A Subjectivity Alignment Approach for Effective Reputation Computation

Hui Fang, Jie Zhang, Murat Sensoy[†], and Nadia Magnenat Thalmann

School of Computer Engineering, Nanyang Technological University, Singapore
[†]Department of Computing Science, University of Aberdeen, United Kingdom
{hfang1}@e.ntu.edu.sg

Abstract. Current reputation systems simply aggregate numerical ratings provided by buyers, but overlook the buyers' subjectivity difference in evaluating the transactions with a seller. To address this problem, we propose a subjectivity alignment approach for reputation computation (SARC). It first requires the buyers to provide ratings and detailed reviews containing values of objective attributes of the transactions. After that, SARC applies Bayesian learning to model the correlations between each rating level and each objective attribute, and adopts a regression analysis model to learn the weights of the attributes, representing each buyer's subjectivity. Ratings provided by one buyer can then be aligned (converted) for another buyer according to the two buyers' subjectivity. Evaluation results indicate that SARC can more accurately and stably model sellers' reputation than the BLADE and TRAVOS approaches. It is also not much affected by deception from dishonest buyers, and more robust to dynamic environments.

1 Introduction

In open e-marketplaces, it is not possible for a buyer to have experience with every seller. On the other hand, dishonest sellers may advertise perfect deals but never deliver the promise. Therefore, there is a significant risk for buyers when selecting a seller among many alternatives. To address the issue, reputation systems [1] have been proposed, where buyers who previously bought products from a seller share their experience, normally in the form of a numerical rating reflecting the level of satisfaction about the transactions with the seller. These ratings are aggregated to represent the seller's reputation. Other buyers can then rely on the reputation values of sellers to make decisions on which sellers to do business with.

A rating is a subjective evaluation of a seller by a buyer within the context of a specific transaction. Therefore, different ratings could be given for the same transaction by different buyers. Subjectivity difference may come from two sources. First, when the buyer evaluates her satisfaction level with a transaction, she considers each attribute related to that transaction. Although the information about each attribute is *objective*, the evaluation (*i.e.*, satisfactory level) of the attribute value may be subjective and change from user to user. This is referred to as *intra-attribute subjectivity* in this paper. For example, a product with the price of "USD1500" may be *expensive* for buyer a, while *not so expensive* for buyer b. Second, when the buyer assigns a satisfaction level to a transaction, she may consider some attributes of the transaction more heavily than others. This is referred to as *extra-attribute subjectivity*. For example, a buyer with better economic conditions may consider a product's *quality* more heavily, while another buyer with worse economic conditions may concern more about the *price* of the product. The above two aspects together contribute to the subjectivity difference among buyers. Due to the subjectivity difference, it may not be effective if a buyer directly aggregates other buyers' ratings to compute seller reputation. The computed reputation values may then mislead the buyer in selecting business partners.

To effectively address the subjectivity difference problem, we propose a subjectivity alignment approach for reputation computation (SARC). In our approach, each buyer is equipped with an intelligent (buying) agent. At the beginning of her interactions with the reputation system, a buyer a is required to provide her buying agent with both a single rating and a detailed review containing values of the objective attributes of transactions with sellers, such as *price* and *delivery time*, for each of a few transactions. Based on these rating-review pairs, the buying agent applies a proposed Bayesian learning approach to model the correlations between buyer a's each rating level and the value of each objective attribute involved in the transactions. The learned correlation function, which represents buyer a's *intra-attribute subjectivity*, will then be shared with the agents of other buyers. The agent of buyer a also applies a regression analysis model to learn the weight of each attribute for buyer a, representing her *extra-attribute subjectivity*. This information will not be shared with other buyers. After the learning phase, buyer a only needs to provide ratings for her transactions with sellers, not detailed reviews.

When another buyer b shares a new rating of her transaction with a seller, the agent of buyer a will first retrieve a rating level for each attribute of the transaction based on the shared rating and the *intra-attribute subjectivity* of buyer b shared by the agent of b. The rating levels of the attributes will then be aggregated according to buyer a's *extra-attribute subjectivity* learned by the agent of a. In this way, the rating shared by buyer b is aligned to that can be used by buyer a for computing the reputation of the seller.

To evaluate the performance of our SARC approach, we simulate an e-commerce environment involving a number of buyers with different subjectivity in evaluating products and a set of sellers selling products with different attribute values. In addition, buyers' subjectivity may change over time, buyers may also intentionally lie about their evaluation of products, and sellers may change the attribute values of their products. Experimental results confirm that our SARC approach provides sufficiently good performance in a general setting. It can more accurately and stably model sellers' reputation than the representative competing approaches of BLADE [2] and TRAVOS [3]. Our approach is not dramatically affected by deceptive buyers because it treats dishonest buyers as the ones with different subjectivity. It is also more robust to dynamic environments.

2 Related Work

Quite a lot of filtering approaches have been proposed to address the problem of subjectivity difference among buyers and unfair ratings intentionally provided by dishonest buyers to mislead other buyers. For example, some of the approaches filter out the ratings of some buyers (advisors) whose past ratings differ significantly from the ratings of all advisors [4, 5], the ratings of a particular buyer [3, 6, 7], or the ratings of both [8]. These filtering approaches generally suffer from the risk of losing or discounting some important information. In contrast, our approach aligns/converts the ratings of the advisor to those that can be directly used by buyers according to the subjectivity of the buyers and the advisor learned by their agents.

Some other alignment approaches have also been proposed to align advisors' advice about the trustworthiness of sellers. For example, Koster et al. [9] propose a trust alignment approach based on the general framework of Channel Theory. In this approach, each agent computes its own user's trust evaluation patterns based on the interactions towards the same sellers (*i.e.*, shared interactions). Then, the generalized patterns are used to align trust advice provided by advisors. The BLADE approach of Regan et al. [2] applies Bayesian learning to model sellers' properties and the correlations between sellers' properties and buyers' ratings. Once a buyer receives a rating from an advisor, she can infer back the target seller's properties, and then compute the rating of her own towards the seller on the basis of the inferred properties of the target seller. One shortcoming of these alignment approaches is that they ignore the intra-attribute subjectivity difference among buyers. Another shortcoming is that they require the buyer and the advisor to have shared interactions, which may not be the case in an e-commerce environment with a large population of sellers. In addition, these approaches generally offer limited flexibility for buyers to deal with the dynamic behavior of sellers and dynamic subjectivity of advisors. In contrast, our SARC approach aligns each rating provided by an advisor towards a transaction with a seller other than an aggregated trust value of the seller. In this way, it is not affected by sellers' changing behavior. Our SARC approach updates the learned subjectivity of buyers (advisors) in certain interval of time to cope with the possible dynamic subjectivity of advisors. Our SARC approach does not rely on shared interactions. Instead, the agent of each buyer makes use of the ratings and detailed reviews provided by the buyer about her transactions with any sellers, to learn the buyer's intra-attribute and extra-attribute subjectivity.

Collaborative filtering [10] and matrix factorization [11] have been proposed to address the subjectivity difference problem in the domain of recommender systems. However, recommender and reputation systems are different in the sense that reputation systems concern about sellers who may change behavior over time whereas recommender systems concentrate on static products. In addition, in reputation systems, a buyer may have several ratings towards one seller whereas a user has only one rating for one product in recommender systems.

3 The SARC Approach

In an e-marketplace, we denote the set of buyers by $\mathcal{B} = \{b_1, b_2, b_3, \ldots\}$. The set of agents (called buying agents) equipped by corresponding buyers is denoted by $\mathcal{A} = \{a_1, a_2, a_3, \ldots\}$, and the set of sellers are referred to as $\mathcal{S} = \{s_1, s_2, s_3, \ldots\}$. The set of objective attributes for describing a transaction between a buyer and a seller is denoted as $\mathcal{F} = \{f_1, f_2, \ldots, f_m\}$, where m represents the total number of objective attributes. Each rating provided by a buyer for a seller is from a set of predefined discrete rating levels $\mathcal{L} = \{r_1, r_2, \ldots, r_n\}$, where n is the total number of different rating levels (*i.e.,* the granularity of rating scale).

For a buyer $b_i \in \mathcal{B}$, the goal of her buying agent $a_i \in \mathcal{A}$ is to accurately compute the reputation value of a target seller $s_j \in \mathcal{S}$, according to b_i's subjectivity. In order to achieve this goal, the buying agent a_i needs to consider the ratings of other buyers (advisors) that evaluate the satisfaction levels about their past transactions with seller s_j. Due to the possible subjectivity difference between buyer b_i and the advisors, agent a_i also needs to align/convert ratings of each advisor (for example b_k) using our SARC approach.

More specifically, at the beginning of buyer b_i's interactions with the system, agent a_i asks b_i to provide a rating for each of her transactions with a seller (which can be any seller in $\mathcal{S}$). Buying agent a_i also asks b_i to provide detailed review information about each transaction containing the values of the set of objective attributes in $\mathcal{F}$. Based on the provided information (rating-review pairs), agent a_i models a set of correlation evaluation functions (CEFs) for buyer b_i, capturing b_i's *intra-attribute subjectivity*. Each correlation evaluation function is represented by a *Bayesian conditional probability density function* that models the correlation between each rating level and each objective attribute. Thus, for each buyer, the total number of the correlation evaluation functions is equal to $m \times n$.

The learned CEFs of buyers will be shared with each other buyer's agent. For a rating provided by the buyer (advisor) b_k, agent a_i can then derive a rating for each attribute, based on the CEFs shared by b_k's agent a_k and those of buyer b_i's own. Note that what is derived for an attribute is in fact a set of probability values, each of which corresponds to a rating level in $\mathcal{L}$. The rating level with the highest probability will be chosen as the rating for the attribute.

Based on the provided rating-review pairs by b_i, agent a_i also learns the *extra-attribute subjectivity* of buyer b_i, which is represented by a set of weights for corresponding attributes in $\mathcal{F}$. The weight of an attribute is determined by two factors: 1) the probability value of the rating derived earlier; and 2) the importance of the attribute learned using a regression analysis model. These weights will not be shared with other buyers. Once the weights are learned, the aligned rating from that of advisor b_k can be computed as the weighted average of the derived ratings for the attributes.

In the next sections, we will describe in great details how SARC models CEFs based on rating-review pairs, derives a rating for each attribute, learns the weights for attributes, and computes a (aligned) rating by aggregating the derived ratings for attributes, organized as intra-attribute subjectivity alignment and extra-attribute subjectivity alignment.

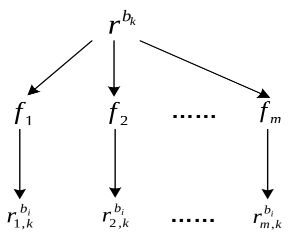

Fig. 1. A Naïve Bayesian Network for Agent a_i of Buyer b_i to Align Buyer b_k's Rating r^{b_k}

3.1 Intra-attribute Subjectivity Alignment

Given a set of rating-review pairs provided by buyer b_i, each of which is for a transaction between b_i and a seller, the rating in a pair indicates b_i's satisfaction level about the corresponding transaction, and the review in the pair is a set of values for the attributes $\mathcal{F}$ of the transaction. Buyer b_i's agent a_i learns the correlation evaluation functions (CEFs) of b_i, each of which is represented by a Bayesian conditional probability density function. Each CEF is the correlation between a rating level and the values of an attribute. More specifically, let us learn $\mathrm{CEF}^{b_i}_{u,v}$, the correlation function between attribute f_u and rating level r_v for buyer b_i, where $1 \leq u \leq m$ and $1 \leq v \leq n$. Buying agent a_i first learns $p^{b_i}(r_v)$ (the probability that buyer b_i provides a rating r_v), $p^{b_i}(f_u)$ (the probability distribution of the values for attribute f_u), and $p^{b_i}(r_v \mid f_u)$ (the conditional probability of rating level r_v given the distribution of the values for attribute f_u). By applying the Bayes' Rule, agent a_i can derive $\mathrm{CEF}^{b_i}_{u,v}$ as the conditional probability distribution of the values for attribute f_u given rating level r_v as follows:

$$\mathrm{CEF}^{b_i}_{u,v} = p^{b_i}(f_u \mid r_v) = \frac{p^{b_i}(r_v \mid f_u) \times p^{b_i}(f_u)}{p^{b_i}(r_v)} \tag{1}$$

In our approach, the agents of buyers share the learned CEFs for their buyers with the agents of other buyers. Suppose that the agent a_k of a buyer b_k shares the learned CEF^{b_k} for b_k with the agent a_i of buyer b_i. For a rating r^{b_k} shared by buyer b_k, agent a_i can then derive a rating level for each attribute in $\mathcal{F}$. We use a Naïve Bayesian Network model to learn the mapping/alignment from r^{b_k} of buyer b_k to the ratings of b_i for the attributes, as illustrated in Figure 1. Although in this model we assume that the attributes are independent given the ratings of buyers, in the next section, we will learn the relative weights of the attributes to capture the dependency among the attributes.

Let us take any $f_u \in \mathcal{F}$ as an example attribute to show how agent a_i derives a rating for attribute f_u. To do so, agent a_i first estimates the conditional probability of a rating level in $\mathcal{L}$ for attribute f_u, given rating r^{b_k} provided by buyer b_k. Take any rating level r_v as an example, agent a_i computes $p^{b_i}(r_{v,f_u}|r^{b_k})$, the conditional probability that buyer b_i will assign the rating level r_{v,f_u} to attribute f_u given the rating r^{b_k} of buyer b_k, as follows:

$$\begin{aligned} p^{b_i}(r_{v,f_u}|r^{b_k}) &= \frac{p^{b_i}(r_v \mid f_u, r^{b_k}) \times p^{b_k}(f_u \mid r^{b_k})}{p^{b_i}(f_u \mid r_v, r^{b_k})} \\ &= \frac{p^{b_i}(r_v \mid f_u) \times p^{b_k}(f_u \mid r^{b_k})}{p^{b_i}(f_u \mid r_v)} \end{aligned} \tag{2}$$

where $p^{b_k}(f_u \mid r^{b_k})$ is learned by agent a_k of buyer b_k using Equation 1 and shared by agent a_k to agent a_i, $p^{b_i}(f_u \mid r_v)$ is learned by a_i itself using Equation 1, and $p^{b_i}(r_v \mid f_u)$ is obtained by agent a_i from the rating-review pairs provided by its buyer b_i. In Equation 2, $p^{b_i}(r_v \mid f_u, r^{b_k})$ is equivalent to $p^{b_i}(r_v \mid f_u)$ and $p^{b_i}(f_u \mid r_v, r^{b_k})$ is equivalent to $p^{b_i}(f_u \mid r_v)$ because buyer b_i provides ratings to corresponding attributes regardless of buyer b_k's ratings. In other words, buyers evaluate transactions independently.

For attribute f_u, agent a_i learns the conditional probability of each rating level $r_v \in \mathcal{L}$ according to Equation 2. The aligned rating of attribute f_u for buyer b_i on the basis of buyer b_k's rating is then determined as the rating level with the highest probability value, as follows:

$$r^{b_i}_{u,k} = \underset{r_v \in \mathcal{L}}{argmax}(p^{b_i}(r_{v,f_u}|r^{b_k})) \tag{3}$$

The aligned ratings for other attributes in $\mathcal{F}$ can also be determined in the same way according to Equations 2 and 3.

3.2 Extra-attribute Subjectivity Alignment

After the ratings of the attributes are obtained, agent a_i of buyer b_i then aggregates the ratings to represent an aligned rating of the rating r^{b_k} shared by buyer b_k. To do this, a_i needs to first determine a weight for each attribute in $\mathcal{F}$ as buyer b_i may concern more about one attribute over another.

The weight of an attribute f_u is determined by two factors. One factor is the confidence C_u about the rating $r^{b_i}_{u,k}$ derived for the attribute f_u using Equations 2 and 3. The confidence can be represented as the conditional probability value of the derived rating, $p^{b_i}(r^{b_i}_{u,k}|r^{b_k})$ estimated using Equation 2. A larger probability value means that it is more probable that the derived rating for attribute f_u should be $r^{b_i}_{u,k}$ according to buyer b_k's rating and the subjectivity of buyers b_i and b_k. In another word, the larger the probability is, the more reliable the derived rating $r^{b_i}_{u,k}$ is. Thus, we have:

$$C_u = p^{b_i}(r^{b_i}_{u,k}|r^{b_k}) \tag{4}$$

Another factor to determine the weight for attribute f_u is the importance I_u of f_u in buyer b_i's view. The importance I_u can be modeled as the coefficient of attribute f_u by a regression analysis model, based on the rating-review pairs provided by b_i. More specifically, given the rating-review pairs, we compute the coefficients for attributes by minimizing the aggregated difference between the true ratings in the rating-review pairs of b_i and the ratings, each of which is predicted for a review by the following equation:

$$r^{b_i}_0 = I_0 + \sum\nolimits_{u=1}^{m} I_u \times V_{f_u} + \varepsilon \tag{5}$$

where $r^{b_i}_0$ is the predicted rating for a review, V_{f_u} is the value of f_u in the review, I_0 is a constant, and ε is residual. So, the coefficients $I = [I_0, I_1, \ldots, I_m]$ can be computed by:

$$I' = (X^{'}X)^{-1}X^{'}Y \tag{6}$$

where if there are c rating-review pairs for buyer b_i in total,

$$X = \begin{bmatrix} 1 & f_{11} & \ldots & f_{m1} \\ 1 & f_{12} & \ldots & f_{m2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & f_{1c} & \ldots & f_{mc} \end{bmatrix}, \quad Y = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_c \end{bmatrix}$$

After the weight (confidence and importance) of each attribute is determined, the aligned rating $r^{b_i}_k$ can be computed as the weighted average of the ratings for attributes derived using Equations 2 and 3, as follows:

$$r^{b_i}_k = \frac{\sum_{u=1}^{m} r^{b_i}_{u,k} \times C_u \times I_u}{\sum_{u=1}^{m} C_u \times I_u} \tag{7}$$

After aligning all ratings shared by all buyers (advisors), the reputation value of seller s_j in the view of b_i can be computed as, for example, the average of the aligned ratings.

4 Experimentation

In this section, we carry out experiments to evaluate the performance of our SARC approach and compare it with some representative competing approaches. We simulate an e-commerce environment involving 50 sellers and 200 buyers. In our simulations, sellers may provide different products. Their products are all different PC configurations with five objective attributes, namely, *Price*, *Speed of CPU*, *Processor Type*, *Graphics Card Type*, and *Hard Drive Size* with ranges presented in Table 1. For each seller, the values of the five attributes of her products are randomly chosen within the ranges.

Buyers may have different subjectivity in evaluating their transactions with (the products of) sellers. We simulate both buyers' intra-attribute subjectivity and extra-attribute subjectivity. To be specific, we assume that a buyer's rating for a transaction with a seller is derived as follows. First, the buyer evaluates each objective attribute according to a specific intrinsic (taste) function. In our experiments, buyers' intra-attribute subjectivity is simulated as an approximate *Gaussian Distribution*. That is, for each attribute, the probability of each rating level given by a buyer is in the form of a normal distribution. Second, the buyer places random weights (in the domain of [0,1]) on different attributes, and computes the weighted average of her evaluations on attributes as a single rating for the transaction. Since buyers can only give ratings under the predefined rating scale in reality, the simulated rating is chosen from the predefined rating scale that is the closest to the weighted average.

Table 1. Product Attributes and Value Ranges

Dimension	Type	Ranges
$Price$	Double	\$100-\$10,000
$Speed\ of\ CPU$	Double	1-10 GHZ
$Processor\ Type$	Char	5 types
$Graphics\ Card\ Type$	Char	2 types
$Hard\ Drive\ Size$	Integer	40-1000GB

In the experiments, we also implement a baseline approach without subjectivity alignment, which computes seller reputation by directly averaging the ratings collected from other buyers. We choose to implement TRAVOS [3], a representative filtering approach (see the Related Work section for details). BLADE [2] is chosen instead of the approach of Koster et al. [9] because they are very similar and the approach of Koster et al. is complicated to implement.

We compare the performance of these approaches with our approach in reputation computation. The performance of an approach is measured as the mean absolute error (MAE) between seller reputation computed for each buyer using the approach, and that using the ratings according to each buyer's own subjectivity (representing the ground truth about seller reputation with respect to the buyer).

To simulate real-world e-commerce environments, we set several important parameters for our simulations, including information availability, dynamic behavior of sellers, dynamic subjectivity of buyers, ratio of liars (dishonest buyers), and granularity of rating scale.

Information availability refers to the amount of available information required by different approaches for subjectivity alignment. Two types of information are needed by

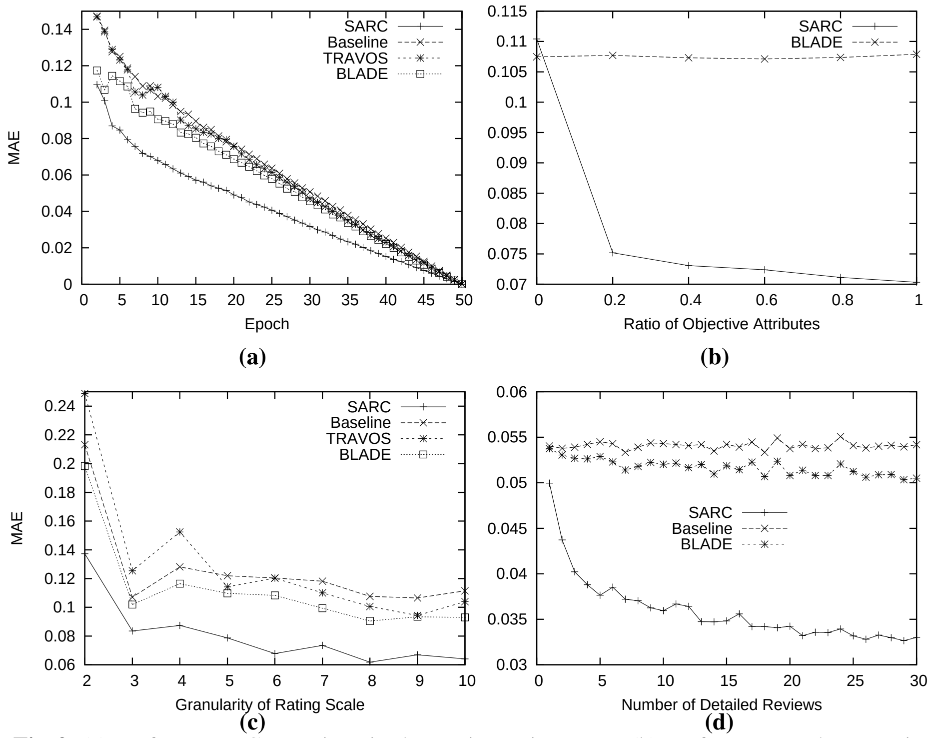

Fig. 2. (a) Performance Comparison in the Basic Environment; (b) Performance When Varying Ratio of Objective Attributes; (c) Performance When Varying Granularity of Rating Scale; (d) Performance When Varying Number of Detailed Reviews

our approach. One is the detailed reviews describing the objective attributes of transactions between buyers and sellers. This information is used by our approach to model the correlation evaluation functions (CEFs) and the importance of the attributes for buyers. We vary the *number of detailed reviews* (N_r) to see how the performance of our approach is affected by this parameter. Another type of information contributing to our approach is the number of objective attributes. In reality, some attributes (*e.g. appearance*) may not be objective. The total number of objective attributes in our simulations may thus be less than 5. We vary the *ratio of objective attributes* (R_{obj}) to be 0%, 20%, 40%, 60%, 80% and 100%, to see how much the performance of our approach will be affected. One type of information required by the BLADE approach is shared interactions where buyers and advisors have interacted with some same sellers. We vary the *ratio of shared interactions* (R_i) to see how this parameter affects the performance of BLADE.

We also set the parameter P_{seller} to capture the *dynamic behavior of sellers*. In real-world e-commerce environments, sellers may change their behavior over time. For example, they may provide products of high quality at first, but those of low quality after earning enough reputation. In our experiments, dynamic behavior of sellers is

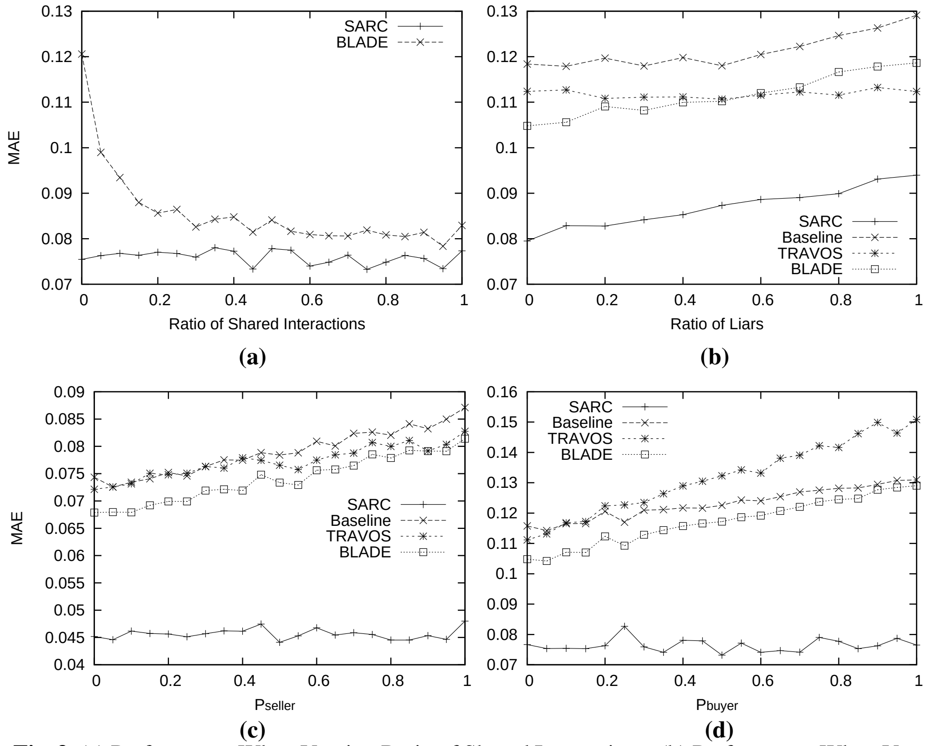

Fig. 3. (a) Performance When Varying Ratio of Shared Interactions; (b) Performance When Varying Ratio of Lying Buyers; (c) Performance for Sellers' Changing Behavior; (d) Performance for Buyers' Changing Subjectivity

simulated by changing the quality of their products (*i.e.* the values of a subset of the objective attributes in Table 1).

Buyers may also adjust their subjectivity over time. Dynamic subjectivity of buyers (P_{buyer}) is captured in their rating procedure by adjusting *intra-attribute subjectivity*, or *extra-attribute subjectivity*, or both.

Ratio of liars (R_{liar}) is adopted to reflect the deception problem in real e-marketplaces where some buyers may lie about their experience with sellers. Following the work of [6, 4, 3], we also simulate the complementary lying behavior where if a true rating to a seller is r in the scale of $[0, 1]$, the liar will modify the rating as $1 - r$.

Granularity of rating scale (G_{scale}) refers to the number of rating levels. It may be different for different reputation systems. In our experiments, we will study the effect of the granularity of rating scale by varying G_{scale} from 2 to 10.

We vary the values of the above parameters to simulate basic, deceptive and dynamic environments, respectively.

Basic Environment We first simulate a basic environment without any variation of the parameters (*i.e.*, $R_{liar} = 0$, $P_{seller} = 0$, $P_{buyer} = 0$), and compare the performance of our approach and that of the three competing approaches, including the baseline approach, TRAVOS and BLADE. We compute their mean absolute error (MAE) values for computing the reputation of sellers in different epoches. In each epoch, each buyer interacts with one seller in the marketplace. From the results shown in Figure 2(a), we can see that our approach performs consistently the best no matter whether buyers have more or less experience with sellers. Because both TRAVOS and BLADE require shared interactions, their performance is limited. Both TRAVOS and BLADE perform slightly better than the baseline approach. The performance difference between the different approaches is reduced when buyers have more experience with sellers in the marketplace.

Based on the basic environment, we then vary some parameters to examine their effects. We first examine how the ratio of objective attributes R_{obj} affects our SARC approach. We vary R_{obj} from 0% to 100% for our SARC approach, while keep R_{obj} to be 100% for BLADE. As shown in Figure 2(b), SARC performs slightly worse than BLADE when there are no objective attributes. However, it performs better than BLADE when there are more than 20% of objective attributes. The performance of SARC consistently increases as the ratio of objective attributes increases. But, the increment becomes smaller when $R_{obj} \geq 20\%$.

The larger the granularity of the rating scale (G_{scale}) is, the easier to learn buyers' subjectivity because buyers' subjectivity can be better captured by the larger granularity of the rating scale. This trend is verified by our experiment. In Figure 2(c), we plot the MAE results of the four approaches when varying G_{scale} from 2 to 10. The figure shows that the performance of SARC is significantly greater than the baseline approach, TRAVOS and BLADE. On average, the performance of SARC improves as G_{scale} increases.

We also vary the number of detailed reviews (N_r) provided by buyers from 1 to 30. We try to figure out a reasonable N_r for SARC. As shown in Figure 2(d), when N_r increases from 1 to 5, the performance of SARC increases significantly. While N_r is larger than 5, as the increase of N_r, the performance of SARC also increases, but in a much smaller degree. This is simply because SARC requires only a few detailed reviews to learn buyers' subjectivity well. After that, any additional information leads to only small improvement. Thus, we can choose 6 as the acceptable minimum N_r. Besides, SARC performs better than the baseline approach and BLADE in all the cases for N_r.

BLADE requires shared interactions in order to learn buyers' subjectivity. However, in real e-marketplaces, shared interactions are generally very sparse. In this experiment, we fix the number of past interactions for each buyer, but vary the ratio of shared interactions (R_i) from 0% to 100%. For each ratio value, MAE is computed as the average of five repeated runs. Figure 3(a) indicates that BLADE performs significantly worse than SARC when R_i is in the range from 0% to 30%. The performance of BLADE increases with the increase of R_i.

Deceptive Environment In this experiment, we examine the effect of deception (buyers lying about their past experience) on different approaches. We vary the ratio of liars (R_{liar}) from 0% to 100%, and plot the MAE results of different approaches in Figure 3(b). We can see that the performance of TRAVOS does not decrease much as R_{liar} increases. Our SARC performs much better than the other three models for any R_{liar}. It is not dramatically affected by lying buyers because SARC learns a buyer's subjectivity from the buyer's own past experience and treats lying buyers as buyers with different subjectivity. When R_{liar} is larger than 0.5, BLADE performs worse than TRAVOS, but consistently better than the baseline approach. Note that in the environment where most buyers are liars, the performances of other models are not so bad. This is mainly because buyers have different subjectivity in our simulations. The effect of buyers' lying behavior may be reduced by the subjectivity difference among buyers, and vice versa.

Dynamic Environment In this experiment, we simulate the environment where sellers may change the quality of their provided products in their transactions with buyers. We define a predefined parameter, P_{seller}, to represent the probability that each seller may vary the values of the five attributes of her provided products. We assume that sellers only change their behavior once in the marketplace. Once their behavior is changed, they will keep the behavior. P_{seller} is ranged from 0 to 1 and increased by 0.05 in our experiment. The MAE results for SARC and other three approaches are plotted in Figure 3(c), which demonstrates that the performance of SARC is not sensitive to the dynamic behavior of sellers, and it performs almost consistently in all cases, while the performance of Baseline, TRAVOS and BLADE gets worse as the increase of P_{seller}. The main reason is that SARC models the rating behavior (subjectivity) of each buyer from the buyer's own experience, which is independent of sellers' behavior change. For TRAVOS and BLADE, they rely on past shared interactions between the buyer and advisors, and these shared interactions may not be suitable source information used for aligning the buyer's subjectivity due to the possible behavior change of sellers in the shared interactions. For example, for a buyer and an advisor with the same subjectivity, if they interact with a seller in different time periods where the seller has changed behavior, TRAVOS may incorrectly treat the advisor as a liar and BLADE may incorrectly conclude that the buyer and the advisor have different subjectivity.

In a marketplace, buyers may also change or adjust their subjectivity after several interactions with sellers. In this experiment, we assume that buyers will change their subjectivity with a certain predefined probability, P_{buyer}. Same as the previous experiment, buyers only change their subjectivity once in the marketplace and then keep their changed subjectivity in the following interactions with sellers. Figure 3(d) shows that the performance of SARC is not affected by buyers' dynamic subjectivity. In SARC, buying agents can update the learned subjectivity of buyers by acquiring their buyers' own recent experience, which provides flexibility to deal with their buyers' dynamic subjectivity. The performance of BLADE becomes almost equivalent to that of Baseline as P_{buyer} increases, and is consistently lower than SARC. In BLADE, once a buyer's subjectivity is changed, her buying agent cannot align ratings from advisors effectively because new shared interactions between the buyer and advisors are needed. TRAVOS performs worse than Baseline as P_{buyer} increases because the learned results of advisors become misleading after they change subjectivity.

5 Conclusion and Future Work

In this paper, we proposed a subjectivity alignment approach for reputation computation, SARC, to address the subjectivity difference problem. In SARC, buyers' subjectivity is learned based on the ratings and detailed reviews they provide about the objective attributes of their transactions with sellers. More specifically, SARC separately learns the *intra-attribute subjectivity* and *extra-attribute subjectivity* of buyers. Buyers' *intra-attribute subjectivity* is modeled using Bayesian learning. Their *extra-attribute subjectivity* is learned using a regression analysis model. We also conducted various experiments to compare the performance of our approach with that of other three competing models, including the baseline approach, TRAVOS and BLADE. Experimental results demonstrate that: 1) SARC performs better than the other three approaches, and can more accurately and stably model sellers' reputation; 2) SARC is capable of coping with environments with deception and dynamic buyer and seller behavior; 3) the requirement of detailed reviews and objective attributes is not very restrictive.

For future work, we will conduct more experiments on real data obtained from, for example, eBay (www.ebay.com) to further validate the effectiveness of our approach.

References

1. Resnick, P., Zeckhauser, R., Friedman, E., Kuwabara, K.: Reputation systems. Advances in Applied Microeconomics **43**(12) (2000) 45–48
2. Regan, K., Poupart, P., Cohen, R.: Bayesian reputation modeling in e-marketplaces sensitive to subjectivity, deception and change. In: Proceedings of the 21st National Conference on Artificial Intelligence (AAAI). (2006)
3. Teacy, W.T.L., Patel, J., Jennings, N.R., Luck, M.: TRAVOS: Trust and reputation in the context of inaccurate information sources. Autonomous Agents and Multi-Agent Systems **12**(2) (2006) 183–198
4. Whitby, A., Jøang, A., Indulska, J.: Filtering out unfair ratings in bayesian reputation systems. In: Proceedings of the 3rd International Joint Conference on Autonomous Agenst Systems Workshop on Trust in Agent Societies (TRUST). (2004)
5. Brennan, M., Wrazien, S., Greenstadt, R.: Using machine learning to augment collaborative filtering of community discussions. In: Proceedings of the 9th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS). (2010)
6. Yu, B., Singh, M.P.: Detecting deception in reputation management. In: Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS). (2003) 73–80
7. Noorian, Z., Marsh, S., Fleming, M.: Multi-layered cognitive filtering by behavioural modelling. In: Proceedings of the 10th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS). (2011)
8. Zhang, J., Cohen, R.: Evaluating the trustworthiness of advice about selling agents in e-marketplaces: A personalized approach. Electronic Commerce Research and Applications **7**(3) (2008) 330–340
9. Koster, A., Sabater-Mir, J., Schorlemmer, M.: Inductively generated trust alignments based on shared interactions. In: Proceedings of 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS). (2010)
10. Li, R., Li, B., Jin, C., Xue, X., Zhu, X.: Tracking user-preference varying speed in collaborative filtering. In: Proceedings of the 25th National Conference on Artificial Intelligence (AAAI). (2011)
11. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. IEEE Computer **42**(8) (2009) 30–37