A Clustering Approach to Filtering Unfair Testimonies for Reputation Systems

(Extended Abstract)

Siyuan Liu
School of Electrical and Electronic Engineering
NTU, Singapore
lius0036@ntu.edu.sg

Chunyan Miao
School of Computer Engineering
NTU, Singapore
ASCYMiao@ntu.edu.sg

Yin-Leng Theng
Wee Kim Wee School of Communication and Information
NTU, Singapore
TYLTheng@ntu.edu.sg

Alex C. Kot
School of Electrical and Electronic Engineering
NTU, Singapore
EACKOT@ntu.edu.sg

ABSTRACT

The problem of unfair testimonies remains an open issue in reputation systems for online trading communities. A common attempt is to use binary ratings to model sellers' reputation. However, this attempt leads to that the research of tackling unfair testimonies also focuses on reputation systems using binary ratings. In this extended abstract, we propose a two-stage clustering approach to filter unfair testimonies for reputation systems using multi-nominal ratings. The proposed approach uses clustering to identify unfair testimonies and further contributes to providing buyers a more accurate reputation evaluation regarding the target seller.

Categories and Subject Descriptors

I.2.11 [**ARTIFICIAL INTELLIGENCE**]: Distributed Artificial Intelligence – Intelligent agents, Multiagent systems

General Terms

Algorithms, Experimentation

Keywords

Reputation System, Rating, Testimonies, Clustering

1. INTRODUCTION

Compared to reputation systems using binary ratings for online trading community, reputation systems using multi-nominal ratings provide buyers richer information regarding a seller's trustworthiness [3]. However, reputation systems using multi-nominal ratings suffer from the problem of "unfair testimonies" as reputation systems using binary ratings do. To cope with this "unfair testimonies" problem, we propose a two-stage clustering approach to filter unfair testimonies for reputation systems using multi-nominal ratings.

Cite as: A clustering approach to filtering unfair testimonies for reputation systems (Extended Abstract), Siyuan Liu, Chunyan Miao, Yin-Leng Theng and Alex C. Kot, *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, van der Hoek, Kaminka, Lespérance, Luck and Sen (eds.), May, 10–14, 2010, Toronto, Canada, pp. 1577-1578

2. TWO-STAGE CLUSTERING APPROACH

Suppose there are k rating levels. The testimonies from one particular buyer for a seller can be organized as a rating vector with its length equal to the number of the rating levels. Each dimension value of the vector is the accumulated count of the buyer's past ratings of the corresponding rating level for the seller. In the first stage, we cluster the normalized rating vectors from all witnesses and the buyer (if any) into predefined fixed number of clusters using hierarchical clustering [2]. More specifically, the rating vector from each witness or the buyer is initially regarded as a cluster, and two clusters with the shortest 2-norm distance are merged together to form a new cluster. Then we continue to select two clusters with the shortest 2-norm distance from the new formed cluster and the remaining clusters in last step. We then merge the two selected clusters together to form a new cluster again. The process continues until the predefined cluster number is met (we suggest that the predefined fixed cluster number should be larger than or equal to the number of rating levels to avoid rating vectors with distinct differences being merged together).

In the second stage, the merging process continues until a different merging criteria is met. We first calculate the furthest distance between any two clusters resulted from the first stage. We then merge two clusters with the smallest furthest distance together if the furthest distance between the two clusters is smaller than the predefined distance threshold which is distinguished between the following two scenarios. Firstly, if one of the two clusters selected for merging is a bounder cluster, whose center is very close to the bounder vector as [1,0,. . . ,0] or [0,0,. . . ,0,1], the distance threshold d_1 is set. Secondly, if neither of the two clusters selected for merging is a bounder cluster, another distance threshold d_2 is set. The two scenarios need to be differentiated because a bounder cluster is more likely the cluster including unfair testimonies. Because of the probable inclusion of unfair testimonies within a bounder cluster, d_1 should be set more strictly than d_2 to avoid inaccurately merging unfairly low ratings with fairly low ratings, or inaccurately merging unfairly high ratings with fairly high ratings. The merging process continues until no furthest distance between any two clusters is smaller than the predefined distance threshold. Finally, the testimonies in the cluster including the buyer's rating vector (if any) or the cluster including the majority witnesses' rating vectors (if the buyer has no ratings regarding the seller) are considered as fair testimonies.

3. EXPERIMENTAL RESULTS

The goal of the experiments is to study the unfair testimonies' influence on the evaluation of sellers' reputation and to study the accuracy of the clustering approach in filtering unfair testimonies. In our experiments, there are two types of unfair witnesses studied – ballot-stuffing witnesses and badmouthing witnesses [1], who provide unfairly high ratings and unfairly low ratings respectively.

A trading community is simulated. There is 1 seller S and 100 witnesses. For the scenario where the buyer has personal ratings for S, one more buyer B is simulated. Each witness or B has 1000 transactions with S. There are five rating levels, denoted as [1,2,3,4,5]. Before each simulation, an initial willingness value is generated, taken from the value set of $[0.2, 0.4, 0.6, 0.8, 1.0]$, representing that the initial rating for S is 1, 2, 3, 4 or 5, respectively. Each transaction's rating for S is controlled by S's willingness. The willingness value for each transaction is generated through a normal distribution whose mean is equal to the initial willingness value subtracting 0.1, and standard deviation is equal to 0.2. The mapping between willingness ranges to rating levels is shown in Table 1. The reputation score for each rating level can be estimated as the integration value of $\int_L^H \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$, where $\mu = initialwillingness - 0.1$ and $\sigma = 0.2$, and L and H are the lower range value and higher range value respectively which correspond to the particular willingness range for an initial willingness value. The mapping between reputation score estimation value for each rating level and the initial willingness value is shown in Table 1.

willingness range	rating level	initial willingness				
		0.2	0.4	0.6	0.8	1.0
$(-\infty, 0.2]$	1	0.692	0.309	0.067	0.006	0
(0.2,0.4]	2	0.242	0.383	0.242	0.061	0.006
(0.4,0.6]	3	0.061	0.242	0.383	0.242	0.061
(0.6,0.8]	4	0.006	0.061	0.242	0.383	0.242
$(0.8, \infty)$	5	0	0.006	0.067	0.309	0.692

Table 1: Mapping among willingness ranges, rating levels, expected reputation scores and initial willingness

In our experiments, the accuracy of the clustering approach is measured by comparing the reputation scores estimated by using Dirichlet Reputaiton System (DRS) [3] after clustering filtering, with the reputation scores estimated by using DRS without clustering filtering, and the expected reputation scores listed in Table 1. The fixed number of clusters is 5 for stage 1 clustering stopping. A cluster is regarded as a bounder cluster if its center vector's first or last value is larger than or equal to 0.95. The distance thresholds for stage 2 clustering stopping are set as $d_1 = 0.283$, which is the distance between the normalized rating vector [1,0,0,0,0] and [0.95,0.05,0,0,0], and $d_2 = 0.612$ which is the distance between the normalized rating vector [0,1,0,0,0] and [0.25,0.5,0.25,0,0]. Figure 1 shows the reputation score changes for rating level 1 in the scenarios where the percentage of the ballot-stuffing witnesses is 20%, the percentage of badmouthing witnesses increases from 0% to 40%, and the initial rating is 3, 4 or 5. B has no transactions with S. As shown in Figure 1, the reputation scores with clustering are almost the same as the expected reputation scores.

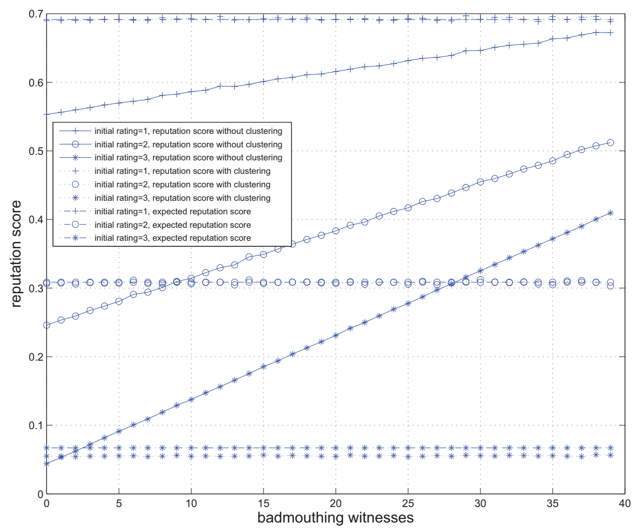

Figure 1: 20% *ballot-stuffing witnesses* and the percentage of *badmouthing witnesses* increases

4. CONCLUSIONS

Reputation systems have contributed much to the success of online trading communities. However, the reliability of reputation systems can easily deteriorate due to the existence of unfair testimonies. To cope with the problem of unfair testimonies, we propose a two-stage clustering approach to filter unfair testimonies for reputation systems using multi-nominal ratings. As the experimental results demonstrate, the approach shows promising results to filter unfair testimonies and provide buyers a more accurate reputation evaluation regarding the seller.

5. REFERENCES

[1] C. Dellarocas. Immunizing online reputation reporting systems against unfair ratings and discriminatory behavior. In *Proceedings of 2nd ACM Conference on Electronic Commerce*, pages 150–157, 2000.

[2] R. O. Duba, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience, 2001.

[3] A. Jøsang and J. Haller. Dirichlet reputation systems. In *Second International Conference on Availability, Reliability and Security (ARES'07)*, pages 112–119, April 2007.