SepaRating: An Approach to Reputation Computation Based on Rating Separation in e-Marketplace

Hyun-Kyo Oh, Yoohan Noh, Sang-Wook Kim*
Department of Electronics and Computer
Engineering, Hanyang University, Seoul, Korea
rapkyo, iamny,
wook@agape.hanyang.ac.kr

Sunju Park
School of Business, Yonsei University, Seoul,
Korea
boxenju@yonsei.ac.kr

ABSTRACT

This paper proposes *SepaRating*, a novel mechanism that separates a buyer's rating on a transaction into two kinds of scores: *seller's score* and *item's score*. SepaRating provides the reputation of sellers correctly based on the seller's score by repetitive separations, which helps potential buyers to find more reliable sellers. We verify the effectiveness of SepaRating via a series of experiments.

Categories and Subject Descriptors

H.2.8 [**DATABASE MANAGEMENT**]: Database applications—*Data mining*

Keywords

Trust, Rating Separation, Reputation System, e-Marketplace

1. INTRODUCTION

The e-marketplace, such as eBay.com, enables a large number of buyers and sellers to trade goods and services on the Internet. For buyers to find more reliable sellers, most e-marketplaces offer a reputation system that provides potential buyers with reputation scores of sellers. Reputation scores of sellers are computed based on the collection of buyers' ratings.

The rating, the buyer's evaluation of a transaction, is based on two criteria. One criterion is based on the capability of the seller, such as the promptness of the answers to customer questions and the safe & fast delivery. The other is based on the quality of an item.

When using ratings to compute reputation scores, however, the existing reputation systems [1, 2, 3] do not distinguish these two evaluation criteria. We claim that the reputation system needs to separate two criteria —whether the rating is about the capability of the seller or about the quality of the item— to provide more precise information about the seller. For example, it is possible that despite of late delivery, the transaction receives a high rating because

---

*Corresponding author

the item is of an excellent quality. On the other hand, it is possible that the transaction receives a high rating because of the seller's detailed answer and fast delivery, even though the item's quality does not meet the buyer's expectation. In both cases, using existing reputation systems, it is impossible to distinguish the cause of the high rating.

In this paper, we propose *SepaRating*, a mechanism that iteratively computes the sellers' reputation by separating the rating into *seller's score* and *item's score*. SepaRating computes the reputation of the seller's capability using the seller's score and helps potential buyers in their search for more reliable sellers.

2. THE PROPOSED MECHANISM

To separate the buyer's rating into seller's score and item's score, we have made an assumption that buyers' evaluation tends to be similar on the quality of the same item. If two buyers have bought the same item from two different sellers and have given different ratings to their transactions, t'his implies that two sellers have different capabilities. We compute the *relative score* among sellers who sell the same item and designate it as the seller's score.

In order to isolate the seller's score from the buyer's rating, we cluster sellers based on the item. The sellers who sell the same item belong to the same cluster. Figure 1 shows an example of seller clustering.
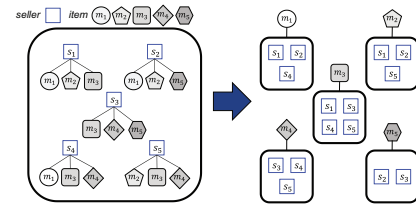
Figure 1: An example of seller clustering based on the item.

Let $\mathbf{S}$ denote the set of sellers and $\mathbf{C}^s$ the set of seller clusters based on the item. Let e_{ik} denote the initial reputation of the i-th seller s_i in the k-th cluster c_k^s. In the beginning, the initial reputation is the average of buyers' ratings on the seller. Let $\mathbf{E}_k$ denote all the initial reputations in c_k^s. The relative score (seller's score) for s_i in c_k^s ($e_{ik}^{'}$) is computed as follows:

$$e_{ik}^{'} = Avg(\mathbf{E}_k \setminus e_{ik}) - e_{ik} \qquad (1)$$

$e_{ik}^{'}$ is defined as the average of the initial reputations of the sellers in the k-th cluster except s_i minus the initial reputation of s_i (e_{ik}). Each seller is assigned with the relative

score within the cluster he belongs to. (Note that a seller may have multiple relative scores.) The i-th seller's reputation, $e_i^{'}$, is computed as follows:

$$e_i^{'} = \frac{1}{|\mathbf{C}^s|} \sum_{c_k^s \in \mathbf{C}^s} e_{ik}^{'} \qquad (2)$$

We call the computation process above as *seller rating separation* (SS). If the sellers are clustered based on their reputations, the item's score can be computed following the process similar to SS. We call the computation process to extract an item's score as *item rating separation* (IS).

For IS, we cluster items based on the sellers' reputations computed by SS. The items with similar[1] sellers' reputations are clustered together. Figure 2 shows an example of item clustering. In Figure 2, we set ϵ as 0.05 such that all items sold by S_1 or S_2 belong to the same cluster.
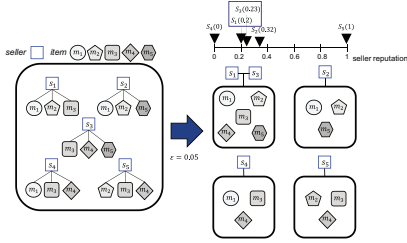
Figure 2: An example of item clustering based on the sellers' reputations.

After performing IS, we proceed with another SS, where the seller clustering is based on item scores similar within ϵ. Because IS changes item scores, and the changed item scores change sellers' scores, and so on, we apply IS and SS successively. The IS and SS processes continues until the sellers' reputations converge.

3. EXPERIMENTS

In existing studies [1, 2, 3], they generated e-marketplace simulation data by considering the simple behavioral patterns between sellers and buyers. In our experiments, we generated more sophisticated simulation data by carefully considering real-world relationships among sellers, buyers, and items. The simulation is divided into days. During a day, each buyer is restricted to trade at most once. The numbers of sellers, buyers, and items can be controlled by parameters.[2]

We generated two sets of e-marketplace data with two different parameter settings. The first one consists of 1,000 items, 500 sellers, 5,000 buyers, and 300 days of a whole transaction period; the second one consists of 2,000 items, 1,000 sellers, 10,000 buyers, and 300 days of a whole transaction period.

As a measure for evaluating the effectiveness of SepaRating (SR), we use the Spearman's rank correlation coefficient, which assesses how closely two sets of rankings agree with each other. As ground truth, we regard the sellers' rankings according to their selling capabilities that are determined in our simulation data.[3] In the experiments, we investigate

---

[1]The term *similar* means the difference between two sellers' reputations is less than ϵ.

[2]Simulation program and simulation data are available at http://goo.gl/zDs1lN

[3]The seller's capability is set according to the normal distribution with an expected value of $\mu = 0.5$ and standard deviation $\sigma = 0.25$.

how much a set of the ground truth rankings and another set of the rankings obtained by the SR agree with each other.

We compare our result with that obtained by a baseline approach (BL), which simply computes a seller's reputation by averaging buyers' ratings on all the transactions in the seller. Figure 3 shows the results. The SR is shown to be more accurate than BL in both data sets. The results indicate that our approach to rating separation is effective in finding true reputation of sellers.
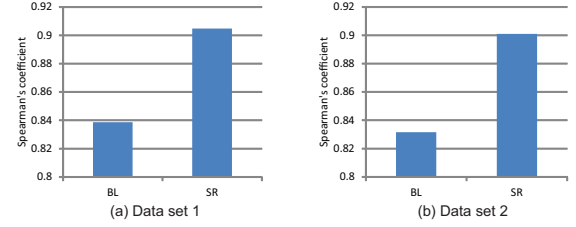
Figure 3: Accuracy with the BL and the SR.

Figure 4 shows the accuracy changes according to the number of iterations in rating separation. We observe the followings: the accuracy of the SR is improved with more iterations; in particular, it is significantly improved at the second iteration; the degree of improvement decreases with further iterations; after the fifth iteration, the accuracy converges to 98 percent and remains stable. The size of clusters tends to increase with a more number of iterations, which leads to higher accuracy.
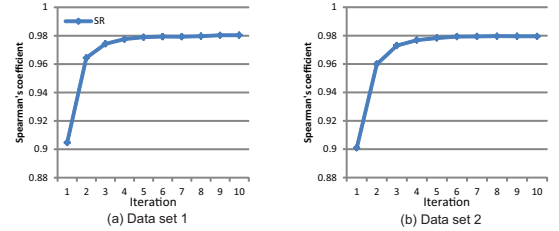
Figure 4: Accuracy changes with the number of iterations.

4. CONCLUSIONS

We have proposed SepaRating for computing the seller's reputation accurately. The main idea is on the separation of a buyer's rating on a transaction into two pieces: seller's score and item's score. Our experimental evaluation shows that the proposed separation idea is fairly effective, thereby helping provide true reputation of sellers to potential buyers.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[1] A. Whitby, A. Jϕsang and J. Indulska. Filtering Out Unfair Ratings in Bayesian Reputation Systems. *The Icfain Journal of Management Research*, Vol. 4, No. 2, pp. 48-64, 2006.

[2] J. Zhang and R. Cohen. Evaluating the Trustworthiness of Advice about Seller Agents in E-marketplaces: A Personalized Approach. *Electronic Commerce Research and Applications*, Vol. 7, No. 3, pp. 330-340, 2008.

[3] S. Liu et al. iCLUB: an Integrated Clustering-based Approach to Improve the Robustness of Reputation Systems. In *Proc. of the 10th Int'l Joint Conf. on Autonomous Agents and Multiagent Systems*, pp. 1151-1152, 2011.