppleSeed [6]. In this method, the set of the closest neighbors is determined for each user. Like Advogato method, at first the trust graph of users to each other is made. Then, in each step, one of the users is considered as source node and the initial energy is propagated based on the trust weight of users to each other in the graph. The calculated energy values for each user in each step are equal to the trust of the source user to that user. We can write the calculation method as:

$$Trust^{AppleSeed}(s, u) = \sum_{v \in incoming\ edges(u)} eng_s(v, u)$$

Where $eng_s(v, u)$ is the propagated energy from user v to u , where user s is selected as the source user. One of the methods is TidalTrust [7]. In this method, for both users like u, v , all paths are considered. For each path, the lowest weight of edges is considered and the highest value is considered between these minimum values. For both users, the highest lowest weight of edge is considered for different paths and is shown by Smaxmin. The paths that will be considered in the trust

Computation are only the paths from with a minimum of the edge weights that is not less

than Smaxmin. Let T^s be all users of s with the above conditions and $Trust^{Explicit}(s, v)$ be explicit trust between s and v . By these values, trust in this method is computed as:

SUNNY method is similar to TidalTrust [8]. In this method, for a source node as S and destination node as B , if all neighbors B are unknown", at first by a Bayesian network, the initial

value of trust between S,B is computed, then each of neighbors tries to change their state to include and again calculates confidence by a Bayesian network. If the resulting confidence bounds are too different from the previous confidence bounds, then node is changed to exclude. Finally, only the neighbors of B that do not change the original confidence bounds too much are selected. This is continued until the conditions of all nodes are defined. Finally, based on all neighbors, we can define the trust to entities.

MoleTrust is similar to Tidal [9]. MoleTrust uses breadth-first search. In this approach, all raters will be considered up to the maximum-depth. The following term can show the behavior of this method well. In this method, only the edges are considered that source node in the edge has extremely low computed trust (the parameter of method).

$$Trust^{TidalTrust}(s,u) = \frac{\sum (Trust^{Explicit}(s,v) \times Trust^{TidalTrust}(v,u))}{\sum (Trust^{Explicit}(s,v))}$$

Another method on [10] is called TaRS and is similar to ModelTrust. With the difference that besides trust applies the similarity of the interest of users to compute the trust between two users. In the presented method in [11], trust is defined from two types. The first type of trust is item-level, it is assumed that each about each item has a set of trust recommendations. The set of recommendations is divided into correct and incorrect and trust is computed as follows:

$$Trust^{Item}(u,i) = \frac{|Predicts^{Correct}(u,i)|}{|Predicts(u,i)|}$$

The second type is profile-level trust and for each user, ratio of correct and incorrect predictions is independent of items. The values of these two trusts are combined well with the similarity of item with the interest of users and have achieved good results.

In [12], a method is presented similar to the previous method but besides the trust to the item, the trust to topic is computed. In other words, if we assume that besides items, we have different meaning fields and each of items belongs to one or some topic range, then each meaning field has trust metric. This metric is computed as follows for each meaning:

$$Trust^{Topic}(u,t) = \frac{\sum_{i \in Items(t)} Trust^{Item}(u,i)}{|Items(t)|}$$

Where trust to item is computed based on the following term:

$$Trust^{Item}(u,i) = \frac{\sum_{v \in Raters(i)} (1 - \frac{|A(u,i,v) - r(u,i)|}{S})}{|Raters(i)|}$$

Where $\hat{r}(u, i, v)$ is the predicted trust of u to i estimated by v. S represents the rating scale.

As similar method to AppleSeedin [3] as TrustWalker is presented. Like other methods, in this method the graph of users is made and then by a random walk on graph, the trust metric is measured. If there is an edge between two u,v users in the graph, the trust of u to v is as:

$$Trust^{TrustWalker}(u, v) = 1/|out(u)|$$

Where out(u) is set of direct friends of given user. In this approach, to predict the rate of user u about item i, if direct ran is given, the same rank is used. Otherwise, if the direct neighbors have given the rank, it is used. If neither the user nor his friends gives any rank to the item, this method is continued about the user friends. To select a friend to be asked, we consider the similarity of items as rated by him and the given item. The similarity between two items is calculated as:

Given that size of the set of common users is also important Jamali and Ester proposed a new

$$similarity(i, j)^{SPCC} = similarity(i, j)^{PCC} \cdot \frac{1}{1 + \exp(-\frac{|CU_{i,j}|}{2})}$$

Where, $CU_{i,j}$ is the number of users who have rated to both I,j, $similarity(i, j)^{PCC}$ is the similarity of two items based on Pearson Correlation Coefficient .

In another method as Mwalker [2] for an item i and user u, if the user has not rated, Top-K most trustworthy users compose the candidate set. If the neighbors rated to i, then returns it directly. Otherwise the user returns his rate about item j that is the most similar item to i. This algorithm continues for all neighbors users till maximum defined as depth. The trust between two users is computed as:

$$Trust^{MWalker}(u, v) = Trust^{Explicit}(u, v) \times Similarity^{Tag}(u, v)$$

In SimTrust, it is assumed that for each item i and user u, there is a text description u regarding i [13].

Thus, the key set in each test is extracted by tf-idf and the trust between two users is computed as:

Where, w denotes set of all explanations of users u,v about different items. Also, n_{uv}^k is the number of explanations with the word k and presented for a common item and n_v^k is the number of explanations of user v with word k.

In Merge, if the user u rated directly the item i , therefore previous rate will be used [1]. Otherwise, aggregation of rates of direct friends will be used. These views are aggregated as linearly and based on the trust between two users and the trust between two users is computed as:

$$W(u, v) = (\alpha \times \text{Similarity}^{\text{Rating}}(u, v)) + (\beta \times \text{Trust}^{\text{MoleTrust}}(u, v)) + ((1 - \alpha - \beta) \times \text{Similarity}^{\text{Social}}(u, v))$$

In RN [14], it is shown that, we cannot find a relationship between the similarity of two users and items by which they are rated. We should use both parameters in trust propagation and using one of them as the example of both of them cannot produce good results. In this method, all neighborhood edges in which the similarity of two users is lower than the limit, are removed and in trust propagation, a coefficient is multiplied by the calculated trust by neighbors as $(d - n + 1)/d$. Where, d is the maximum propagation distance and n the distance of the source user from the destination user.

In AgeTrust [15], it is shown that older friends are more reliable than new one. The following formula is presented to determine the trust between two users. P is rank v among friends u in terms of time in friendship and n is the number of direct friends of u and α is the initial trust to a new friend.

$$W(u, v) = \text{Similarity}(u, v) \times \text{Trust}(u, v) \times \left(1 - \left(p \times \frac{1 - \alpha}{n - 1}\right)\right)$$

In this section, different methods are reviews to calculate trust. These methods are based on this assumption that a set of users are related and each one has different views about the different items. To measure the trust of a user to item without any view regarding it, trust propagation methods regarding the trust of two users are used. This study aims to measure correlation of reputation with trust metrics. As the reviewed criteria are based on direct views of users. It seems that views are based on reputation. Here, we mention two limitations in trust metrics. Two main limitations in the real world are cold-start and sparse matrix of the relationship between users with each other and rating of users to items. In the real world, in most conditions, the initial data are low and the mentioned methods show an approximate of the trust to item. Under these conditions, evaluation of proposed values with reputation and its correction can be of great importance. In other words, it should be evaluated how much reputation can show trust. The correlation of these two criteria can show the rate of validity of different items in comparison with each other. The reputation criterion in websites is based on their visit frequency and Alexa rate is a good example.

Later, the correlation test is explained. In this section, the set of learning data and test data set are made.

4. Preparation of test

As it was said, the aim of this section and the next section is the evaluation of the correlation between trust and reputation and creating required conditions of the test. To do this, the following steps are taken:

- Selection of target websites
- Selection of rating reference people
- Establish relationship among people
- Create conditions of rating to the items similar to real conditions
- Select correlation criterion
- Select trust prediction method

In the first step, it is required to create a set of websites as items in which trust is computed. Based on the goal of this study as the evaluation of the correlation between reputation rate and trust, it seems that the required websites shouldn't have high reputation. In other words, these websites shouldn't be recognized as a source in e-commerce. The reputation and trust to such websites have high correlation. Thus, based on the registered websites in e-commerce in the e-trust website, 100 websites with Alexa rank ranging 10000 to 100000 are selected. These websites are selected randomly among the given websites. Some examples of selected websites are as:

www.plus.ir	www.netshahr.com	www.netbarg.com
www.modiseh.com	www.mihanstore.net	www.shixon.com
www.bigtheme.ir	www.mihanmarket.com	www.goldtag.net
www.zanbil.ir	www.jeeb.ir	www.faranesh.com
www.ponisha.ir	www.kimiaonline.com	www.alldigital.ir

Experts in rating are fully aware of internet and e-commerce and they have at least 10 e-purchases and they can evaluate the trust of a website. To establish communication, they are asked to present a CV and the communication is established based on the similarity of the interest of these people with each other. To do this, based on cosines distance between the vector of CV as made based on tfidf, this task is performed.

To create similar conditions with real conditions, it is attempted to use the standard data set with the same features. One of the existing standard data is Epinion dataset as published by Stanford University [16]. In these data sets, the average output degree of each node is $\log_2 |v|$ in which v is the set of graph nodes. The graph is connected and its diameter is $\log_2 |v|$. In the applied dataset, it is attempted to use features similar to Epinion graph. The users are asked to define their trust rate regarding other users and websites. Then, a graph with the above features is extracted and this set is tested as training data set and the rest is used as reference of test.

The applied criteria in determining trust are classified into three groups. The first group includes the values ranging 0-5 and second type is regarding yes or no questions. The third type is dedicated to values 0-20.

First type, the questions or criteria of trust assessment (ranging 0-5):

- The investigation of the validity and response of telephone numbers registered in the website
- The evaluation of the validity of E-mails in the website
- The investigation of the response of contact forms with website
- The complete information regarding the features of website
- The exact map of site and guidance about purchase method
- Advertisement in valid internet websites
- Complete features about each product
- Comparison of products with each other
- Variety of goods in each classification
- Complete explanation of all purchase costs via the relevant website
- Complete explanation of goods send back method
- Exact follow up of order state
- Validity of the websites with links to this website
- Correct classification of products

Second type of questions (yes or no):

- Evaluate the identity of accounts owners and consistency with the owners of store
- Evaluate the identity and contact numbers as the owner of internet address
- Present invoice and pre-invoice
- Determine the after sale and support services
- Using safe communication protocols

Third type of questions or trust assessment criterion (0-20):

- Your trust based on the previous feedbacks to the evaluated website

After changing the scale of responses in each criterion, all the values of criteria are added and are dedicated as trust criterion by a person to the given website. Then, the trust to website is computed based on one of the mentioned methods in the previous section and then, the correlation with reputation is calculated. Later, we can explain about the selected function to calculate trust.

For correlation criterion, Copula Function is used. These functions have special features showing different aspects of correlation. Simply, Copula Functions are functions connecting marginal distribution functions of some random variables. This definition cannot show the features of Copula Function. We can say for two random variables as u, v with distribution functions as $F(u)$, $G(v)$ and binominal distribution function $H(u, v)$, we can show a mapping of 2-D space of $[0, 1] \times [0, 1]$ to $[0, 1]$ as the mapping between the marginal distribution functions of these two variables and their binominal distribution function. We can say this mapping is a function and

such functions are called Copula Function. The most important feature of Copula Function is their independence from the type of distribution of marginal functions. This feature causes that Copula Functions are good instruments to be used in heterogeneous spaces as web. One of the important features of Copula Function is to identify dependence in occurrence or non-occurrence of random variables. In other words, if two random variables have similar events, they are identified as two dependent variables. Even if the similar non-occurrences can show the similarity of two variables, various general copula functions are presented and in this study, Gaussian function is used. The following shows a general view of the formula of this Copula Function.

$$\int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi(1-\theta^2)^{\frac{1}{2}}} \times e^{-\frac{(s^2-2\theta st+t^2)}{2(1-\theta^2)}} ds dt$$

Where, θ is the input parameter. By this Copula Function, we can calculate correlation as follows:

$$\int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) du_1 du_2$$

The method to determine trust is MoleTrust. This method is selected based on the problem conditions and good capability. This method can deal with the sparse data and based on trust propagation stage, this method can solve this problem. Based on the weakness of applied data set regarding meta data and not using them in this method, this method can fulfill the conditions of test.

5. Evaluation of the correlation between reputation (reputation rank) and trust

This section evaluates the correlation of reputation and trust. To achieve this correlation, at first the calculated reputation and trust is normalized. The reputation values as being the rate of reputation have inverse trend in Alexa. The higher the reputation, the smaller than reputation rate in Alexa or any other rating site. Thus, based on zero average I normal values, by multiplication of rate values by -1, we can make these values similar to the reputation value. Then,, these values are observed in the following chart.

As shown in the Chart, the correlation between trust and reputation is revealed. To be sure of this correlation, Gaussian function is computed for it. To do this, at first the values should have scale change. The Copula Function is applied on probable values and it is ranging 0-1. After making these changes, the calculated correlation is 0.9123 and it shows high correlation.

6. Conclusion

This paper evaluates the correlation of reputation (reputation rate) and validity criteria. To do this, a set of questions is designed and they by the data of a set of websites and people, the rating

is performed based on the results of these questions. People are connected based on the CV similarities and a graph is formed. Finally, by MoleTrust, final trust values for each website are calculated.

The results are normalized and then by Copula Function, correlation is computed. The set of these results and data show that correlation between reputation and trust in websites with acceptable value of reputation is high. Thus, people can trust in reputation of E-commerce website and start e-purchase. We can criticize also reputation criterion or sometimes reputation of a website is high but there is not trust.

Based on the general correlation between reputation and trust, we can use this criterion in future as a measure to determine the significance of each of trust criteria. In other words, in providing a good questionnaire to measure trust, we can evaluate the correlation of the given criterion with reputation in the famous websites and determine the weight of relevant criterion in trust measurement.

References

- [1] G. Guo, J. Zhang, and D. Thalmann, "Merging trust in collaborative filtering to alleviate data sparsity and cold start," *Knowledge-Based Syst*, Dec. 2013.
- [2] J. Jin and Q. Chen, "A trust-based Top-K recommender system using social tagging network," in *Proceedings of the 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012)*, 2012, no. Fskd, pp. 1270–1274.
- [3] M. Jamali and M. Ester, "TrustWalker: a random walk model for combining trust-based and item-based recommendation," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 397–406.
- [4] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work - CSCW '94*, 1994, pp. 175–186.
- [5] R. Levien, "Attack resistant trust metrics," UC Berkeley, 2004.
- [6] C. Ziegler, "Towards decentralized recommender systems," 2005.
- [7] J. Golbeck, "Generating predictive movie recommendations from trust in social networks," in *proceedings of the 4th international conference on Trust Management (Trust'06)*, 2006, pp. 93–104.
- [8] U. Kuter and J. Golbeck, "Sunny: A new algorithm for trust inference in social networks using probabilistic confidence models," in *AAAI'07 Proceedings of the 22nd national conference on Artificial intelligence*, 2007, vol. 2, pp. 1377–1382.
- [9] P. Massa and P. Avesani, "Trust metrics on controversial users: Balancing between tyranny of the majority," *Int J Semant Web Inf Syst*, vol. 3, no. 1, pp. 1–21, 2007.
- [10] P. Massa and P. Avesani, "Trust-aware recommender systems," in *Proceedings of the 2007 ACM conference on Recommender systems*, 2007, pp. 17–24.
- [11] J. O'Donovan and B. Smyth, "Trust in recommender systems," in *Proceedings of the 10th international conference on Intelligent user interfaces - IUI '05*, 2005, p. 167.
- [12] Z. Fu-guo and X. Sheng-hua, "Topic-level Trust in Recommender Systems," in *2007 International Conference on Management Science and Engineering*, 2007, pp. 156–161.

- [13]
- [14] T. Bhuiyan, Y. Xu, A. Jøsang, H. Liang, and C. Cox, “Developing trust networks based on user tagging information for recommendation making,” in *Web Information Systems Engineering – WISE 2010 LNCS*, 2010, pp. 357–364.
- [15] S. Ray and A. Mahanti, “Improving Prediction Accuracy in Trust-Aware Recommender Systems,” in the *43rd Hawaii International Conference on System Sciences*, 2010, pp. 1–9.
- [16] M. G. Moghaddam, N. Mustapha, A. Mustapha, N. Mohd Sharef, and A. Elahian, “AgeTrust: A new temporal trust-based collaborative filtering approach- In press,” in the *proceeding of 5th International Conference on Information Science & Applications (ICISA’ 14)*, 2014, pp. 1–4.
- [17] P. Massa, P. Avesani, “Trust Metrics on Controversial User: Balancing Between Tyranny of the Majority and Echo Chambers”. *International Journal on Semantic Web and Information Systems* 3 (1): 39. (2006).