

Reputation Rating Mode and Aggregating Method of Online Reputation Management System*

Guang-xing Song

*School of MBA Education, Yunnan University of Finance and Economics,
Kunming 650221, P.R.China*

Abstract: With the rapid development of electronic commerce, online reputation management systems are of increasing importance in building trust and managing risk. Reputation rating mode and aggregating method are the most crucial parts of a reputation management system. In this paper, we analyze the merits and disadvantages associated with the rating mode and aggregating approach of current reputation management systems, and put forward some suggestions. These suggestions are helpful in improving current reputation management systems and developing new reputation management systems.

Keywords: Electronic Commerce; Reputation Management System; Rating; Aggregating

1. Introduction

In electronic marketplace, sellers and buyers are often unrelated to each other, they may have never met and have no information on each other's credit. Trust may be more important in electronic commerce than in traditional business due to the spatial and temporal separation imposed between buyers and sellers. In general, sellers have more information about the quality and the cost of products or services than buyers do. Information asymmetry often results in sellers' fraudulent behaviors and increase of trading cost. Meanwhile, the separation of payment and delivery in most of online transactions may lead to the opportunistic behaviors of sellers or buyers. Therefore, both buyers and sellers face more risk in online transaction than in conventional face-to-face transaction.

A variety of governance systems have emerged to deal with the increased risk associated with online transactions. Reputation also serves as a governance mechanism capable of assuring trustworthy behavior [2]. When relying on reputation as an intermediary, additional costs to the buyer are not incurred, and this makes reputation an important mechanism for reducing the risk associated with online trades [11].

Although widely used, current reputation management systems have some drawbacks. R.A. Malaga [9] outlines six main problems with them and put forwards corresponding solutions. C. Dellarocas [4] proposes two classes of reputation system fraud, i.e., unfair rating and discriminatory seller behavior, and suggests eliminating or reducing the effects of these frauds by means of controlled anonymity and cluster filtering.

In fact, most problems from which current online reputation management systems might suffer are related to rating mode and aggregating approach. In this paper, we focus our

* This work is supported by National Natural Science Foundation of China, Grant No. 70572068

research on these two aspects. Section 2 describes the rationale and roles of reputation management system. Section 3 analyzes the reputation rating modes and aggregating methods that most electronic marketplaces use and points out their advantages and drawbacks, meanwhile, puts forwards solutions to related problems. Section 4 summarizes the contributions of the paper and presents the future research directions.

2. Reputation Management System in E-commerce

2.1 Origin of Online Reputation Management System

Reputation is the current assessment of an entity's desirability as established by some external person or group of persons [11]. Operationally, reputation is usually represented as a prediction about likely future behavior. Its predictive power depends on the supposition that past behavior is indicative of future behavior.

In the setting of asymmetric information, reputation is a kind of intangible asset of a firm. In traditional commerce, reputation helps to build trust. If a seller has a good reputation, the customers will trust what he/she advertises, and are more likely to trade with him/her. Under e-commerce environment, due to the trading parties are invisible to each other, buyers would face much uncertainty about the quality of products and the reliability of information. In the absence of an inspection or personal interaction, one good way to assess the quality and reliability of products or information is reputation. Reputation is obviously more important for building and maintaining trust in electronic marketplace than in conventional marketplace.

In traditional trade environment, reputation information is distributed via person-to-person gossip and mass medias. In e-commerce environment, reputation information can be generated and disseminated through the use of online reputation management system. This type of system allows a user to present his/her opinions (feedbacks) in the quality of products, information and service of an enterprise or individual related to a specific transaction. Responses from all users are aggregated into an overall reputation rating, which is available to public and can be used as an indicator of the entity's reputation. At present, various reputation management systems are used in electronic marketplaces (e.g., eBay, Amazon.com, Yahoo!Auction, eachnet and BizRate.com) and online communities (e.g., Slashdot and Epinions). Some reputation systems allow bi-directional ratings (e.g., eBay and Amazon), while others allow only one-directional ratings (e.g., Onsale only allows buyers to rate sellers, and BizRate only allows consumers to rate businesses). Here we primarily discuss reputation management systems used in electronic marketplaces.

2.2 Roles of Reputation Management System

As an important quality signaling and quality control mechanism in e-commerce, a reputation management system mainly plays the following roles:

- Constraining the behaviors of all participants of a transaction and encouraging trustworthy/good behaviors.
- Helping the user to understand his/her counterpart's credit and reliability so as to determine who is trustworthy, consequently, increasing the rate of successful online transaction.

To sum up, a reputation management system can be viewed as a kind of risk management mechanism and effective way to build trust. Implementation of reputation management system can reduce transaction costs (such as the searching costs for an acceptable product) and decrease transaction risk to both sellers and buyer through building trust.

3. Analysis on Reputation Rating Mode and Aggregating Method

When designing a reputation management system, one must pay more attention to how a user rates his/her trading partner and how all the ratings are aggregated into an overall reputation score.

3.1 Reputation Rating Mode

At the end of a transaction, one party can leave a reputation feedback on another with a specified rating mode, such as (-1, 0, 1) scale, 1-5 scale, (-2, -1, 0, 1, 2) scale, or 1-10 scale, etc. The first two scales are widely used in present electronic marketplaces and online communities.

The more the grades of a scale are, the better it reflects the details of reputation of the entity being rated. But a rating scale with more than seven grades is not easy to use. In most cases, a rating scale with five or seven grades is appropriate. A scale permitting negative feedback is in favor of punishing bad behavior, thus effectively constraints opportunistic behaviors.

It is more convenient for users to post feedbacks in the form of linguistic variable instead of numerical values. A linguistic variable is a variable whose values are not numbers but words or phrases in natural or artificial language [14]. For example, if a scale with five grades is used in a reputation system, the feedback will be one of the following linguistic variable values: very unsatisfactory, unsatisfactory, moderate, satisfactory and very satisfactory.

Afterwards, the reputation system itself converts linguistic variable values into numerical values (i.e., scale values). In the five-grade-scale, four kinds of possible relationship between linguistic variable values and scale values are shown in Table 1.

Table 1. Relationship between Linguistic Variable Values and Scale Values

Linguistic variable values		Very unsatisfactory	Unsatisfactory	Moderate	Satisfactory	Very satisfactory
Scale values	(I)	-2	-1	0	1	2
	(II)	1	2	3	4	5
	(III)	0.2	0.4	0.6	0.8	1
	(IV)	0	0.25	0.5	0.75	1

Since “good” and “satisfactory” are fuzzy concepts, we had better convert linguistic variable values into triangular fuzzy numbers rather than trivial single values. A triangular fuzzy number is defined as a triplets (a, b, c) , where $a \leq b \leq c$. One possible transformation for converting linguistic variable values into corresponding triangular fuzzy numbers is shown in Table 2. Considering the fact that the linear combination of several triangular fuzzy numbers is also a triangular fuzzy number [8], we can aggregate all feedbacks via weighed average method, in which the weight is used to reflect the relative importance of different feedbacks. The triangular fuzzy number obtained can be converted into a proper linguistic variable value so that it is easy to understand for users of the system.

Table 2. Relationship between Linguistic Variable Values and Triangular Fuzzy Numbers

Five-grade-scale		Seven-grade-scale	
Linguistic variable values	Triangular fuzzy numbers	Linguistic variable values	Triangular fuzzy numbers
Very poor	(0, 0, 0)	Very poor	(0, 0, 0)
Poor	(0, 0.25, 0.5)	Poor	(0, 0.166, 0.332)

Medium	(0.25, 0.5, 0.75)	Medium poor	(0.166, 0.332, 0.5)
Good	(0.5, 0.75, 1)	Moderate	(0.332, 0.5, 0.668)
Very good	(1, 1, 1)	Medium good	(0.5, 0.668, 0.834)
		Good	(0.668, 0.834, 1)
		Very good	(1, 1, 1)

In general, the reputation of an entity should be conceived as multidimensional value [13]. An individual may enjoy a very high reputation for his/her expertise in one domain, while having a low reputation in another. For instance, a person/firm may have a reputation for selling quality goods, but a reputation for high price and for late delivery of the goods. A user can evaluate this person/firm's reputation from many aspects, and integrate these evaluations into an overall reputation rating score. In such cases, above approach involving linguistic variable and triangular fuzzy number can be also used to aggregate multidimensional reputation ratings into a comprehensive reputation rating.

3.2 Reputation Aggregating Methods

Reputation aggregating methods mainly include summation and average. Each of them has advantages and disadvantages.

(1) *Summation*. It means that the overall reputation score of an entity is obtained by summing up all the rating scores the entity has received.

(2) *Average*. It means that the overall reputation score of an entity is a quotient of the total amount of reputation points and the number of ratings.

The method of summation is not suitable to the reputation systems that permit negative feedback. For example, a user who has received 10 positive feedbacks will have the same overall reputation score as a person who has received 25 positive and 15 negative feedbacks. Moreover, this method can not accurately reflect a user's actual reputation because it does not utilize information about the number of ratings. As an example, we assume that certain reputation management adopts 1-5 scale. The number of ratings two merchants s_1, s_2 have received are 5 and 8, respectively. Rating scores are as follows, s_1 : 5, 5, 4, 5, 4, and s_2 : 3, 2, 4, 5, 4, 3, 1, 1. Total reputation scores of these two merchants make no difference, i.e., 23. If we observe the feedback profile, we will find that s_1 is more reputable than s_2 . In fact, the average score of s_1 and s_2 are 4.6 and 2.875, respectively. Therefore, when the total reputation score is used as the aggregation of all feedback information, it is better to look over the number of ratings or calculate the average score. Sometimes, it is necessary for a user to read the feedback profile and textual comments to gain an accurate picture of the reliability of his/her potential trading partner. Our conclusion is that one should review average reputation score when two entities examined have equal or close total reputation scores.

Although the method of average is suitable to all rating modes, its ability to differentiate entities is weak. For example, if 1-5 scale is used to express reputation ratings and average method is used to aggregate these ratings, the overall reputation scores of all registered users will lie in interval $[0, 5]$ (here we assume the starting reputation score for a newcomer is 0). In an electronic marketplace with millions of registered members (e.g., eBay has more than 10 million registered members and over 4 million items auctioned daily), a user could hardly compare potential trading partners according to their average reputation scores and make a correct choice.

Aggregating method can be described as the following recursive relation:

$$R_n = f(R_{n-1}, r_n), \quad n = 1, 2, \dots$$

Where R_0 is the starting reputation score, R_n is the overall reputation score when the number of ratings an entity received is just n , and r_n is the n th rating an entity received.

For example, the aggregating method at eBay can be represented as : $R_n = R_{n-1} + r_n$, $n = 1, 2, \dots$, where $R_0 = 0$, and r_n is -1 , 0 , or 1 . Moreover, we can deduce: $R_n = R_0 + \sum_{i=1}^n r_i$.

3.3 Assignment of Starting Reputation Score

In a reputation management system, it is crucial to decide the starting reputation score given to a newcomer. If a user's reputation score were lower than that of a beginner, he/she would have an incentive to discard his/her initial identity and start from the beginning. Therefore, although a user's reputation may decrease after a transaction, it will never fall below a newcomer's value. Let R_0 be the newcomer's reputation score given by the system, R_{\min} be the minimum reputation rating allowed by the system. Then they must satisfy the following condition: $R_0 \leq R_{\min}$.

At eBay, the starting reputation score of each newcomer is zero, and the minimum reputation rating allowed by the system is -1 . They do not meet above condition. Although the system stipulates that participants with overall score of -4 or lower are banned from future transactions on eBay, a participant would change his/her identity as long as his/her reputation score is negative.

3.4 Other Drawbacks of Current Reputation Systems and Corresponding Solutions

The reputation management system of eBay can't effectively help buyers estimate the product quality of sellers because it does not publicize the total number of transactions of a seller. Let N be the total number of sales transactions of a given seller in most recent time window (e.g., six month at eBay). A potential buyer estimates an item's quality according to the item's advertised quality and the seller's feedback profile. Dellarocas C. [5] has shown that N is essential for forming reliable quality estimation. But since eBay does not currently publish the number of transactions that have received no rating (at eBay, raters who would submit "neutral" rating can be subsumed into the set of raters who don't submit any rating at all), N is actually not available to buyers at eBay. Therefore, N should be open to users so as to improve the reputation system of eBay.

When a trade is completed, one may have no incentive to rate his/her counterpart, although just spending a few minutes filling out an online form. This problem can be solved through paying the users who provide feedbacks. Another solution is to build incentives into the reputation equation by giving a small increase in reputation whenever a user provides reputation feedback to others [9].

For most reputation management systems, it is very difficult to elicit negative feedback. For example, ebay tells us: "if you were treated poorly, try to resolve the problem first by contacting the other person. Most problems can be corrected by improving communication. If things are still not resolved, consider leaving a negative comment." Amazon.com Auctions and Amazon.com zShop also recommend raters negotiating with the sellers before they post critical feedback. Therefore, only really bad behavior is reported. In addition, even if poorly treated, users might post neutral feedback because of fear of retaliatory or just a desire to avoid further unpleasant interactions. Researches on the mechanism to elicit fair feedback are needed.

When calculating overall reputation scores, almost all current reputation management systems use each individual's all feedback information. If summation is the aggregating method, a veteran user behaving well continuously will get a very high reputation score in a

busy electronic marketplace like eBay. For newcomers, their reputation scores are of great disparity compared with those of veteran users. Thus it is very difficult for newcomers to trade because many users do not want to deal with individuals with lower reputation scores. Therefore, it is more reasonable to consider the most recent ratings rather than all rating information. For the entities that behave well or bad all the time, these two manners show little difference. Aggregating with the most recent ratings has the following advantages:

- It is fair for newcomers and it allows a person with a bad reputation to change.
- It can prevent those people who have higher reputation score from behaving opportunistically for a short period while not obviously affecting their overall reputation scores.

Suppose M to be the threshold, the equation of computing overall reputation score according to the latest M ratings can be described as:

$$R_n = \begin{cases} g(r_1, r_2, \dots, r_n), & n \leq M \\ g(r_{n-M+1}, r_{n-M+2}, \dots, r_n), & n > M \end{cases}$$

Where g is a function that represents aggregating method.

4. Conclusion

This paper studies on the rating mode and aggregating method of an online reputation management system. Advantages and disadvantages regarding the current reputation management systems are discussed in detail. Primarily, we recommend using linguistic variable to express reputation ratings, and converting them into triangular fuzzy numbers so as to aggregate them into an overall reputation score more precisely. Some other new ideas and suggestions are given. Our research is helpful to improve present reputation systems and design new ones, as well as for users to make good use of these systems.

When average method is used to aggregate reputation ratings, weighted average method is more reasonable than simple average method. The weight embodies relative importance of a feedback by considering the value of a transaction and the reputation of the rater. The larger the value of a transaction or the reputation score of the rater, the bigger the weight should be. Moreover, more recent feedback should be given higher weight. Future research is needed to properly determine the weight reflects the transaction value and the rater's reputation, as well as the time at which the rating is submitted.

References:

- [1] Amazon.com <http://www.amazon.com>
- [2] Barney J., Hansen M. Trustworthiness as a Source of Competitive Advantage. *Strategic Management Journal*, 1994,15: 175~190
- [3] BizRate.com <http://www.bizrate.com>
- [4] Dellarocas C. Immunizing Online Reputation Reporting Systems Against Unfair Ratings and Discriminatory Behavior. *Proceedings of the 2nd ACM Conference on Electronic Commerce*, Minneapolis, USA, Oct. 2000: 150~157
- [5] Dellarocas C. Analyzing the Economic Efficiency of eBay-like Online Reputation Reporting Mechanisms. *Proceedings of the 3rd ACM Conference on Electronic Commerce*, Tampa, Florida, USA, Oct. 2001: 171~179
- [6] eBay <http://www.ebay.com>
- [7] eachnet <http://www.eachnet.com>

- [8] Kaufman A., Gupta M.M. Introduction to Fuzzy Arithmetic — Theory and Applications. New York: Van Nostrand Reinhold Company Inc., 1985: 68~75
- [9] Malaga R.A. Web-based Reputation Management Systems: Problems and Suggested Solutions. Electronic Commerce Research, 2001,1: 403~417
- [10] Resnick P., Zeckhauser R., Friedman E., Kuwabara K. Reputation Systems. Communications of the ACM, 2002,43(12): 45~48
- [11] Standifird S. Reputation and e-commerce: eBay Auction and the Asymmetrical Impact of Positive and Negative Rating. Journal of Management, 2001,27(3): 279~296
- [12] Yahoo!Auction <http://auctions.yahoo.com>
- [13] Zacharia G., Moukas A., and Maes P. Collaborative Reputation Mechanisms in Electronic Marketplaces. Decision Support System, 2000,29(4): 371~388
- [14] Zadeh L.A. The Concept of a Linguistic Variable and Its Application to Approximate Reasoning, Part I. Information Science, 1975(8): 199~249